

logistic-regression - Newton

February 25, 2022

Name: Rahul Govindkumar

```
[74]: import numpy as np
import matplotlib.pyplot as plt
from matplotlib import style
style.use('ggplot')
%matplotlib inline
import re
import pandas as pd
```

Numpy is library for scientific computing in Python. It has efficient implementation of n-dimensional array (tensor) manipulations, which is useful for machine learning applications.

```
[75]: import numpy as np
```

We can convert a list into numpy array (tensor)

```
[76]: b = [[1, 2, 4], [2, 6, 9]]
a = np.array(b)
a
```

```
[76]: array([[1, 2, 4],
          [2, 6, 9]])
```

We can check the dimensions of the array

```
[77]: a.shape
```

```
[77]: (2, 3)
```

We can apply simple arithmetic operation on all element of a tensor

```
[78]: a * 3
```

```
[78]: array([[ 3,  6, 12],
          [ 6, 18, 27]])
```

You can transpose a tensor

```
[79]: print(a.T.shape)
a.T
```

(3, 2)

```
[79]: array([[1, 2],
           [2, 6],
           [4, 9]])
```

You can apply aggregate functions on the whole tensor

```
[80]: np.sum(a)
```

```
[80]: 24
```

or on one dimension of it

```
[81]: np.sum(a, axis=0)
```

```
[81]: array([ 3,  8, 13])
```

```
[82]: np.sum(a, axis=1)
```

```
[82]: array([ 7, 17])
```

We can do element-wise arithmetic operation on two tensors (of the same size)

```
[83]: c1 = np.array([[1, 2, 4], [2, 6, 9]])
      c2 = np.array([[2, 3, 5], [1, 2, 1]])
      c1 * c2
```

```
[83]: array([[ 2,  6, 20],
           [ 2, 12,  9]])
```

If you want to multiply all columns of a tensor by vector (for example if you want to multiply all data features by their labels) you need a trick. This multiplication shows up in calculating the gradients.

```
[84]: a = np.array([[1, 2, 4], [2, 6, 9]])
      b = np.array([1, -1])
      print(a)
      print(b)
```

```
[[1 2 4]
 [2 6 9]]
[ 1 -1]
```

Here we want to multiply the first row of a by 1 and the second row of a by -1. Simply multiplying a by b does not work because a and b do not have the same dimension

```
[85]: # a * b
```

To do this multiplication we first have to assume b has one column and then repeat the column of b with the number of columns in a. We use tile function to do that

```
[86]: b_repeat = np.tile(b, (a.shape[1],1)).T
      print(b_repeat.shape)
      b_repeat
```

(2, 3)

```
[86]: array([[ 1,  1,  1],
            [-1, -1, -1]])
```

Now we can multiply each column of a by b:

```
[87]: a * b_repeat
```

```
[87]: array([[ 1,  2,  4],
            [-2, -6, -9]])
```

You can create initial random vector using numpy (using $N(0,1)$):

```
[88]: mu = 0 #mean
      sigma = 1 #standard deviation
      r = np.random.normal(mu,sigma, 1000) #draws 1000 samples from a normal
      ↪distribution
```

We can apply functions on tensors

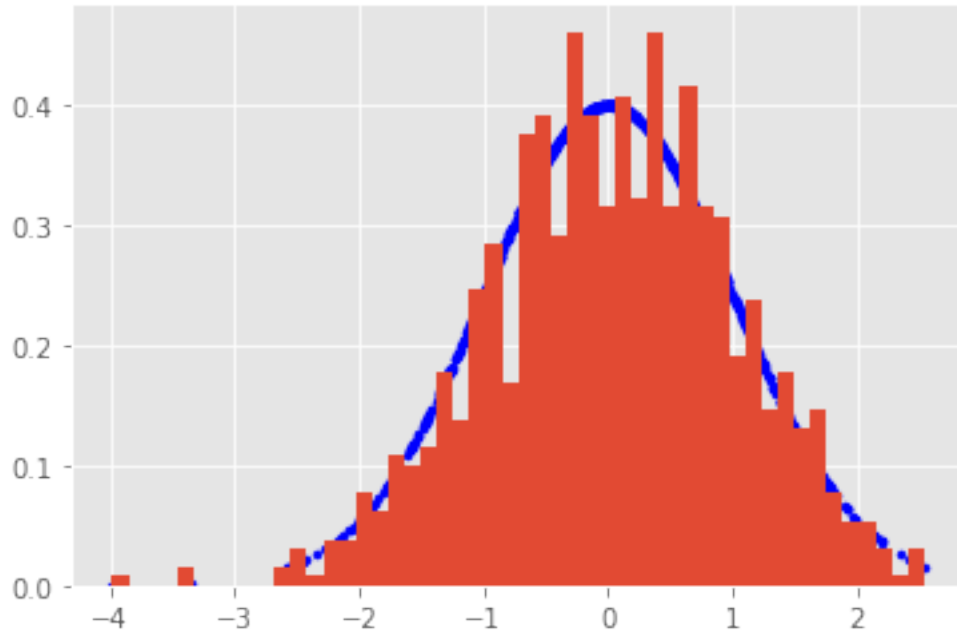
```
[89]: #implementation of Normal distribution
      def normal(x, mu, sigma):
          return np.exp( -0.5 * ((x-mu)/sigma)**2)/np.sqrt(2.0*np.pi*sigma**2)

      #probability of samples on the Normal distribution
      probabilities = normal(r, mu, sigma)
```

Numpy has useful APIs for analysis. Here we plot the histogram of samples and also plot the probabilities to see if the samples follow the normal distribution.

```
[90]: counts, bins = np.histogram(r,50,density=True)
      plt.hist(bins[:-1], bins, weights=counts)
      plt.scatter(r, probabilities, c='b', marker='.')
```

```
[90]: <matplotlib.collections.PathCollection at 0x1b78000cfd0>
```



```
[91]: def read_data(filename):
    f = open(filename, 'r')
    p = re.compile(',')
    xdata = []
    ydata = []
    header = f.readline().strip()
    varnames = p.split(header)
    namehash = {}
    for l in f:
        li = p.split(l.strip())
        xdata.append([float(x) for x in li[:-1]])
        ydata.append(float(li[-1]))

    return np.array(xdata), np.array(ydata)
```

Assuming our data is x is available in numpy we use numpy to implement logistic regression

```
[92]: (xtrain_whole, ytrain_whole) = read_data('datasets/spambase-train.csv')
(xtest, ytest) = read_data('datasets/spambase-test.csv')
```

```
[93]: print("The shape of xtrain:", xtrain_whole.shape)
print("The shape of ytrain:", ytrain_whole.shape)
print("The shape of xtest:", xtest.shape)
print("The shape of ytest:", ytest.shape)
```

```
The shape of xtrain: (3601, 54)
The shape of ytrain: (3601,)
```

The shape of xtest: (1000, 54)

The shape of ytest: (1000,)

before training make we normalize the input data (features)

```
[94]: xmean = np.mean(xtrain_whole, axis=0)
      xstd = np.std(xtrain_whole, axis=0)
      xtrain_normal_whole = (xtrain_whole-xmean) / xstd
      xtest_normal = (xtest-xmean) / xstd
```

We need to create a validation set. We create an array of indecies and permute it.

```
[95]: premute_indicies = np.random.permutation(np.arange(xtrain_whole.shape[0]))
```

We keep the first 2600 data points as the training data and rest as the validation data

```
[96]: xtrain_normal = xtrain_normal_whole[premute_indicies[:2600]]
      ytrain = ytrain_whole[premute_indicies[:2600]]
      xval_normal = xtrain_normal_whole[premute_indicies[2600:]]
      yval = ytrain_whole[premute_indicies[2600:]]
```

Initiallizing the weights and bias with random values from $N(0,1)$

```
[97]: weights = np.random.normal(0, 1, xtrain_normal.shape[1]);
      bias = np.random.normal(0,1,1)
```

```
[98]: #the sigmoid function
      def sigmoid(v):
          #return np.exp(-np.logaddexp(0, -v)) #numerically stable implementation of
          ↪sigmoid function
          return 1.0 / (1+np.exp(-v))
```

We can use dot-product from numpy to calculate the margin and pass it to the sigmoid function

```
[99]: #w: weight vector (numpy array of size n)
      #b: numpy array of size 1
      #returns p(y=1|x, w, b)
      def prob(x, w, b):
          return sigmoid(np.dot(x,w) + b);
```

You can also calculate l_2 penalty using linalg library of numpy

```
[100]: np.linalg.norm(weights)
```

```
[100]: 7.278174103965956
```

$$\text{Cross Entropy Loss} = -\frac{1}{|D|} \left[\sum_{(y^i, \mathbf{x}^i) \in \mathcal{D}} y^i \log p(y = 1 | \mathbf{x}^i; \mathbf{w}, b) + (1 - y^i) \log(1 - p(y = 1 | \mathbf{x}^i; \mathbf{w}, b)) \right] + \frac{\lambda}{2} \|\mathbf{w}\|^2$$

```
[101]: #w: weight vector (numpy array of size n)
#y_prob: p(y/x, w, b)
#y_true: class variable data
#lambda_: l2 penalty coefficient
#returns the cross entropy loss
def loss(w, y_prob, y_true, lambda_):

    n_sample=len(w)

    cross_entropy_Loss=-(1/abs(n_sample))*np.sum((y_true * np.log(y_prob + 0.
↪00000000000000000001)) + (1-y_true)*np.log(1-y_prob + 0.00000000000000000001))
↪+ lambda_*0.5*np.linalg.norm(w)

    return cross_entropy_Loss
```

```
[102]: #x: input variables (data of size m x n with m data point and n features)
#w: weight vector (numpy array of size n)
#y_prob: p(y/x, w, b)
#y_true: class variable data
#lambda_: l2 penalty coefficient
#returns tuple of gradient w.r.t w and w.r.t to bias

def grad_w_b(x, w, y_prob, y_true, lambda_):

    n_samples=len(x)
    # grad_w = np.zeros(w.shape)
    # grad_b = 0.0

    grad_w = (1/n_samples) * np.dot(x.T, (y_prob - y_true))
    grad_b = (1/n_samples) * np.sum(y_prob -y_true)

    return (grad_w,grad_b)
```

```
[103]: #x: input variables (data of size m x n with m data point and n features)
#w: weight vector (numpy array of size n)
#y_prob: p(y/x, w, b)
#y_true: class variable data
#lambda_: l2 penalty coefficient
#returns tuple of gradient w.r.t w and w.r.t to bias

def second_grad_w(x, w, y_prob, y_true, lambda_):
    n_sample = len(x)
    second_grad_w = np.zeros(w.shape)
    second_grad_b = 0.0
```

```

        second_grad_w = (1/n_sample) * -1 * np.dot(np.dot(x.T,x), np.
        ↪dot(y_prob,(1-y_prob) ))
        second_grad_b = y_prob - y_true
        return (second_grad_w,second_grad_b)

```

```
[104]: cost_list=[]
```

```

[105]: #lambda_ is the coeffienct of l2 norm penalty
#learning_rate is learning rate of gradient descent algorithm
#max_iter determines the maximum number of iterations if the gradients descent_
    ↪does not converge.
#continue the training while gradient > 0.1 or the number steps is less max_iter

#returns model as tuple of (weights,bias)

def fit(x, y_true, learning_rate, lambda_, max_iter, verbose=0):
    weights = np.random.normal(0, 1, x.shape[1]);
    bias = np.random.normal(0,1,1)

    for i in range (max_iter):

        y_prod = prob(x, weights, bias)
        dw,db = grad_w_b(x, weights, y_prod, y_true, lambda_)

        H, second_db = second_grad_w(x,weights,y_prod,y_true,lambda_)
        second_gradident = np.dot(np.linalg.inv(H),dw)

        old_weight = weights
        weights += second_gradident

        cross_loss = loss(weights, y_prod, y_true, lambda_)

        if verbose:
            print("Iteration Number: "+str(i)+"\nLoss: "+ str(cross_loss)+"\nl2_
            ↪norm of gradients: "+str(np.linalg.norm(dw))+"\nl2 norm of weights: "+str(np.
            ↪linalg.norm(weights)))
            print("-----")
            cost_list.append(cross_loss)
        #change the condition appropriately
        # while True:

        #         if verbose: #verbose is used for debugging purposes
        #             #print iteration number, loss, l2 norm of gradients, l2 norm of_
            ↪weights
        #                 pass
        return (weights, bias)

```

```
[106]: def accuracy(x, y_true, model):  
        w, b = model  
        return np.sum((prob(x, w, b)>0.5).astype(np.double) == y_true) / y_true.  
        ↪shape[0]
```

```
[107]: learning_rate = 0.001  
        lambda_ = 1.0  
        max_iter = 10000  
        model = fit(xtrain_normal, ytrain, learning_rate, lambda_, max_iter, verbose=1)↵  
        ↪#keep the verbose on here for your submissions
```

```
Iteration Number: 0  
Loss: 105.02215387521849  
l2 norm of gradients: 1.042082558802105  
l2 norm of weights: 6.623264099107694  
-----  
Iteration Number: 1  
Loss: 104.96040358850966  
l2 norm of gradients: 1.0417264475069021  
l2 norm of weights: 6.621168814145506  
-----  
Iteration Number: 2  
Loss: 104.8986925689002  
l2 norm of gradients: 1.0413704425779455  
l2 norm of weights: 6.619074284909974  
-----  
Iteration Number: 3  
Loss: 104.83702078614908  
l2 norm of gradients: 1.0410145439273442  
l2 norm of weights: 6.616980510728018  
-----  
Iteration Number: 4  
Loss: 104.77538817983537  
l2 norm of gradients: 1.0406587514670602  
l2 norm of weights: 6.614887490928327  
-----  
Iteration Number: 5  
Loss: 104.71379471709757  
l2 norm of gradients: 1.040303065108911  
l2 norm of weights: 6.612795224841365  
-----  
Iteration Number: 6  
Loss: 104.65224036861085  
l2 norm of gradients: 1.0399474847645698  
l2 norm of weights: 6.610703711799365  
-----  
Iteration Number: 7  
Loss: 104.59072509121279
```


l2 norm of gradients: 1.0395920103455705
l2 norm of weights: 6.608612951136326

Iteration Number: 8
Loss: 104.52924881906924
l2 norm of gradients: 1.0392366417633083
l2 norm of weights: 6.606522942188009

Iteration Number: 9
Loss: 104.46781154872754
l2 norm of gradients: 1.0388813789290423
l2 norm of weights: 6.604433684291944

Iteration Number: 10
Loss: 104.4064132224843
l2 norm of gradients: 1.0385262217538964
l2 norm of weights: 6.602345176787415

Iteration Number: 11
Loss: 104.34505379753972
l2 norm of gradients: 1.0381711701488643
l2 norm of weights: 6.600257419015467

Iteration Number: 12
Loss: 104.28373324681381
l2 norm of gradients: 1.0378162240248088
l2 norm of weights: 6.598170410318899

Iteration Number: 13
Loss: 104.22245151761734
l2 norm of gradients: 1.0374613832924655
l2 norm of weights: 6.596084150042268

Iteration Number: 14
Loss: 104.161208564329
l2 norm of gradients: 1.0371066478624449
l2 norm of weights: 6.593998637531875

Iteration Number: 15
Loss: 104.100004375302
l2 norm of gradients: 1.0367520176452336
l2 norm of weights: 6.591913872135776

Iteration Number: 16
Loss: 104.0388388805733
l2 norm of gradients: 1.0363974925511978
l2 norm of weights: 6.58982985320377

```
Iteration Number: 17
Loss: 103.97771206340836
l2 norm of gradients: 1.0360430724905847
l2 norm of weights: 6.587746580087401
-----
Iteration Number: 18
Loss: 103.91662387231034
l2 norm of gradients: 1.0356887573735245
l2 norm of weights: 6.585664052139958
-----
Iteration Number: 19
Loss: 103.85557427094378
l2 norm of gradients: 1.0353345471100326
l2 norm of weights: 6.583582268716465
-----
Iteration Number: 20
Loss: 103.79456321856392
l2 norm of gradients: 1.0349804416100121
l2 norm of weights: 6.581501229173683
-----
Iteration Number: 21
Loss: 103.73359066956571
l2 norm of gradients: 1.0346264407832557
l2 norm of weights: 6.579420932870114
-----
Iteration Number: 22
Loss: 103.67265660231695
l2 norm of gradients: 1.0342725445394474
l2 norm of weights: 6.5773413791659845
-----
Iteration Number: 23
Loss: 103.6117609781275
l2 norm of gradients: 1.0339187527881655
l2 norm of weights: 6.575262567423257
-----
Iteration Number: 24
Loss: 103.55090373430679
l2 norm of gradients: 1.033565065438884
l2 norm of weights: 6.573184497005618
-----
Iteration Number: 25
Loss: 103.49008484818808
l2 norm of gradients: 1.0332114824009748
l2 norm of weights: 6.571107167278482
-----
Iteration Number: 26
Loss: 103.42930428390821
l2 norm of gradients: 1.0328580035837096
```

```

12 norm of weights: 6.569030577608984
-----
Iteration Number: 27
Loss: 103.36856199423579
12 norm of gradients: 1.0325046288962627
12 norm of weights: 6.56695472736598
-----
Iteration Number: 28
Loss: 103.30785794900721
12 norm of gradients: 1.0321513582477122
12 norm of weights: 6.564879615920042
-----
Iteration Number: 29
Loss: 103.24719211271639
12 norm of gradients: 1.0317981915470429
12 norm of weights: 6.562805242643462
-----
Iteration Number: 30
Loss: 103.18656442985001
12 norm of gradients: 1.0314451287031474
12 norm of weights: 6.56073160691024
-----
Iteration Number: 31
Loss: 103.12597488094816
12 norm of gradients: 1.031092169624829
12 norm of weights: 6.558658708096091
-----
Iteration Number: 32
Loss: 103.0654234210975
12 norm of gradients: 1.0307393142208032
12 norm of weights: 6.556586545578433
-----
Iteration Number: 33
Loss: 103.00491000917978
12 norm of gradients: 1.0303865623997002
12 norm of weights: 6.5545151187363935
-----
Iteration Number: 34
Loss: 102.94443459837967
12 norm of gradients: 1.030033914070066
12 norm of weights: 6.5524444269507995
-----
Iteration Number: 35
Loss: 102.883997180486
12 norm of gradients: 1.029681369140366
12 norm of weights: 6.55037446960418
-----
Iteration Number: 36

```

```
Loss: 102.82359767883662
l2 norm of gradients: 1.0293289275189852
l2 norm of weights: 6.548305246080765
-----
Iteration Number: 37
Loss: 102.76323609016261
l2 norm of gradients: 1.0289765891142317
l2 norm of weights: 6.546236755766473
-----
Iteration Number: 38
Loss: 102.70291236240054
l2 norm of gradients: 1.0286243538343374
l2 norm of weights: 6.544168998048919
-----
Iteration Number: 39
Loss: 102.6426264434663
l2 norm of gradients: 1.0282722215874618
l2 norm of weights: 6.54210197231741
-----
Iteration Number: 40
Loss: 102.58237832804754
l2 norm of gradients: 1.0279201922816914
l2 norm of weights: 6.540035677962937
-----
Iteration Number: 41
Loss: 102.52216794887838
l2 norm of gradients: 1.027568265825044
l2 norm of weights: 6.537970114378178
-----
Iteration Number: 42
Loss: 102.46199528620885
l2 norm of gradients: 1.0272164421254697
l2 norm of weights: 6.53590528095749
-----
Iteration Number: 43
Loss: 102.40186029705747
l2 norm of gradients: 1.0268647210908526
l2 norm of weights: 6.533841177096915
-----
Iteration Number: 44
Loss: 102.3417629383633
l2 norm of gradients: 1.0265131026290135
l2 norm of weights: 6.531777802194169
-----
Iteration Number: 45
Loss: 102.28170319642172
l2 norm of gradients: 1.0261615866477107
l2 norm of weights: 6.529715155648641
```

```

-----
Iteration Number: 46
Loss: 102.22168099959502
l2 norm of gradients: 1.0258101730546434
l2 norm of weights: 6.527653236861395
-----
Iteration Number: 47
Loss: 102.16169634069722
l2 norm of gradients: 1.0254588617574525
l2 norm of weights: 6.525592045235163
-----
Iteration Number: 48
Loss: 102.10174916418144
l2 norm of gradients: 1.0251076526637228
l2 norm of weights: 6.5235315801743425
-----
Iteration Number: 49
Loss: 102.04183944340733
l2 norm of gradients: 1.0247565456809853
l2 norm of weights: 6.521471841084995
-----
Iteration Number: 50
Loss: 101.98196714031987
l2 norm of gradients: 1.024405540716718
l2 norm of weights: 6.519412827374843
-----
Iteration Number: 51
Loss: 101.9221322162468
l2 norm of gradients: 1.0240546376783501
l2 norm of weights: 6.5173545384532705
-----
Iteration Number: 52
Loss: 101.862334635282
l2 norm of gradients: 1.0237038364732602
l2 norm of weights: 6.515296973731312
-----
Iteration Number: 53
Loss: 101.80257436188228
l2 norm of gradients: 1.0233531370087823
l2 norm of weights: 6.513240132621659
-----
Iteration Number: 54
Loss: 101.74285135607032
l2 norm of gradients: 1.0230025391922044
l2 norm of weights: 6.511184014538652
-----
Iteration Number: 55
Loss: 101.68316558397697

```

l2 norm of gradients: 1.0226520429307722
l2 norm of weights: 6.509128618898279

Iteration Number: 56
Loss: 101.62351701009457
l2 norm of gradients: 1.0223016481316902
l2 norm of weights: 6.507073945118171

Iteration Number: 57
Loss: 101.56390559505269
l2 norm of gradients: 1.021951354702124
l2 norm of weights: 6.505019992617606

Iteration Number: 58
Loss: 101.50433130997928
l2 norm of gradients: 1.0216011625492012
l2 norm of weights: 6.502966760817497

Iteration Number: 59
Loss: 101.4447941103
l2 norm of gradients: 1.0212510715800143
l2 norm of weights: 6.5009142491403935

Iteration Number: 60
Loss: 101.38529396717865
l2 norm of gradients: 1.0209010817016222
l2 norm of weights: 6.498862457010482

Iteration Number: 61
Loss: 101.32583084641075
l2 norm of gradients: 1.020551192821052
l2 norm of weights: 6.4968113838535775

Iteration Number: 62
Loss: 101.26640470395805
l2 norm of gradients: 1.0202014048452994
l2 norm of weights: 6.494761029097123

Iteration Number: 63
Loss: 101.20701550430627
l2 norm of gradients: 1.0198517176813338
l2 norm of weights: 6.49271139217019

Iteration Number: 64
Loss: 101.14766322081675
l2 norm of gradients: 1.0195021312360961
l2 norm of weights: 6.490662472503469

Iteration Number: 65
Loss: 101.08834781617207
l2 norm of gradients: 1.0191526454165032
l2 norm of weights: 6.48861426952927

Iteration Number: 66
Loss: 101.02906924870157
l2 norm of gradients: 1.0188032601294497
l2 norm of weights: 6.486566782681526

Iteration Number: 67
Loss: 100.96982748415496
l2 norm of gradients: 1.018453975281807
l2 norm of weights: 6.484520011395776

Iteration Number: 68
Loss: 100.91062249516827
l2 norm of gradients: 1.0181047907804288
l2 norm of weights: 6.4824739551091755

Iteration Number: 69
Loss: 100.8514542383368
l2 norm of gradients: 1.0177557065321496
l2 norm of weights: 6.4804286132604885

Iteration Number: 70
Loss: 100.79232268266058
l2 norm of gradients: 1.017406722443789
l2 norm of weights: 6.4783839852900815

Iteration Number: 71
Loss: 100.73322778856972
l2 norm of gradients: 1.0170578384221511
l2 norm of weights: 6.476340070639928

Iteration Number: 72
Loss: 100.67416951951517
l2 norm of gradients: 1.0167090543740274
l2 norm of weights: 6.474296868753599

Iteration Number: 73
Loss: 100.61514786064666
l2 norm of gradients: 1.0163603702061987
l2 norm of weights: 6.47225437907626

Iteration Number: 74
Loss: 100.55616274899793
l2 norm of gradients: 1.0160117858254365

```

12 norm of weights: 6.470212601054676
-----
Iteration Number: 75
Loss: 100.4972141647668
12 norm of gradients: 1.0156633011385043
12 norm of weights: 6.468171534137201
-----
Iteration Number: 76
Loss: 100.43830208272587
12 norm of gradients: 1.0153149160521595
12 norm of weights: 6.466131177773776
-----
Iteration Number: 77
Loss: 100.37942644756089
12 norm of gradients: 1.0149666304731555
12 norm of weights: 6.46409153141593
-----
Iteration Number: 78
Loss: 100.32058723727756
12 norm of gradients: 1.0146184443082422
12 norm of weights: 6.462052594516773
-----
Iteration Number: 79
Loss: 100.26178440562143
12 norm of gradients: 1.0142703574641694
12 norm of weights: 6.460014366530993
-----
Iteration Number: 80
Loss: 100.20301794361842
12 norm of gradients: 1.0139223698476862
12 norm of weights: 6.45797684691486
-----
Iteration Number: 81
Loss: 100.14428779039413
12 norm of gradients: 1.0135744813655452
12 norm of weights: 6.455940035126215
-----
Iteration Number: 82
Loss: 100.08559392549712
12 norm of gradients: 1.0132266919245008
12 norm of weights: 6.4539039306244685
-----
Iteration Number: 83
Loss: 100.02693630848388
12 norm of gradients: 1.0128790014313145
12 norm of weights: 6.451868532870601
-----
Iteration Number: 84

```



```
Loss: 99.96831490681913
l2 norm of gradients: 1.0125314097927538
l2 norm of weights: 6.449833841327158
-----
Iteration Number: 85
Loss: 99.90972969072274
l2 norm of gradients: 1.0121839169155942
l2 norm of weights: 6.447799855458246
-----
Iteration Number: 86
Loss: 99.85118062810609
l2 norm of gradients: 1.0118365227066222
l2 norm of weights: 6.445766574729531
-----
Iteration Number: 87
Loss: 99.79266766685988
l2 norm of gradients: 1.0114892270726348
l2 norm of weights: 6.443733998608239
-----
Iteration Number: 88
Loss: 99.7341908050314
l2 norm of gradients: 1.0111420299204428
l2 norm of weights: 6.441702126563142
-----
Iteration Number: 89
Loss: 99.6757499827147
l2 norm of gradients: 1.0107949311568705
l2 norm of weights: 6.439670958064569
-----
Iteration Number: 90
Loss: 99.61734517598079
l2 norm of gradients: 1.0104479306887602
l2 norm of weights: 6.437640492584393
-----
Iteration Number: 91
Loss: 99.55897634840888
l2 norm of gradients: 1.0101010284229694
l2 norm of weights: 6.435610729596032
-----
Iteration Number: 92
Loss: 99.50064346559493
l2 norm of gradients: 1.0097542242663768
l2 norm of weights: 6.4335816685744485
-----
Iteration Number: 93
Loss: 99.44234650262476
l2 norm of gradients: 1.0094075181258797
l2 norm of weights: 6.431553308996137
```

```

-----
Iteration Number: 94
Loss: 99.38408541855114
l2 norm of gradients: 1.009060909908399
l2 norm of weights: 6.429525650339135
-----
Iteration Number: 95
Loss: 99.3258601839033
l2 norm of gradients: 1.0087143995208783
l2 norm of weights: 6.427498692083006
-----
Iteration Number: 96
Loss: 99.26767076241433
l2 norm of gradients: 1.008367986870286
l2 norm of weights: 6.425472433708847
-----
Iteration Number: 97
Loss: 99.20951712143378
l2 norm of gradients: 1.008021671863617
l2 norm of weights: 6.42344687469928
-----
Iteration Number: 98
Loss: 99.1513992294811
l2 norm of gradients: 1.0076754544078945
l2 norm of weights: 6.421422014538452
-----
Iteration Number: 99
Loss: 99.09331705709945
l2 norm of gradients: 1.0073293344101701
l2 norm of weights: 6.419397852712028
-----
Iteration Number: 100
Loss: 99.03527056125739
l2 norm of gradients: 1.0069833117775266
l2 norm of weights: 6.417374388707193
-----
Iteration Number: 101
Loss: 98.97725972151918
l2 norm of gradients: 1.006637386417078
l2 norm of weights: 6.415351622012644
-----
Iteration Number: 102
Loss: 98.91928449447708
l2 norm of gradients: 1.0062915582359722
l2 norm of weights: 6.413329552118594
-----
Iteration Number: 103
Loss: 98.8613448555771

```

l2 norm of gradients: 1.0059458271413917
l2 norm of weights: 6.411308178516758

Iteration Number: 104
Loss: 98.80344076571708
l2 norm of gradients: 1.0056001930405551
l2 norm of weights: 6.409287500700361

Iteration Number: 105
Loss: 98.74557220129667
l2 norm of gradients: 1.0052546558407187
l2 norm of weights: 6.40726751816413

Iteration Number: 106
Loss: 98.68773911819098
l2 norm of gradients: 1.0049092154491772
l2 norm of weights: 6.405248230404291

Iteration Number: 107
Loss: 98.62994149353864
l2 norm of gradients: 1.0045638717732648
l2 norm of weights: 6.403229636918565

Iteration Number: 108
Loss: 98.57217929362872
l2 norm of gradients: 1.0042186247203582
l2 norm of weights: 6.401211737206168

Iteration Number: 109
Loss: 98.5144524804658
l2 norm of gradients: 1.003873474197876
l2 norm of weights: 6.399194530767806

Iteration Number: 110
Loss: 98.45676102498913
l2 norm of gradients: 1.0035284201132812
l2 norm of weights: 6.397178017105673

Iteration Number: 111
Loss: 98.39910490038237
l2 norm of gradients: 1.0031834623740814
l2 norm of weights: 6.395162195723443

Iteration Number: 112
Loss: 98.341484066723
l2 norm of gradients: 1.0028386008878314
l2 norm of weights: 6.393147066126279

Iteration Number: 113
Loss: 98.28389849328573
l2 norm of gradients: 1.0024938355621333
l2 norm of weights: 6.391132627820814

Iteration Number: 114
Loss: 98.22634815597974
l2 norm of gradients: 1.0021491663046387
l2 norm of weights: 6.389118880315161

Iteration Number: 115
Loss: 98.1688330181021
l2 norm of gradients: 1.0018045930230484
l2 norm of weights: 6.3871058231189055

Iteration Number: 116
Loss: 98.11135304131265
l2 norm of gradients: 1.0014601156251155
l2 norm of weights: 6.385093455743098

Iteration Number: 117
Loss: 98.05390820487699
l2 norm of gradients: 1.0011157340186456
l2 norm of weights: 6.383081777700258

Iteration Number: 118
Loss: 97.99649847449118
l2 norm of gradients: 1.0007714481114978
l2 norm of weights: 6.381070788504369

Iteration Number: 119
Loss: 97.93912381390169
l2 norm of gradients: 1.0004272578115863
l2 norm of weights: 6.379060487670872

Iteration Number: 120
Loss: 97.88178419591259
l2 norm of gradients: 1.0000831630268814
l2 norm of weights: 6.377050874716663

Iteration Number: 121
Loss: 97.82447957944186
l2 norm of gradients: 0.999739163665411
l2 norm of weights: 6.375041949160098

Iteration Number: 122
Loss: 97.76720995267449
l2 norm of gradients: 0.9993952596352611

```

12 norm of weights: 6.373033710520977
-----
Iteration Number: 123
Loss: 97.7099752676704
12 norm of gradients: 0.9990514508445772
12 norm of weights: 6.371026158320553
-----
Iteration Number: 124
Loss: 97.65277549956055
12 norm of gradients: 0.9987077372015661
12 norm of weights: 6.36901929208152
-----
Iteration Number: 125
Loss: 97.59561062069943
12 norm of gradients: 0.9983641186144957
12 norm of weights: 6.3670131113280135
-----
Iteration Number: 126
Loss: 97.5384805867825
12 norm of gradients: 0.9980205949916969
12 norm of weights: 6.365007615585611
-----
Iteration Number: 127
Loss: 97.48138538450246
12 norm of gradients: 0.9976771662415652
12 norm of weights: 6.363002804381322
-----
Iteration Number: 128
Loss: 97.42432497356312
12 norm of gradients: 0.9973338322725602
12 norm of weights: 6.360998677243591
-----
Iteration Number: 129
Loss: 97.36729931946599
12 norm of gradients: 0.9969905929932084
12 norm of weights: 6.358995233702288
-----
Iteration Number: 130
Loss: 97.31030839935488
12 norm of gradients: 0.9966474483121026
12 norm of weights: 6.356992473288713
-----
Iteration Number: 131
Loss: 97.25335217959923
12 norm of gradients: 0.9963043981379048
12 norm of weights: 6.354990395535588
-----
Iteration Number: 132

```

```

Loss: 97.19643062806345
12 norm of gradients: 0.995961442379345
12 norm of weights: 6.352988999977056
-----
Iteration Number: 133
Loss: 97.1395437160874
12 norm of gradients: 0.9956185809452244
12 norm of weights: 6.350988286148674
-----
Iteration Number: 134
Loss: 97.08269141100129
12 norm of gradients: 0.9952758137444149
12 norm of weights: 6.348988253587414
-----
Iteration Number: 135
Loss: 97.02587368925975
12 norm of gradients: 0.9949331406858604
12 norm of weights: 6.346988901831663
-----
Iteration Number: 136
Loss: 96.9690905081136
12 norm of gradients: 0.9945905616785781
12 norm of weights: 6.344990230421211
-----
Iteration Number: 137
Loss: 96.91234184842263
12 norm of gradients: 0.9942480766316592
12 norm of weights: 6.3429922388972555
-----
Iteration Number: 138
Loss: 96.85562767342176
12 norm of gradients: 0.9939056854542705
12 norm of weights: 6.340994926802395
-----
Iteration Number: 139
Loss: 96.79894795707028
12 norm of gradients: 0.9935633880556537
12 norm of weights: 6.338998293680627
-----
Iteration Number: 140
Loss: 96.74230267450453
12 norm of gradients: 0.993221184345128
12 norm of weights: 6.337002339077344
-----
Iteration Number: 141
Loss: 96.68569177964795
12 norm of gradients: 0.9928790742320902
12 norm of weights: 6.33500706253933

```

```
-----
Iteration Number: 142
Loss: 96.6291152543541
l2 norm of gradients: 0.9925370576260162
l2 norm of weights: 6.333012463614764
-----
Iteration Number: 143
Loss: 96.57257306611778
l2 norm of gradients: 0.9921951344364603
l2 norm of weights: 6.331018541853205
-----
Iteration Number: 144
Loss: 96.51606518691958
l2 norm of gradients: 0.9918533045730583
l2 norm of weights: 6.329025296805599
-----
Iteration Number: 145
Loss: 96.45959158574465
l2 norm of gradients: 0.991511567945527
l2 norm of weights: 6.327032728024272
-----
Iteration Number: 146
Loss: 96.40315223130327
l2 norm of gradients: 0.9911699244636646
l2 norm of weights: 6.325040835062926
-----
Iteration Number: 147
Loss: 96.34674709453287
l2 norm of gradients: 0.9908283740373534
l2 norm of weights: 6.323049617476639
-----
Iteration Number: 148
Loss: 96.29037614915083
l2 norm of gradients: 0.9904869165765583
l2 norm of weights: 6.321059074821861
-----
Iteration Number: 149
Loss: 96.23403936024161
l2 norm of gradients: 0.9901455519913293
l2 norm of weights: 6.319069206656406
-----
Iteration Number: 150
Loss: 96.17773670390832
l2 norm of gradients: 0.9898042801918014
l2 norm of weights: 6.317080012539458
-----
Iteration Number: 151
Loss: 96.12146814573619
```

l2 norm of gradients: 0.9894631010881962
l2 norm of weights: 6.315091492031562

Iteration Number: 152
Loss: 96.06523365957021
l2 norm of gradients: 0.9891220145908218
l2 norm of weights: 6.313103644694622

Iteration Number: 153
Loss: 96.0090332163314
l2 norm of gradients: 0.9887810206100737
l2 norm of weights: 6.311116470091895

Iteration Number: 154
Loss: 95.95286678498881
l2 norm of gradients: 0.9884401190564363
l2 norm of weights: 6.309129967787998

Iteration Number: 155
Loss: 95.89673433639777
l2 norm of gradients: 0.988099309840482
l2 norm of weights: 6.307144137348891

Iteration Number: 156
Loss: 95.84063584609655
l2 norm of gradients: 0.9877585928728743
l2 norm of weights: 6.305158978341884

Iteration Number: 157
Loss: 95.78457127737443
l2 norm of gradients: 0.9874179680643664
l2 norm of weights: 6.303174490335634

Iteration Number: 158
Loss: 95.72854060862008
l2 norm of gradients: 0.9870774353258028
l2 norm of weights: 6.301190672900133

Iteration Number: 159
Loss: 95.67254380874388
l2 norm of gradients: 0.9867369945681198
l2 norm of weights: 6.299207525606716

Iteration Number: 160
Loss: 95.61658084270076
l2 norm of gradients: 0.9863966457023459
l2 norm of weights: 6.297225048028052

Iteration Number: 161
Loss: 95.56065169165574
l2 norm of gradients: 0.9860563886396034
l2 norm of weights: 6.295243239738139

Iteration Number: 162
Loss: 95.50475632025687
l2 norm of gradients: 0.9857162232911082
l2 norm of weights: 6.293262100312307

Iteration Number: 163
Loss: 95.44889470575166
l2 norm of gradients: 0.9853761495681695
l2 norm of weights: 6.291281629327213

Iteration Number: 164
Loss: 95.3930668094653
l2 norm of gradients: 0.9850361673821931
l2 norm of weights: 6.289301826360833

Iteration Number: 165
Loss: 95.33727261502864
l2 norm of gradients: 0.9846962766446793
l2 norm of weights: 6.287322690992467

Iteration Number: 166
Loss: 95.28151208373122
l2 norm of gradients: 0.9843564772672247
l2 norm of weights: 6.28534422280273

Iteration Number: 167
Loss: 95.2257851965022
l2 norm of gradients: 0.9840167691615235
l2 norm of weights: 6.283366421373552

Iteration Number: 168
Loss: 95.17009191882177
l2 norm of gradients: 0.9836771522393657
l2 norm of weights: 6.2813892862881735

Iteration Number: 169
Loss: 95.11443222241549
l2 norm of gradients: 0.9833376264126402
l2 norm of weights: 6.279412817131142

Iteration Number: 170
Loss: 95.05880608152151
l2 norm of gradients: 0.9829981915933342

```
l2 norm of weights: 6.277437013488313
-----
Iteration Number: 171
Loss: 95.0032134672111
l2 norm of gradients: 0.9826588476935333
l2 norm of weights: 6.275461874946841
-----
Iteration Number: 172
Loss: 94.94765435155219
l2 norm of gradients: 0.9823195946254227
l2 norm of weights: 6.273487401095181
-----
Iteration Number: 173
Loss: 94.89212870702701
l2 norm of gradients: 0.9819804323012877
l2 norm of weights: 6.271513591523086
-----
Iteration Number: 174
Loss: 94.83663650705314
l2 norm of gradients: 0.9816413606335137
l2 norm of weights: 6.2695404458216
-----
Iteration Number: 175
Loss: 94.78117771765602
l2 norm of gradients: 0.9813023795345868
l2 norm of weights: 6.267567963583058
-----
Iteration Number: 176
Loss: 94.72575231837826
l2 norm of gradients: 0.9809634889170945
l2 norm of weights: 6.265596144401083
-----
Iteration Number: 177
Loss: 94.67036027707604
l2 norm of gradients: 0.9806246886937261
l2 norm of weights: 6.26362498787058
-----
Iteration Number: 178
Loss: 94.61500156545317
l2 norm of gradients: 0.980285978777273
l2 norm of weights: 6.26165449358774
-----
Iteration Number: 179
Loss: 94.5596761617461
l2 norm of gradients: 0.979947359080629
l2 norm of weights: 6.259684661150028
-----
Iteration Number: 180
```

Loss: 94.50438403358609
l2 norm of gradients: 0.979608829516791
l2 norm of weights: 6.257715490156187

Iteration Number: 181
Loss: 94.44912514918418
l2 norm of gradients: 0.9792703899988594
l2 norm of weights: 6.255746980206233

Iteration Number: 182
Loss: 94.39389949105104
l2 norm of gradients: 0.9789320404400378
l2 norm of weights: 6.253779130901452

Iteration Number: 183
Loss: 94.33870702546395
l2 norm of gradients: 0.9785937807536341
l2 norm of weights: 6.2518119418443945

Iteration Number: 184
Loss: 94.28354772980781
l2 norm of gradients: 0.9782556108530611
l2 norm of weights: 6.249845412638877

Iteration Number: 185
Loss: 94.22842156745497
l2 norm of gradients: 0.9779175306518354
l2 norm of weights: 6.247879542889978

Iteration Number: 186
Loss: 94.1733285220956
l2 norm of gradients: 0.9775795400635798
l2 norm of weights: 6.245914332204031

Iteration Number: 187
Loss: 94.11826856020706
l2 norm of gradients: 0.9772416390020218
l2 norm of weights: 6.243949780188628

Iteration Number: 188
Loss: 94.06324165781547
l2 norm of gradients: 0.9769038273809945
l2 norm of weights: 6.241985886452612

Iteration Number: 189
Loss: 94.00824778288309
l2 norm of gradients: 0.9765661051144374
l2 norm of weights: 6.240022650606076

```
-----  
Iteration Number: 190  
Loss: 93.95328691442162  
l2 norm of gradients: 0.9762284721163962  
l2 norm of weights: 6.238060072260359  
-----  
Iteration Number: 191  
Loss: 93.89835902278438  
l2 norm of gradients: 0.9758909283010231  
l2 norm of weights: 6.236098151028044  
-----  
Iteration Number: 192  
Loss: 93.84346408076588  
l2 norm of gradients: 0.9755534735825767  
l2 norm of weights: 6.234136886522956  
-----  
Iteration Number: 193  
Loss: 93.78860205978954  
l2 norm of gradients: 0.9752161078754232  
l2 norm of weights: 6.232176278360157  
-----  
Iteration Number: 194  
Loss: 93.73377293957716  
l2 norm of gradients: 0.9748788310940357  
l2 norm of weights: 6.230216326155944  
-----  
Iteration Number: 195  
Loss: 93.67897668757577  
l2 norm of gradients: 0.974541643152995  
l2 norm of weights: 6.228257029527848  
-----  
Iteration Number: 196  
Loss: 93.62421327657209  
l2 norm of gradients: 0.9742045439669892  
l2 norm of weights: 6.226298388094627  
-----  
Iteration Number: 197  
Loss: 93.5694826857964  
l2 norm of gradients: 0.9738675334508143  
l2 norm of weights: 6.22434040147627  
-----  
Iteration Number: 198  
Loss: 93.51478488559754  
l2 norm of gradients: 0.9735306115193748  
l2 norm of weights: 6.222383069293984  
-----  
Iteration Number: 199  
Loss: 93.4601198470088
```

l2 norm of gradients: 0.9731937780876829
l2 norm of weights: 6.220426391170201

Iteration Number: 200
Loss: 93.40548754539641
l2 norm of gradients: 0.9728570330708595
l2 norm of weights: 6.218470366728572

Iteration Number: 201
Loss: 93.35088795606488
l2 norm of gradients: 0.9725203763841332
l2 norm of weights: 6.216514995593962

Iteration Number: 202
Loss: 93.29632104918845
l2 norm of gradients: 0.9721838079428423
l2 norm of weights: 6.214560277392448

Iteration Number: 203
Loss: 93.24178680379728
l2 norm of gradients: 0.9718473276624332
l2 norm of weights: 6.212606211751317

Iteration Number: 204
Loss: 93.18728518616608
l2 norm of gradients: 0.9715109354584611
l2 norm of weights: 6.210652798299065

Iteration Number: 205
Loss: 93.13281617912652
l2 norm of gradients: 0.9711746312465902
l2 norm of weights: 6.20870003666539

Iteration Number: 206
Loss: 93.07837975005431
l2 norm of gradients: 0.9708384149425938
l2 norm of weights: 6.206747926481193

Iteration Number: 207
Loss: 93.02397587670404
l2 norm of gradients: 0.9705022864623545
l2 norm of weights: 6.204796467378574

Iteration Number: 208
Loss: 92.96960452916251
l2 norm of gradients: 0.9701662457218629
l2 norm of weights: 6.202845658990827

Iteration Number: 209
Loss: 92.915265688518
l2 norm of gradients: 0.9698302926372206
l2 norm of weights: 6.200895500952443

Iteration Number: 210
Loss: 92.86095931903802
l2 norm of gradients: 0.9694944271246367
l2 norm of weights: 6.198945992899098

Iteration Number: 211
Loss: 92.80668540112717
l2 norm of gradients: 0.9691586491004304
l2 norm of weights: 6.196997134467661

Iteration Number: 212
Loss: 92.75244391185309
l2 norm of gradients: 0.96882295848103
l2 norm of weights: 6.195048925296184

Iteration Number: 213
Loss: 92.6982348187593
l2 norm of gradients: 0.9684873551829727
l2 norm of weights: 6.1931013650239

Iteration Number: 214
Loss: 92.64405809990616
l2 norm of gradients: 0.9681518391229055
l2 norm of weights: 6.191154453291225

Iteration Number: 215
Loss: 92.58991372809461
l2 norm of gradients: 0.9678164102175841
l2 norm of weights: 6.189208189739749

Iteration Number: 216
Loss: 92.53580167853839
l2 norm of gradients: 0.9674810683838735
l2 norm of weights: 6.187262574012236

Iteration Number: 217
Loss: 92.48172192617591
l2 norm of gradients: 0.9671458135387476
l2 norm of weights: 6.1853176057526245

Iteration Number: 218
Loss: 92.4276744472454
l2 norm of gradients: 0.9668106455992901

```

12 norm of weights: 6.183373284606017
-----
Iteration Number: 219
Loss: 92.37365921173024
12 norm of gradients: 0.9664755644826928
12 norm of weights: 6.1814296102186885
-----
Iteration Number: 220
Loss: 92.3196761991725
12 norm of gradients: 0.9661405701062568
12 norm of weights: 6.179486582238073
-----
Iteration Number: 221
Loss: 92.2657253827927
12 norm of gradients: 0.9658056623873917
12 norm of weights: 6.177544200312764
-----
Iteration Number: 222
Loss: 92.21180673668238
12 norm of gradients: 0.9654708412436163
12 norm of weights: 6.175602464092518
-----
Iteration Number: 223
Loss: 92.15792023549989
12 norm of gradients: 0.9651361065925572
12 norm of weights: 6.173661373228243
-----
Iteration Number: 224
Loss: 92.10406585412575
12 norm of gradients: 0.9648014583519506
12 norm of weights: 6.171720927372002
-----
Iteration Number: 225
Loss: 92.05024356751694
12 norm of gradients: 0.96446689643964
12 norm of weights: 6.169781126177006
-----
Iteration Number: 226
Loss: 91.99645335289516
12 norm of gradients: 0.9641324207735774
12 norm of weights: 6.167841969297617
-----
Iteration Number: 227
Loss: 91.94269518275341
12 norm of gradients: 0.9637980312718233
12 norm of weights: 6.16590345638934
-----
Iteration Number: 228

```

Loss: 91.88896903487303
l2 norm of gradients: 0.9634637278525449
l2 norm of weights: 6.16396558710882

Iteration Number: 229
Loss: 91.83527488083637
l2 norm of gradients: 0.9631295104340183
l2 norm of weights: 6.162028361113846

Iteration Number: 230
Loss: 91.78161269825114
l2 norm of gradients: 0.962795378934626
l2 norm of weights: 6.160091778063341

Iteration Number: 231
Loss: 91.72798246026298
l2 norm of gradients: 0.9624613332728587
l2 norm of weights: 6.1581558376173655

Iteration Number: 232
Loss: 91.6743841456214
l2 norm of gradients: 0.9621273733673136
l2 norm of weights: 6.156220539437108

Iteration Number: 233
Loss: 91.62081772908441
l2 norm of gradients: 0.9617934991366949
l2 norm of weights: 6.1542858831848894

Iteration Number: 234
Loss: 91.56728318259934
l2 norm of gradients: 0.9614597104998134
l2 norm of weights: 6.152351868524156

Iteration Number: 235
Loss: 91.51378048329532
l2 norm of gradients: 0.961126007375586
l2 norm of weights: 6.150418495119481

Iteration Number: 236
Loss: 91.46030961046021
l2 norm of gradients: 0.9607923896830364
l2 norm of weights: 6.148485762636555

Iteration Number: 237
Loss: 91.4068705353407
l2 norm of gradients: 0.9604588573412934
l2 norm of weights: 6.1465536707421915


```

-----
Iteration Number: 238
Loss: 91.35346323385566
l2 norm of gradients: 0.9601254102695916
l2 norm of weights: 6.144622219104317
-----
Iteration Number: 239
Loss: 91.3000876837371
l2 norm of gradients: 0.9597920483872714
l2 norm of weights: 6.142691407391977
-----
Iteration Number: 240
Loss: 91.2467438611064
l2 norm of gradients: 0.9594587716137774
l2 norm of weights: 6.140761235275323
-----
Iteration Number: 241
Loss: 91.19343173832891
l2 norm of gradients: 0.9591255798686593
l2 norm of weights: 6.138831702425622
-----
Iteration Number: 242
Loss: 91.14015129343782
l2 norm of gradients: 0.9587924730715711
l2 norm of weights: 6.136902808515241
-----
Iteration Number: 243
Loss: 91.08690250388348
l2 norm of gradients: 0.9584594511422714
l2 norm of weights: 6.1349745532176545
-----
Iteration Number: 244
Loss: 91.03368534327542
l2 norm of gradients: 0.9581265140006213
l2 norm of weights: 6.133046936207438
-----
Iteration Number: 245
Loss: 90.98049978866541
l2 norm of gradients: 0.9577936615665866
l2 norm of weights: 6.131119957160266
-----
Iteration Number: 246
Loss: 90.92734581544552
l2 norm of gradients: 0.9574608937602352
l2 norm of weights: 6.129193615752912
-----
Iteration Number: 247
Loss: 90.87422340078098

```

l2 norm of gradients: 0.9571282105017387
l2 norm of weights: 6.127267911663239

Iteration Number: 248
Loss: 90.82113251742356
l2 norm of gradients: 0.9567956117113694
l2 norm of weights: 6.125342844570205

Iteration Number: 249
Loss: 90.76807314649157
l2 norm of gradients: 0.9564630973095033
l2 norm of weights: 6.123418414153855

Iteration Number: 250
Loss: 90.71504526329245
l2 norm of gradients: 0.956130667216617
l2 norm of weights: 6.1214946200953255

Iteration Number: 251
Loss: 90.662048840096
l2 norm of gradients: 0.9557983213532881
l2 norm of weights: 6.1195714620768324

Iteration Number: 252
Loss: 90.60908385841165
l2 norm of gradients: 0.9554660596401957
l2 norm of weights: 6.117648939781675

Iteration Number: 253
Loss: 90.55615029284058
l2 norm of gradients: 0.9551338819981187
l2 norm of weights: 6.115727052894233

Iteration Number: 254
Loss: 90.50324811609731
l2 norm of gradients: 0.9548017883479356
l2 norm of weights: 6.113805801099962

Iteration Number: 255
Loss: 90.45037731098374
l2 norm of gradients: 0.954469778610625
l2 norm of weights: 6.111885184085393

Iteration Number: 256
Loss: 90.3975378502253
l2 norm of gradients: 0.9541378527072645
l2 norm of weights: 6.109965201538131

Iteration Number: 257
Loss: 90.34472971154727
l2 norm of gradients: 0.9538060105590301
l2 norm of weights: 6.108045853146848

Iteration Number: 258
Loss: 90.29195287025678
l2 norm of gradients: 0.9534742520871957
l2 norm of weights: 6.106127138601285

Iteration Number: 259
Loss: 90.23920730460193
l2 norm of gradients: 0.9531425772131333
l2 norm of weights: 6.10420905759225

Iteration Number: 260
Loss: 90.18649299150921
l2 norm of gradients: 0.9528109858583119
l2 norm of weights: 6.10229160981161

Iteration Number: 261
Loss: 90.13380990673316
l2 norm of gradients: 0.9524794779442972
l2 norm of weights: 6.1003747949522955

Iteration Number: 262
Loss: 90.08115802699874
l2 norm of gradients: 0.9521480533927512
l2 norm of weights: 6.098458612708294

Iteration Number: 263
Loss: 90.02853732986561
l2 norm of gradients: 0.9518167121254315
l2 norm of weights: 6.096543062774648

Iteration Number: 264
Loss: 89.97594779181968
l2 norm of gradients: 0.9514854540641909
l2 norm of weights: 6.094628144847458

Iteration Number: 265
Loss: 89.92338939109707
l2 norm of gradients: 0.9511542791309773
l2 norm of weights: 6.0927138586238705

Iteration Number: 266
Loss: 89.87086210221705
l2 norm of gradients: 0.9508231872478323

```

12 norm of weights: 6.090800203802081
-----
Iteration Number: 267
Loss: 89.81836590582128
12 norm of gradients: 0.9504921783368913
12 norm of weights: 6.088887180081336
-----
Iteration Number: 268
Loss: 89.76590077649966
12 norm of gradients: 0.950161252320383
12 norm of weights: 6.086974787161922
-----
Iteration Number: 269
Loss: 89.71346669083931
12 norm of gradients: 0.9498304091206283
12 norm of weights: 6.085063024745171
-----
Iteration Number: 270
Loss: 89.66106362770816
12 norm of gradients: 0.9494996486600407
12 norm of weights: 6.083151892533451
-----
Iteration Number: 271
Loss: 89.60869156445558
12 norm of gradients: 0.9491689708611245
12 norm of weights: 6.08124139023017
-----
Iteration Number: 272
Loss: 89.5563504800412
12 norm of gradients: 0.9488383756464756
12 norm of weights: 6.079331517539771
-----
Iteration Number: 273
Loss: 89.5040403469299
12 norm of gradients: 0.9485078629387796
12 norm of weights: 6.07742227416773
-----
Iteration Number: 274
Loss: 89.45176114673727
12 norm of gradients: 0.948177432660812
12 norm of weights: 6.075513659820551
-----
Iteration Number: 275
Loss: 89.39951285502781
12 norm of gradients: 0.9478470847354378
12 norm of weights: 6.073605674205772
-----
Iteration Number: 276

```

Loss: 89.34729544898147
l2 norm of gradients: 0.9475168190856104
l2 norm of weights: 6.0716983170319505

Iteration Number: 277
Loss: 89.29510890848597
l2 norm of gradients: 0.9471866356343709
l2 norm of weights: 6.069791588008674

Iteration Number: 278
Loss: 89.24295320902448
l2 norm of gradients: 0.946856534304848
l2 norm of weights: 6.067885486846549

Iteration Number: 279
Loss: 89.19082832942954
l2 norm of gradients: 0.9465265150202572
l2 norm of weights: 6.065980013257201

Iteration Number: 280
Loss: 89.13873424688964
l2 norm of gradients: 0.9461965777038998
l2 norm of weights: 6.064075166953275

Iteration Number: 281
Loss: 89.08667094075706
l2 norm of gradients: 0.9458667222791631
l2 norm of weights: 6.062170947648428

Iteration Number: 282
Loss: 89.03463838569826
l2 norm of gradients: 0.9455369486695191
l2 norm of weights: 6.060267355057334

Iteration Number: 283
Loss: 88.98263656265468
l2 norm of gradients: 0.9452072567985236
l2 norm of weights: 6.058364388895676

Iteration Number: 284
Loss: 88.93066544633442
l2 norm of gradients: 0.9448776465898167
l2 norm of weights: 6.056462048880144

Iteration Number: 285
Loss: 88.87872501622209
l2 norm of gradients: 0.9445481179671206
l2 norm of weights: 6.054560334728437

```
-----
Iteration Number: 286
Loss: 88.82681525073346
l2 norm of gradients: 0.9442186708542407
l2 norm of weights: 6.0526592461592585
-----
Iteration Number: 287
Loss: 88.77493612722634
l2 norm of gradients: 0.9438893051750633
l2 norm of weights: 6.050758782892311
-----
Iteration Number: 288
Loss: 88.72308762528964
l2 norm of gradients: 0.9435600208535558
l2 norm of weights: 6.048858944648303
-----
Iteration Number: 289
Loss: 88.67126972190333
l2 norm of gradients: 0.9432308178137663
l2 norm of weights: 6.046959731148934
-----
Iteration Number: 290
Loss: 88.61948239662406
l2 norm of gradients: 0.9429016959798222
l2 norm of weights: 6.045061142116904
-----
Iteration Number: 291
Loss: 88.56772562397828
l2 norm of gradients: 0.9425726552759295
l2 norm of weights: 6.043163177275906
-----
Iteration Number: 292
Loss: 88.51599938436406
l2 norm of gradients: 0.9422436956263728
l2 norm of weights: 6.0412658363506235
-----
Iteration Number: 293
Loss: 88.46430365780859
l2 norm of gradients: 0.9419148169555139
l2 norm of weights: 6.0393691190667305
-----
Iteration Number: 294
Loss: 88.41263842322543
l2 norm of gradients: 0.9415860191877923
l2 norm of weights: 6.037473025150888
-----
Iteration Number: 295
Loss: 88.36100365213434
```

l2 norm of gradients: 0.9412573022477225
l2 norm of weights: 6.035577554330744

Iteration Number: 296
Loss: 88.30939933142663
l2 norm of gradients: 0.940928666059895
l2 norm of weights: 6.033682706334926

Iteration Number: 297
Loss: 88.25782543360178
l2 norm of gradients: 0.9406001105489753
l2 norm of weights: 6.031788480893048

Iteration Number: 298
Loss: 88.20628194248381
l2 norm of gradients: 0.940271635639702
l2 norm of weights: 6.029894877735698

Iteration Number: 299
Loss: 88.15476883083427
l2 norm of gradients: 0.9399432412568877
l2 norm of weights: 6.028001896594445

Iteration Number: 300
Loss: 88.10328608211955
l2 norm of gradients: 0.9396149273254172
l2 norm of weights: 6.02610953720183

Iteration Number: 301
Loss: 88.0518336738658
l2 norm of gradients: 0.9392866937702473
l2 norm of weights: 6.02421779929137

Iteration Number: 302
Loss: 88.00041158091895
l2 norm of gradients: 0.9389585405164057
l2 norm of weights: 6.022326682597551

Iteration Number: 303
Loss: 87.94901978741363
l2 norm of gradients: 0.9386304674889905
l2 norm of weights: 6.02043618685583

Iteration Number: 304
Loss: 87.89765826750046
l2 norm of gradients: 0.9383024746131694
l2 norm of weights: 6.01854631180263

Iteration Number: 305
Loss: 87.84632700445238
l2 norm of gradients: 0.9379745618141788
l2 norm of weights: 6.016657057175335

Iteration Number: 306
Loss: 87.79502597459911
l2 norm of gradients: 0.9376467290173234
l2 norm of weights: 6.014768422712299

Iteration Number: 307
Loss: 87.7437551573942
l2 norm of gradients: 0.9373189761479751
l2 norm of weights: 6.012880408152834

Iteration Number: 308
Loss: 87.6925145331089
l2 norm of gradients: 0.9369913031315723
l2 norm of weights: 6.0109930132372105

Iteration Number: 309
Loss: 87.6413040763406
l2 norm of gradients: 0.936663709893619
l2 norm of weights: 6.0091062377066535

Iteration Number: 310
Loss: 87.5901237718426
l2 norm of gradients: 0.9363361963596851
l2 norm of weights: 6.00722008130335

Iteration Number: 311
Loss: 87.53897359444105
l2 norm of gradients: 0.9360087624554038
l2 norm of weights: 6.005334543770433

Iteration Number: 312
Loss: 87.48785352626292
l2 norm of gradients: 0.9356814081064722
l2 norm of weights: 6.0034496248519895

Iteration Number: 313
Loss: 87.43676354462843
l2 norm of gradients: 0.9353541332386501
l2 norm of weights: 6.0015653242930584

Iteration Number: 314
Loss: 87.385703628802
l2 norm of gradients: 0.9350269377777592


```

12 norm of weights: 5.99968164183962
-----
Iteration Number: 315
Loss: 87.33467375773151
12 norm of gradients: 0.9346998216496822
12 norm of weights: 5.997798577238607
-----
Iteration Number: 316
Loss: 87.28367391435712
12 norm of gradients: 0.9343727847803628
12 norm of weights: 5.995916130237888
-----
Iteration Number: 317
Loss: 87.23270407207025
12 norm of gradients: 0.9340458270958032
12 norm of weights: 5.9940343005862795
-----
Iteration Number: 318
Loss: 87.18176421631745
12 norm of gradients: 0.9337189485220656
12 norm of weights: 5.9921530880335325
-----
Iteration Number: 319
Loss: 87.13085432152539
12 norm of gradients: 0.9333921489852691
12 norm of weights: 5.99027249233034
-----
Iteration Number: 320
Loss: 87.07997436979143
12 norm of gradients: 0.9330654284115909
12 norm of weights: 5.9883925132283276
-----
Iteration Number: 321
Loss: 87.0291243411114
12 norm of gradients: 0.9327387867272641
12 norm of weights: 5.986513150480055
-----
Iteration Number: 322
Loss: 86.97830421358977
12 norm of gradients: 0.9324122238585772
12 norm of weights: 5.984634403839016
-----
Iteration Number: 323
Loss: 86.92751396750123
12 norm of gradients: 0.9320857397318745
12 norm of weights: 5.982756273059634
-----
Iteration Number: 324

```

Loss: 86.87675358204775
l2 norm of gradients: 0.9317593342735533
l2 norm of weights: 5.980878757897257

Iteration Number: 325
Loss: 86.82602303802375
l2 norm of gradients: 0.9314330074100643
l2 norm of weights: 5.979001858108164

Iteration Number: 326
Loss: 86.77532231385503
l2 norm of gradients: 0.9311067590679107
l2 norm of weights: 5.977125573449556

Iteration Number: 327
Loss: 86.72465138851128
l2 norm of gradients: 0.9307805891736477
l2 norm of weights: 5.975249903679558

Iteration Number: 328
Loss: 86.67401024473602
l2 norm of gradients: 0.9304544976538807
l2 norm of weights: 5.973374848557214

Iteration Number: 329
Loss: 86.62339886102824
l2 norm of gradients: 0.930128484435265
l2 norm of weights: 5.971500407842488

Iteration Number: 330
Loss: 86.57281721641263
l2 norm of gradients: 0.9298025494445055
l2 norm of weights: 5.96962658129626

Iteration Number: 331
Loss: 86.52226529164473
l2 norm of gradients: 0.9294766926083553
l2 norm of weights: 5.967753368680328

Iteration Number: 332
Loss: 86.47174306751604
l2 norm of gradients: 0.9291509138536147
l2 norm of weights: 5.9658807697574

Iteration Number: 333
Loss: 86.42125052272957
l2 norm of gradients: 0.9288252131071311
l2 norm of weights: 5.9640087842910985

```

-----
Iteration Number: 334
Loss: 86.3707876377133
l2 norm of gradients: 0.9284995902957974
l2 norm of weights: 5.962137412045953
-----
Iteration Number: 335
Loss: 86.32035439172355
l2 norm of gradients: 0.9281740453465521
l2 norm of weights: 5.960266652787403
-----
Iteration Number: 336
Loss: 86.26995076665287
l2 norm of gradients: 0.9278485781863772
l2 norm of weights: 5.958396506281793
-----
Iteration Number: 337
Loss: 86.2195767412912
l2 norm of gradients: 0.927523188742299
l2 norm of weights: 5.956526972296374
-----
Iteration Number: 338
Loss: 86.16923229611906
l2 norm of gradients: 0.9271978769413857
l2 norm of weights: 5.9546580505992965
-----
Iteration Number: 339
Loss: 86.11891741335096
l2 norm of gradients: 0.9268726427107473
l2 norm of weights: 5.952789740959614
-----
Iteration Number: 340
Loss: 86.06863207091557
l2 norm of gradients: 0.9265474859775351
l2 norm of weights: 5.950922043147278
-----
Iteration Number: 341
Loss: 86.01837625013006
l2 norm of gradients: 0.9262224066689406
l2 norm of weights: 5.9490549569331375
-----
Iteration Number: 342
Loss: 85.96814993097503
l2 norm of gradients: 0.9258974047121944
l2 norm of weights: 5.947188482088937
-----
Iteration Number: 343
Loss: 85.91795309519672

```

```

12 norm of gradients: 0.9255724800345655
12 norm of weights: 5.945322618387316
-----
Iteration Number: 344
Loss: 85.8677857205705
12 norm of gradients: 0.9252476325633604
12 norm of weights: 5.943457365601803
-----
Iteration Number: 345
Loss: 85.8176477898993
12 norm of gradients: 0.9249228622259231
12 norm of weights: 5.941592723506819
-----
Iteration Number: 346
Loss: 85.76753928430065
12 norm of gradients: 0.9245981689496334
12 norm of weights: 5.939728691877672
-----
Iteration Number: 347
Loss: 85.71746018051452
12 norm of gradients: 0.9242735526619059
12 norm of weights: 5.93786527049056
-----
Iteration Number: 348
Loss: 85.66741046362843
12 norm of gradients: 0.9239490132901896
12 norm of weights: 5.936002459122561
-----
Iteration Number: 349
Loss: 85.61739011161329
12 norm of gradients: 0.9236245507619678
12 norm of weights: 5.934140257551641
-----
Iteration Number: 350
Loss: 85.56739910862304
12 norm of gradients: 0.9233001650047556
12 norm of weights: 5.932278665556642
-----
Iteration Number: 351
Loss: 85.51743742957406
12 norm of gradients: 0.9229758559461009
12 norm of weights: 5.930417682917294
-----
Iteration Number: 352
Loss: 85.46750506125326
12 norm of gradients: 0.922651623513582
12 norm of weights: 5.928557309414195
-----

```

Iteration Number: 353
Loss: 85.41760197991384
l2 norm of gradients: 0.9223274676348077
l2 norm of weights: 5.926697544828829

Iteration Number: 354
Loss: 85.36772816987728
l2 norm of gradients: 0.9220033882374165
l2 norm of weights: 5.924838388943547

Iteration Number: 355
Loss: 85.31788360939672
l2 norm of gradients: 0.9216793852490753
l2 norm of weights: 5.922979841541578

Iteration Number: 356
Loss: 85.2680682813071
l2 norm of gradients: 0.9213554585974787
l2 norm of weights: 5.9211219024070205

Iteration Number: 357
Loss: 85.2182821638512
l2 norm of gradients: 0.9210316082103491
l2 norm of weights: 5.919264571324843

Iteration Number: 358
Loss: 85.16852524187222
l2 norm of gradients: 0.9207078340154342
l2 norm of weights: 5.917407848080879

Iteration Number: 359
Loss: 85.11879749558285
l2 norm of gradients: 0.9203841359405075
l2 norm of weights: 5.915551732461832

Iteration Number: 360
Loss: 85.06909890281362
l2 norm of gradients: 0.9200605139133669
l2 norm of weights: 5.913696224255271

Iteration Number: 361
Loss: 85.01942944925564
l2 norm of gradients: 0.9197369678618346
l2 norm of weights: 5.911841323249622

Iteration Number: 362
Loss: 84.96978911153373
l2 norm of gradients: 0.919413497713755

```

12 norm of weights: 5.909987029234178
-----
Iteration Number: 363
Loss: 84.92017787421078
12 norm of gradients: 0.9190901033969956
12 norm of weights: 5.908133341999089
-----
Iteration Number: 364
Loss: 84.87059571865416
12 norm of gradients: 0.9187667848394446
12 norm of weights: 5.906280261335363
-----
Iteration Number: 365
Loss: 84.82104262358797
12 norm of gradients: 0.9184435419690109
12 norm of weights: 5.904427787034864
-----
Iteration Number: 366
Loss: 84.77151857252885
12 norm of gradients: 0.9181203747136231
12 norm of weights: 5.902575918890312
-----
Iteration Number: 367
Loss: 84.72202354590988
12 norm of gradients: 0.9177972830012292
12 norm of weights: 5.900724656695277
-----
Iteration Number: 368
Loss: 84.67255752547162
12 norm of gradients: 0.917474266759795
12 norm of weights: 5.898874000244184
-----
Iteration Number: 369
Loss: 84.62312049283646
12 norm of gradients: 0.917151325917304
12 norm of weights: 5.897023949332306
-----
Iteration Number: 370
Loss: 84.57371242889772
12 norm of gradients: 0.9168284604017558
12 norm of weights: 5.895174503755762
-----
Iteration Number: 371
Loss: 84.52433331557374
12 norm of gradients: 0.9165056701411666
12 norm of weights: 5.893325663311521
-----
Iteration Number: 372

```

```

Loss: 84.47498313473942
12 norm of gradients: 0.9161829550635667
12 norm of weights: 5.891477427797394
-----
Iteration Number: 373
Loss: 84.42566186636739
12 norm of gradients: 0.9158603150970017
12 norm of weights: 5.889629797012036
-----
Iteration Number: 374
Loss: 84.37636949488343
12 norm of gradients: 0.9155377501695295
12 norm of weights: 5.887782770754945
-----
Iteration Number: 375
Loss: 84.32710600008927
12 norm of gradients: 0.915215260209222
12 norm of weights: 5.885936348826456
-----
Iteration Number: 376
Loss: 84.2778713632979
12 norm of gradients: 0.914892845144162
12 norm of weights: 5.884090531027742
-----
Iteration Number: 377
Loss: 84.2286655681128
12 norm of gradients: 0.9145705049024441
12 norm of weights: 5.882245317160819
-----
Iteration Number: 378
Loss: 84.17948859495154
12 norm of gradients: 0.9142482394121729
12 norm of weights: 5.880400707028529
-----
Iteration Number: 379
Loss: 84.13034042528653
12 norm of gradients: 0.9139260486014631
12 norm of weights: 5.878556700434554
-----
Iteration Number: 380
Loss: 84.08122104191395
12 norm of gradients: 0.9136039323984374
12 norm of weights: 5.8767132971834055
-----
Iteration Number: 381
Loss: 84.03213042623524
12 norm of gradients: 0.913281890731228
12 norm of weights: 5.874870497080426

```

```

-----
Iteration Number: 382
Loss: 83.98306856093727
l2 norm of gradients: 0.9129599235279731
l2 norm of weights: 5.873028299931786
-----
Iteration Number: 383
Loss: 83.93403542828815
l2 norm of gradients: 0.9126380307168183
l2 norm of weights: 5.871186705544484
-----
Iteration Number: 384
Loss: 83.8850310079329
l2 norm of gradients: 0.9123162122259149
l2 norm of weights: 5.869345713726341
-----
Iteration Number: 385
Loss: 83.83605528390434
l2 norm of gradients: 0.9119944679834192
l2 norm of weights: 5.8675053242860065
-----
Iteration Number: 386
Loss: 83.78710823764156
l2 norm of gradients: 0.9116727979174923
l2 norm of weights: 5.86566553703295
-----
Iteration Number: 387
Loss: 83.73818985187249
l2 norm of gradients: 0.9113512019562985
l2 norm of weights: 5.86382635177746
-----
Iteration Number: 388
Loss: 83.68930010823465
l2 norm of gradients: 0.911029680028005
l2 norm of weights: 5.861987768330649
-----
Iteration Number: 389
Loss: 83.64043898925424
l2 norm of gradients: 0.9107082320607821
l2 norm of weights: 5.860149786504442
-----
Iteration Number: 390
Loss: 83.59160647667096
l2 norm of gradients: 0.9103868579828002
l2 norm of weights: 5.858312406111582
-----
Iteration Number: 391
Loss: 83.54280255353275

```



```

12 norm of gradients: 0.9100655577222315
12 norm of weights: 5.856475626965629
-----
Iteration Number: 392
Loss: 83.49402720054154
12 norm of gradients: 0.909744331207248
12 norm of weights: 5.854639448880952
-----
Iteration Number: 393
Loss: 83.44528040209387
12 norm of gradients: 0.9094231783660212
12 norm of weights: 5.852803871672735
-----
Iteration Number: 394
Loss: 83.39656213991961
12 norm of gradients: 0.9091020991267204
12 norm of weights: 5.850968895156966
-----
Iteration Number: 395
Loss: 83.3478723953323
12 norm of gradients: 0.9087810934175141
12 norm of weights: 5.84913451915045
-----
Iteration Number: 396
Loss: 83.2992111515594
12 norm of gradients: 0.9084601611665672
12 norm of weights: 5.84730074347079
-----
Iteration Number: 397
Loss: 83.25057839152217
12 norm of gradients: 0.9081393023020413
12 norm of weights: 5.845467567936401
-----
Iteration Number: 398
Loss: 83.2019740968005
12 norm of gradients: 0.907818516752094
12 norm of weights: 5.843634992366497
-----
Iteration Number: 399
Loss: 83.1533982509514
12 norm of gradients: 0.9074978044448783
12 norm of weights: 5.841803016581097
-----
Iteration Number: 400
Loss: 83.10485083664922
12 norm of gradients: 0.9071771653085411
12 norm of weights: 5.839971640401019
-----

```

Iteration Number: 401
Loss: 83.05633183484036
l2 norm of gradients: 0.9068565992712234
l2 norm of weights: 5.8381408636478795

Iteration Number: 402
Loss: 83.00784122961848
l2 norm of gradients: 0.90653610626106
l2 norm of weights: 5.836310686144095

Iteration Number: 403
Loss: 82.9593790037492
l2 norm of gradients: 0.9062156862061771
l2 norm of weights: 5.834481107712875

Iteration Number: 404
Loss: 82.91094513910434
l2 norm of gradients: 0.9058953390346937
l2 norm of weights: 5.832652128178226

Iteration Number: 405
Loss: 82.86253961918531
l2 norm of gradients: 0.9055750646747195
l2 norm of weights: 5.830823747364944

Iteration Number: 406
Loss: 82.81416242650128
l2 norm of gradients: 0.9052548630543549
l2 norm of weights: 5.828995965098619

Iteration Number: 407
Loss: 82.76581354436426
l2 norm of gradients: 0.90493473410169
l2 norm of weights: 5.827168781205632

Iteration Number: 408
Loss: 82.71749295463317
l2 norm of gradients: 0.9046146777448046
l2 norm of weights: 5.82534219551315

Iteration Number: 409
Loss: 82.66920064068044
l2 norm of gradients: 0.904294693911767
l2 norm of weights: 5.823516207849126

Iteration Number: 410
Loss: 82.62093658592998
l2 norm of gradients: 0.9039747825306333

```

12 norm of weights: 5.8216908180423
-----
Iteration Number: 411
Loss: 82.57270077330705
12 norm of gradients: 0.9036549435294472
12 norm of weights: 5.819866025922197
-----
Iteration Number: 412
Loss: 82.52449318396808
12 norm of gradients: 0.9033351768362391
12 norm of weights: 5.81804183131912
-----
Iteration Number: 413
Loss: 82.47631380321901
12 norm of gradients: 0.9030154823790256
12 norm of weights: 5.8162182340641575
-----
Iteration Number: 414
Loss: 82.42816261369322
12 norm of gradients: 0.9026958600858092
12 norm of weights: 5.814395233989173
-----
Iteration Number: 415
Loss: 82.38003959812153
12 norm of gradients: 0.9023763098845766
12 norm of weights: 5.81257283092681
-----
Iteration Number: 416
Loss: 82.33194474039837
12 norm of gradients: 0.9020568317032998
12 norm of weights: 5.810751024710489
-----
Iteration Number: 417
Loss: 82.28387802185449
12 norm of gradients: 0.9017374254699343
12 norm of weights: 5.8089298151744035
-----
Iteration Number: 418
Loss: 82.23583942701504
12 norm of gradients: 0.9014180911124183
12 norm of weights: 5.807109202153519
-----
Iteration Number: 419
Loss: 82.18782893934862
12 norm of gradients: 0.9010988285586735
12 norm of weights: 5.805289185483576
-----
Iteration Number: 420

```

Loss: 82.1398465426251
l2 norm of gradients: 0.9007796377366032
l2 norm of weights: 5.803469765001083

Iteration Number: 421
Loss: 82.09189221843363
l2 norm of gradients: 0.9004605185740922
l2 norm of weights: 5.801650940543316

Iteration Number: 422
Loss: 82.0439659504836
l2 norm of gradients: 0.9001414709990067
l2 norm of weights: 5.799832711948321

Iteration Number: 423
Loss: 81.99606772366064
l2 norm of gradients: 0.8998224949391926
l2 norm of weights: 5.798015079054908

Iteration Number: 424
Loss: 81.94819752004433
l2 norm of gradients: 0.8995035903224768
l2 norm of weights: 5.796198041702651

Iteration Number: 425
Loss: 81.90035532336975
l2 norm of gradients: 0.8991847570766646
l2 norm of weights: 5.794381599731886

Iteration Number: 426
Loss: 81.85254111763179
l2 norm of gradients: 0.8988659951295404
l2 norm of weights: 5.792565752983712

Iteration Number: 427
Loss: 81.80475488563816
l2 norm of gradients: 0.898547304408867
l2 norm of weights: 5.790750501299987

Iteration Number: 428
Loss: 81.7569966116881
l2 norm of gradients: 0.8982286848423849
l2 norm of weights: 5.788935844523326

Iteration Number: 429
Loss: 81.70926627823717
l2 norm of gradients: 0.8979101363578122
l2 norm of weights: 5.787121782497102

```

-----
Iteration Number: 430
Loss: 81.66156386992272
l2 norm of gradients: 0.8975916588828434
l2 norm of weights: 5.785308315065442
-----
Iteration Number: 431
Loss: 81.6138893698302
l2 norm of gradients: 0.8972732523451493
l2 norm of weights: 5.783495442073225
-----
Iteration Number: 432
Loss: 81.56624276204401
l2 norm of gradients: 0.8969549166723771
l2 norm of weights: 5.78168316336609
-----
Iteration Number: 433
Loss: 81.51862402979125
l2 norm of gradients: 0.8966366517921485
l2 norm of weights: 5.779871478790415
-----
Iteration Number: 434
Loss: 81.47103315687123
l2 norm of gradients: 0.8963184576320603
l2 norm of weights: 5.778060388193336
-----
Iteration Number: 435
Loss: 81.42347012868196
l2 norm of gradients: 0.896000334119684
l2 norm of weights: 5.776249891422733
-----
Iteration Number: 436
Loss: 81.37593492606578
l2 norm of gradients: 0.8956822811825651
l2 norm of weights: 5.774439988327232
-----
Iteration Number: 437
Loss: 81.32842753608917
l2 norm of gradients: 0.8953642987482218
l2 norm of weights: 5.772630678756206
-----
Iteration Number: 438
Loss: 81.28094794003465
l2 norm of gradients: 0.895046386744146
l2 norm of weights: 5.770821962559768
-----
Iteration Number: 439
Loss: 81.23349612276131

```

l2 norm of gradients: 0.8947285450978022
l2 norm of weights: 5.769013839588775

Iteration Number: 440
Loss: 81.18607206847894
l2 norm of gradients: 0.8944107737366267
l2 norm of weights: 5.767206309694824

Iteration Number: 441
Loss: 81.13867576111666
l2 norm of gradients: 0.8940930725880277
l2 norm of weights: 5.76539937273025

Iteration Number: 442
Loss: 81.09130718450864
l2 norm of gradients: 0.8937754415793847
l2 norm of weights: 5.763593028548125

Iteration Number: 443
Loss: 81.04396632232229
l2 norm of gradients: 0.8934578806380482
l2 norm of weights: 5.761787277002257

Iteration Number: 444
Loss: 80.99665315966053
l2 norm of gradients: 0.893140389691339
l2 norm of weights: 5.759982117947188

Iteration Number: 445
Loss: 80.94936767969739
l2 norm of gradients: 0.8928229686665483
l2 norm of weights: 5.758177551238194

Iteration Number: 446
Loss: 80.90210986689614
l2 norm of gradients: 0.8925056174909366
l2 norm of weights: 5.7563735767312805

Iteration Number: 447
Loss: 80.8548797051709
l2 norm of gradients: 0.8921883360917342
l2 norm of weights: 5.754570194283184

Iteration Number: 448
Loss: 80.80767717892506
l2 norm of gradients: 0.89187112439614
l2 norm of weights: 5.752767403751369

Iteration Number: 449
Loss: 80.76050227335945
l2 norm of gradients: 0.8915539823313213
l2 norm of weights: 5.750965204994025

Iteration Number: 450
Loss: 80.71335497079886
l2 norm of gradients: 0.8912369098244145
l2 norm of weights: 5.74916359787007

Iteration Number: 451
Loss: 80.66623525684145
l2 norm of gradients: 0.8909199068025233
l2 norm of weights: 5.747362582239143

Iteration Number: 452
Loss: 80.61914311608653
l2 norm of gradients: 0.8906029731927189
l2 norm of weights: 5.745562157961606

Iteration Number: 453
Loss: 80.57207853214446
l2 norm of gradients: 0.8902861089220397
l2 norm of weights: 5.743762324898543

Iteration Number: 454
Loss: 80.52504148874708
l2 norm of gradients: 0.8899693139174912
l2 norm of weights: 5.741963082911756

Iteration Number: 455
Loss: 80.47803197216915
l2 norm of gradients: 0.8896525881060453
l2 norm of weights: 5.740164431863765

Iteration Number: 456
Loss: 80.43104996620606
l2 norm of gradients: 0.8893359314146407
l2 norm of weights: 5.738366371617807

Iteration Number: 457
Loss: 80.384095453662
l2 norm of gradients: 0.8890193437701812
l2 norm of weights: 5.736568902037832

Iteration Number: 458
Loss: 80.33716842127122
l2 norm of gradients: 0.8887028250995367

```

12 norm of weights: 5.734772022988507
-----
Iteration Number: 459
Loss: 80.29026885337343
12 norm of gradients: 0.8883863753295423
12 norm of weights: 5.732975734335205
-----
Iteration Number: 460
Loss: 80.24339673258964
12 norm of gradients: 0.8880699943869987
12 norm of weights: 5.731180035944015
-----
Iteration Number: 461
Loss: 80.19655204576625
12 norm of gradients: 0.8877536821986708
12 norm of weights: 5.729384927681734
-----
Iteration Number: 462
Loss: 80.14973477603712
12 norm of gradients: 0.8874374386912884
12 norm of weights: 5.727590409415861
-----
Iteration Number: 463
Loss: 80.10294490954524
12 norm of gradients: 0.8871212637915454
12 norm of weights: 5.725796481014607
-----
Iteration Number: 464
Loss: 80.05618242940145
12 norm of gradients: 0.8868051574261001
12 norm of weights: 5.724003142346884
-----
Iteration Number: 465
Loss: 80.0094473212749
12 norm of gradients: 0.8864891195215742
12 norm of weights: 5.722210393282309
-----
Iteration Number: 466
Loss: 79.96273957024879
12 norm of gradients: 0.8861731500045531
12 norm of weights: 5.720418233691199
-----
Iteration Number: 467
Loss: 79.9160591593912
12 norm of gradients: 0.8858572488015858
12 norm of weights: 5.71862666344457
-----
Iteration Number: 468

```


Loss: 79.86940607579659
l2 norm of gradients: 0.8855414158391843
l2 norm of weights: 5.716835682414138

Iteration Number: 469
Loss: 79.82278030260404
l2 norm of gradients: 0.8852256510438233
l2 norm of weights: 5.7150452904723155

Iteration Number: 470
Loss: 79.77618182612605
l2 norm of gradients: 0.8849099543419408
l2 norm of weights: 5.713255487492209

Iteration Number: 471
Loss: 79.72961062963051
l2 norm of gradients: 0.8845943256599367
l2 norm of weights: 5.71146627334762

Iteration Number: 472
Loss: 79.68306669998776
l2 norm of gradients: 0.8842787649241738
l2 norm of weights: 5.7096776479130416

Iteration Number: 473
Loss: 79.63655001992966
l2 norm of gradients: 0.8839632720609768
l2 norm of weights: 5.707889611063658

Iteration Number: 474
Loss: 79.5900605762482
l2 norm of gradients: 0.8836478469966327
l2 norm of weights: 5.706102162675345

Iteration Number: 475
Loss: 79.5435983538854
l2 norm of gradients: 0.8833324896573899
l2 norm of weights: 5.704315302624662

Iteration Number: 476
Loss: 79.49716333716042
l2 norm of gradients: 0.8830171999694592
l2 norm of weights: 5.702529030788855

Iteration Number: 477
Loss: 79.45075551076971
l2 norm of gradients: 0.8827019778590122
l2 norm of weights: 5.70074334704586

```
-----  
Iteration Number: 478  
Loss: 79.40437486112152  
l2 norm of gradients: 0.8823868232521828  
l2 norm of weights: 5.698958251274289  
-----  
Iteration Number: 479  
Loss: 79.35802137303949  
l2 norm of gradients: 0.8820717360750655  
l2 norm of weights: 5.697173743353442  
-----  
Iteration Number: 480  
Loss: 79.31169503058246  
l2 norm of gradients: 0.8817567162537165  
l2 norm of weights: 5.6953898231632945  
-----  
Iteration Number: 481  
Loss: 79.26539582075684  
l2 norm of gradients: 0.8814417637141533  
l2 norm of weights: 5.693606490584504  
-----  
Iteration Number: 482  
Loss: 79.21912372763525  
l2 norm of gradients: 0.8811268783823534  
l2 norm of weights: 5.691823745498403  
-----  
Iteration Number: 483  
Loss: 79.17287873663618  
l2 norm of gradients: 0.8808120601842564  
l2 norm of weights: 5.690041587787  
-----  
Iteration Number: 484  
Loss: 79.12666083308392  
l2 norm of gradients: 0.8804973090457621  
l2 norm of weights: 5.688260017332979  
-----  
Iteration Number: 485  
Loss: 79.0804700025064  
l2 norm of gradients: 0.8801826248927312  
l2 norm of weights: 5.686479034019696  
-----  
Iteration Number: 486  
Loss: 79.03430623006894  
l2 norm of gradients: 0.8798680076509857  
l2 norm of weights: 5.6846986377311755  
-----  
Iteration Number: 487  
Loss: 78.98816950165275
```

```
12 norm of gradients: 0.879553457246307
12 norm of weights: 5.682918828352116
-----
Iteration Number: 488
Loss: 78.94205980190598
12 norm of gradients: 0.879238973604438
12 norm of weights: 5.681139605767879
-----
Iteration Number: 489
Loss: 78.89597711746451
12 norm of gradients: 0.8789245566510825
12 norm of weights: 5.679360969864498
-----
Iteration Number: 490
Loss: 78.8499214315751
12 norm of gradients: 0.8786102063119038
12 norm of weights: 5.6775829205286685
-----
Iteration Number: 491
Loss: 78.8038927323433
12 norm of gradients: 0.8782959225125267
12 norm of weights: 5.675805457647748
-----
Iteration Number: 492
Loss: 78.7578910035485
12 norm of gradients: 0.8779817051785361
12 norm of weights: 5.674028581109756
-----
Iteration Number: 493
Loss: 78.71191623226106
12 norm of gradients: 0.8776675542354775
12 norm of weights: 5.672252290803376
-----
Iteration Number: 494
Loss: 78.665968402259
12 norm of gradients: 0.8773534696088564
12 norm of weights: 5.670476586617948
-----
Iteration Number: 495
Loss: 78.62004750019031
12 norm of gradients: 0.87703945122414
12 norm of weights: 5.6687014684434684
-----
Iteration Number: 496
Loss: 78.57415351213143
12 norm of gradients: 0.8767254990067551
12 norm of weights: 5.66692693617059
-----
```

Iteration Number: 497
Loss: 78.52828642315008
l2 norm of gradients: 0.8764116128820895
l2 norm of weights: 5.66515298969062

Iteration Number: 498
Loss: 78.48244621844756
l2 norm of gradients: 0.8760977927754914
l2 norm of weights: 5.663379628895519

Iteration Number: 499
Loss: 78.43663288495006
l2 norm of gradients: 0.8757840386122698
l2 norm of weights: 5.661606853677897

Iteration Number: 500
Loss: 78.3908464075535
l2 norm of gradients: 0.8754703503176948
l2 norm of weights: 5.659834663931014

Iteration Number: 501
Loss: 78.3450867722937
l2 norm of gradients: 0.8751567278169969
l2 norm of weights: 5.6580630595487795

Iteration Number: 502
Loss: 78.29935396543779
l2 norm of gradients: 0.8748431710353675
l2 norm of weights: 5.656292040425746

Iteration Number: 503
Loss: 78.2536479725535
l2 norm of gradients: 0.874529679897959
l2 norm of weights: 5.654521606457115

Iteration Number: 504
Loss: 78.20796877837675
l2 norm of gradients: 0.8742162543298855
l2 norm of weights: 5.652751757538728

Iteration Number: 505
Loss: 78.16231637092613
l2 norm of gradients: 0.8739028942562213
l2 norm of weights: 5.650982493567071

Iteration Number: 506
Loss: 78.1166907346274
l2 norm of gradients: 0.8735895996020026

```
12 norm of weights: 5.649213814439267
-----
Iteration Number: 507
Loss: 78.0710918558636
12 norm of gradients: 0.8732763702922266
12 norm of weights: 5.64744572005308
-----
Iteration Number: 508
Loss: 78.02551972096882
12 norm of gradients: 0.8729632062518526
12 norm of weights: 5.64567821030691
-----
Iteration Number: 509
Loss: 77.97997431555893
12 norm of gradients: 0.8726501074058016
12 norm of weights: 5.643911285099794
-----
Iteration Number: 510
Loss: 77.93445562552934
12 norm of gradients: 0.8723370736789556
12 norm of weights: 5.6421449443314
-----
Iteration Number: 511
Loss: 77.88896363716752
12 norm of gradients: 0.8720241049961593
12 norm of weights: 5.64037918790203
-----
Iteration Number: 512
Loss: 77.84349833683815
12 norm of gradients: 0.8717112012822196
12 norm of weights: 5.638614015712618
-----
Iteration Number: 513
Loss: 77.79805971018926
12 norm of gradients: 0.8713983624619054
12 norm of weights: 5.636849427664725
-----
Iteration Number: 514
Loss: 77.75264774364948
12 norm of gradients: 0.8710855884599482
12 norm of weights: 5.6350854236605405
-----
Iteration Number: 515
Loss: 77.70726242325044
12 norm of gradients: 0.8707728792010424
12 norm of weights: 5.6333220036028795
-----
Iteration Number: 516
```

Loss: 77.66190373558938
l2 norm of gradients: 0.8704602346098449
l2 norm of weights: 5.631559167395183

Iteration Number: 517
Loss: 77.6165716665721
l2 norm of gradients: 0.8701476546109762
l2 norm of weights: 5.629796914941513

Iteration Number: 518
Loss: 77.57126620252004
l2 norm of gradients: 0.8698351391290199
l2 norm of weights: 5.628035246146553

Iteration Number: 519
Loss: 77.5259873299894
l2 norm of gradients: 0.8695226880885231
l2 norm of weights: 5.626274160915606

Iteration Number: 520
Loss: 77.48073503482989
l2 norm of gradients: 0.869210301413997
l2 norm of weights: 5.624513659154595

Iteration Number: 521
Loss: 77.43550930326238
l2 norm of gradients: 0.8688979790299165
l2 norm of weights: 5.622753740770058

Iteration Number: 522
Loss: 77.39031012274411
l2 norm of gradients: 0.8685857208607208
l2 norm of weights: 5.620994405669147

Iteration Number: 523
Loss: 77.34513747841636
l2 norm of gradients: 0.868273526830814
l2 norm of weights: 5.6192356537596275

Iteration Number: 524
Loss: 77.29999135694581
l2 norm of gradients: 0.8679613968645651
l2 norm of weights: 5.617477484949878

Iteration Number: 525
Loss: 77.25487174546657
l2 norm of gradients: 0.8676493308863075
l2 norm of weights: 5.615719899148886

```

-----
Iteration Number: 526
Loss: 77.20977863036482
l2 norm of gradients: 0.8673373288203408
l2 norm of weights: 5.613962896266249
-----
Iteration Number: 527
Loss: 77.16471199699178
l2 norm of gradients: 0.8670253905909301
l2 norm of weights: 5.612206476212167
-----
Iteration Number: 528
Loss: 77.11967183328896
l2 norm of gradients: 0.8667135161223062
l2 norm of weights: 5.61045063889745
-----
Iteration Number: 529
Loss: 77.07465812532052
l2 norm of gradients: 0.8664017053386667
l2 norm of weights: 5.60869538423351
-----
Iteration Number: 530
Loss: 77.02967085934065
l2 norm of gradients: 0.8660899581641757
l2 norm of weights: 5.606940712132359
-----
Iteration Number: 531
Loss: 76.98471002223906
l2 norm of gradients: 0.865778274522964
l2 norm of weights: 5.605186622506612
-----
Iteration Number: 532
Loss: 76.93977560090089
l2 norm of gradients: 0.8654666543391303
l2 norm of weights: 5.603433115269483
-----
Iteration Number: 533
Loss: 76.89486758117745
l2 norm of gradients: 0.8651550975367405
l2 norm of weights: 5.601680190334778
-----
Iteration Number: 534
Loss: 76.84998595105466
l2 norm of gradients: 0.864843604039829
l2 norm of weights: 5.599927847616906
-----
Iteration Number: 535
Loss: 76.80513069592205

```

l2 norm of gradients: 0.8645321737723982
l2 norm of weights: 5.598176087030862

Iteration Number: 536
Loss: 76.76030180307737
l2 norm of gradients: 0.8642208066584197
l2 norm of weights: 5.59642490849224

Iteration Number: 537
Loss: 76.71549925930685
l2 norm of gradients: 0.863909502621834
l2 norm of weights: 5.594674311917221

Iteration Number: 538
Loss: 76.67072305180689
l2 norm of gradients: 0.8635982615865513
l2 norm of weights: 5.5929242972225754

Iteration Number: 539
Loss: 76.62597316676373
l2 norm of gradients: 0.8632870834764521
l2 norm of weights: 5.59117486432566

Iteration Number: 540
Loss: 76.58124959117764
l2 norm of gradients: 0.8629759682153867
l2 norm of weights: 5.589426013144418

Iteration Number: 541
Loss: 76.53655231212538
l2 norm of gradients: 0.8626649157271769
l2 norm of weights: 5.587677743597378

Iteration Number: 542
Loss: 76.49188131646824
l2 norm of gradients: 0.8623539259356153
l2 norm of weights: 5.585930055603646

Iteration Number: 543
Loss: 76.44723659112003
l2 norm of gradients: 0.8620429987644663
l2 norm of weights: 5.5841829490829165

Iteration Number: 544
Loss: 76.40261812248907
l2 norm of gradients: 0.8617321341374671
l2 norm of weights: 5.582436423955456

Iteration Number: 545
Loss: 76.35802589828424
l2 norm of gradients: 0.8614213319783263
l2 norm of weights: 5.580690480142111

Iteration Number: 546
Loss: 76.31345990500334
l2 norm of gradients: 0.8611105922107267
l2 norm of weights: 5.578945117564302

Iteration Number: 547
Loss: 76.26892013006767
l2 norm of gradients: 0.8607999147583243
l2 norm of weights: 5.577200336144028

Iteration Number: 548
Loss: 76.22440655990503
l2 norm of gradients: 0.8604892995447492
l2 norm of weights: 5.575456135803854

Iteration Number: 549
Loss: 76.17991918260681
l2 norm of gradients: 0.8601787464936058
l2 norm of weights: 5.57371251646692

Iteration Number: 550
Loss: 76.1354579844129
l2 norm of gradients: 0.8598682555284741
l2 norm of weights: 5.571969478056934

Iteration Number: 551
Loss: 76.09102295265208
l2 norm of gradients: 0.8595578265729089
l2 norm of weights: 5.57022702049817

Iteration Number: 552
Loss: 76.04661407442674
l2 norm of gradients: 0.8592474595504418
l2 norm of weights: 5.56848514371547

Iteration Number: 553
Loss: 76.0022313373966
l2 norm of gradients: 0.8589371543845808
l2 norm of weights: 5.5667438476342355

Iteration Number: 554
Loss: 75.9578747281497
l2 norm of gradients: 0.858626910998811

```

12 norm of weights: 5.565003132180434
-----
Iteration Number: 555
Loss: 75.91354423418845
12 norm of gradients: 0.8583167293165952
12 norm of weights: 5.563262997280593
-----
Iteration Number: 556
Loss: 75.86923984232774
12 norm of gradients: 0.8580066092613745
12 norm of weights: 5.561523442861799
-----
Iteration Number: 557
Loss: 75.82496154046879
12 norm of gradients: 0.8576965507565689
12 norm of weights: 5.559784468851692
-----
Iteration Number: 558
Loss: 75.7807093158216
12 norm of gradients: 0.8573865537255779
12 norm of weights: 5.558046075178473
-----
Iteration Number: 559
Loss: 75.73648315538479
12 norm of gradients: 0.8570766180917804
12 norm of weights: 5.556308261770893
-----
Iteration Number: 560
Loss: 75.69228304643136
12 norm of gradients: 0.8567667437785367
12 norm of weights: 5.5545710285582555
-----
Iteration Number: 561
Loss: 75.64810897679929
12 norm of gradients: 0.8564569307091876
12 norm of weights: 5.552834375470415
-----
Iteration Number: 562
Loss: 75.60396093346371
12 norm of gradients: 0.8561471788070563
12 norm of weights: 5.551098302437772
-----
Iteration Number: 563
Loss: 75.5598389041602
12 norm of gradients: 0.855837487995448
12 norm of weights: 5.549362809391279
-----
Iteration Number: 564

```

Loss: 75.51574287565754
l2 norm of gradients: 0.8555278581976509
l2 norm of weights: 5.547627896262428

Iteration Number: 565
Loss: 75.47167283671075
l2 norm of gradients: 0.8552182893369371
l2 norm of weights: 5.545893562983259

Iteration Number: 566
Loss: 75.42762877314055
l2 norm of gradients: 0.8549087813365627
l2 norm of weights: 5.54415980948635

Iteration Number: 567
Loss: 75.38361067384129
l2 norm of gradients: 0.8545993341197691
l2 norm of weights: 5.542426635704821

Iteration Number: 568
Loss: 75.33961852542598
l2 norm of gradients: 0.8542899476097832
l2 norm of weights: 5.54069404157233

Iteration Number: 569
Loss: 75.29565231606855
l2 norm of gradients: 0.8539806217298179
l2 norm of weights: 5.53896202702307

Iteration Number: 570
Loss: 75.25171203336988
l2 norm of gradients: 0.8536713564030736
l2 norm of weights: 5.53723059199177

Iteration Number: 571
Loss: 75.20779766449355
l2 norm of gradients: 0.853362151552738
l2 norm of weights: 5.535499736413692

Iteration Number: 572
Loss: 75.16390919714715
l2 norm of gradients: 0.853053007101987
l2 norm of weights: 5.53376946022463

Iteration Number: 573
Loss: 75.1200466190215
l2 norm of gradients: 0.8527439229739859
l2 norm of weights: 5.532039763360905

```
-----  
Iteration Number: 574  
Loss: 75.07620991785386  
l2 norm of gradients: 0.8524348990918896  
l2 norm of weights: 5.530310645759368  
-----  
Iteration Number: 575  
Loss: 75.03239908153768  
l2 norm of gradients: 0.8521259353788433  
l2 norm of weights: 5.528582107357394  
-----  
Iteration Number: 576  
Loss: 74.98861409766332  
l2 norm of gradients: 0.8518170317579837  
l2 norm of weights: 5.526854148092883  
-----  
Iteration Number: 577  
Loss: 74.9448549537952  
l2 norm of gradients: 0.8515081881524392  
l2 norm of weights: 5.525126767904258  
-----  
Iteration Number: 578  
Loss: 74.90112163780002  
l2 norm of gradients: 0.8511994044853312  
l2 norm of weights: 5.523399966730463  
-----  
Iteration Number: 579  
Loss: 74.85741413756395  
l2 norm of gradients: 0.850890680679774  
l2 norm of weights: 5.52167374451096  
-----  
Iteration Number: 580  
Loss: 74.81373244086635  
l2 norm of gradients: 0.8505820166588768  
l2 norm of weights: 5.519948101185727  
-----  
Iteration Number: 581  
Loss: 74.77007653533587  
l2 norm of gradients: 0.8502734123457428  
l2 norm of weights: 5.518223036695259  
-----  
Iteration Number: 582  
Loss: 74.72644640892153  
l2 norm of gradients: 0.8499648676634721  
l2 norm of weights: 5.5164985509805655  
-----  
Iteration Number: 583  
Loss: 74.68284204965015
```

l2 norm of gradients: 0.8496563825351606
l2 norm of weights: 5.5147746439831655

Iteration Number: 584
Loss: 74.63926344568374
l2 norm of gradients: 0.8493479568839015
l2 norm of weights: 5.513051315645088

Iteration Number: 585
Loss: 74.5957105842285
l2 norm of gradients: 0.8490395906327864
l2 norm of weights: 5.5113285659088715

Iteration Number: 586
Loss: 74.55218345380099
l2 norm of gradients: 0.8487312837049057
l2 norm of weights: 5.509606394717561

Iteration Number: 587
Loss: 74.50868204187498
l2 norm of gradients: 0.8484230360233495
l2 norm of weights: 5.507884802014703

Iteration Number: 588
Loss: 74.46520633687865
l2 norm of gradients: 0.8481148475112086
l2 norm of weights: 5.506163787744353

Iteration Number: 589
Loss: 74.42175632687776
l2 norm of gradients: 0.8478067180915754
l2 norm of weights: 5.504443351851062

Iteration Number: 590
Loss: 74.37833199968541
l2 norm of gradients: 0.8474986476875438
l2 norm of weights: 5.502723494279882

Iteration Number: 591
Loss: 74.33493334316722
l2 norm of gradients: 0.8471906362222121
l2 norm of weights: 5.501004214976362

Iteration Number: 592
Loss: 74.29156034589698
l2 norm of gradients: 0.8468826836186812
l2 norm of weights: 5.499285513886547

Iteration Number: 593
Loss: 74.24821299526027
l2 norm of gradients: 0.8465747898000578
l2 norm of weights: 5.497567390956974

Iteration Number: 594
Loss: 74.20489128038119
l2 norm of gradients: 0.8462669546894539
l2 norm of weights: 5.495849846134677

Iteration Number: 595
Loss: 74.16159518860427
l2 norm of gradients: 0.8459591782099881
l2 norm of weights: 5.494132879367174

Iteration Number: 596
Loss: 74.11832470835711
l2 norm of gradients: 0.8456514602847868
l2 norm of weights: 5.492416490602474

Iteration Number: 597
Loss: 74.07507982777761
l2 norm of gradients: 0.8453438008369843
l2 norm of weights: 5.4907006797890725

Iteration Number: 598
Loss: 74.03186053518488
l2 norm of gradients: 0.8450361997897244
l2 norm of weights: 5.488985446875948

Iteration Number: 599
Loss: 73.98866681868952
l2 norm of gradients: 0.8447286570661613
l2 norm of weights: 5.487270791812565

Iteration Number: 600
Loss: 73.94549866668962
l2 norm of gradients: 0.8444211725894606
l2 norm of weights: 5.485556714548867

Iteration Number: 601
Loss: 73.90235606737507
l2 norm of gradients: 0.8441137462827991
l2 norm of weights: 5.483843215035274

Iteration Number: 602
Loss: 73.85923900896202
l2 norm of gradients: 0.843806378069367

```

12 norm of weights: 5.4821302932226885
-----
Iteration Number: 603
Loss: 73.81614747981372
12 norm of gradients: 0.8434990678723691
12 norm of weights: 5.480417949062484
-----
Iteration Number: 604
Loss: 73.77308146854278
12 norm of gradients: 0.8431918156150239
12 norm of weights: 5.478706182506508
-----
Iteration Number: 605
Loss: 73.73004096313731
12 norm of gradients: 0.8428846212205672
12 norm of weights: 5.476994993507082
-----
Iteration Number: 606
Loss: 73.6870259521937
12 norm of gradients: 0.8425774846122502
12 norm of weights: 5.475284382016997
-----
Iteration Number: 607
Loss: 73.64403642365905
12 norm of gradients: 0.8422704057133428
12 norm of weights: 5.4735743479895085
-----
Iteration Number: 608
Loss: 73.60107236674139
12 norm of gradients: 0.8419633844471337
12 norm of weights: 5.471864891378341
-----
Iteration Number: 609
Loss: 73.55813376922669
12 norm of gradients: 0.8416564207369313
12 norm of weights: 5.470156012137684
-----
Iteration Number: 610
Loss: 73.51522061996575
12 norm of gradients: 0.8413495145060643
12 norm of weights: 5.468447710222185
-----
Iteration Number: 611
Loss: 73.47233290701286
12 norm of gradients: 0.841042665677884
12 norm of weights: 5.4667399855869565
-----
Iteration Number: 612

```

```

Loss: 73.42947061948188
l2 norm of gradients: 0.8407358741757641
l2 norm of weights: 5.465032838187568
-----
Iteration Number: 613
Loss: 73.3866337453105
l2 norm of gradients: 0.8404291399231021
l2 norm of weights: 5.463326267980043
-----
Iteration Number: 614
Loss: 73.34382227352896
l2 norm of gradients: 0.8401224628433206
l2 norm of weights: 5.461620274920865
-----
Iteration Number: 615
Loss: 73.3010361924411
l2 norm of gradients: 0.8398158428598681
l2 norm of weights: 5.459914858966965
-----
Iteration Number: 616
Loss: 73.25827549049777
l2 norm of gradients: 0.8395092798962199
l2 norm of weights: 5.458210020075728
-----
Iteration Number: 617
Loss: 73.21554015669797
l2 norm of gradients: 0.8392027738758795
l2 norm of weights: 5.456505758204987
-----
Iteration Number: 618
Loss: 73.17283017943083
l2 norm of gradients: 0.8388963247223795
l2 norm of weights: 5.454802073313022
-----
Iteration Number: 619
Loss: 73.13014554716281
l2 norm of gradients: 0.8385899323592827
l2 norm of weights: 5.45309896535856
-----
Iteration Number: 620
Loss: 73.08748624889606
l2 norm of gradients: 0.8382835967101829
l2 norm of weights: 5.4513964343007695
-----
Iteration Number: 621
Loss: 73.04485227330373
l2 norm of gradients: 0.8379773176987061
l2 norm of weights: 5.44969448009926

```



```
-----
Iteration Number: 622
Loss: 73.0022436089001
l2 norm of gradients: 0.8376710952485124
l2 norm of weights: 5.447993102714082
-----
Iteration Number: 623
Loss: 72.95966024418942
l2 norm of gradients: 0.8373649292832956
l2 norm of weights: 5.4462923021057215
-----
Iteration Number: 624
Loss: 72.91710216858269
l2 norm of gradients: 0.8370588197267858
l2 norm of weights: 5.444592078235103
-----
Iteration Number: 625
Loss: 72.87456937049812
l2 norm of gradients: 0.8367527665027493
l2 norm of weights: 5.442892431063583
-----
Iteration Number: 626
Loss: 72.83206183848294
l2 norm of gradients: 0.8364467695349904
l2 norm of weights: 5.441193360552947
-----
Iteration Number: 627
Loss: 72.7895795616387
l2 norm of gradients: 0.836140828747353
l2 norm of weights: 5.439494866665417
-----
Iteration Number: 628
Loss: 72.7471225287679
l2 norm of gradients: 0.8358349440637196
l2 norm of weights: 5.437796949363636
-----
Iteration Number: 629
Loss: 72.70469072859359
l2 norm of gradients: 0.8355291154080158
l2 norm of weights: 5.436099608610678
-----
Iteration Number: 630
Loss: 72.66228415035638
l2 norm of gradients: 0.835223342704208
l2 norm of weights: 5.434402844370037
-----
Iteration Number: 631
Loss: 72.61990278222956
```

l2 norm of gradients: 0.8349176258763074
l2 norm of weights: 5.432706656605632

Iteration Number: 632
Loss: 72.57754661382882
l2 norm of gradients: 0.834611964848369
l2 norm of weights: 5.431011045281801

Iteration Number: 633
Loss: 72.53521563356017
l2 norm of gradients: 0.8343063595444942
l2 norm of weights: 5.4293160103633

Iteration Number: 634
Loss: 72.49290983040673
l2 norm of gradients: 0.8340008098888314
l2 norm of weights: 5.4276215518153

Iteration Number: 635
Loss: 72.45062919400017
l2 norm of gradients: 0.8336953158055773
l2 norm of weights: 5.4259276696033885

Iteration Number: 636
Loss: 72.40837371239466
l2 norm of gradients: 0.8333898772189775
l2 norm of weights: 5.4242343636935635

Iteration Number: 637
Loss: 72.36614337505908
l2 norm of gradients: 0.833084494053329
l2 norm of weights: 5.422541634052234

Iteration Number: 638
Loss: 72.32393817118792
l2 norm of gradients: 0.8327791662329799
l2 norm of weights: 5.420849480646217

Iteration Number: 639
Loss: 72.28175808937439
l2 norm of gradients: 0.8324738936823324
l2 norm of weights: 5.419157903442734

Iteration Number: 640
Loss: 72.23960311883793
l2 norm of gradients: 0.8321686763258418
l2 norm of weights: 5.417466902409416

Iteration Number: 641
Loss: 72.19747324864267
l2 norm of gradients: 0.8318635140880195
l2 norm of weights: 5.4157764775142905

Iteration Number: 642
Loss: 72.15536846835175
l2 norm of gradients: 0.8315584068934335
l2 norm of weights: 5.414086628725788

Iteration Number: 643
Loss: 72.11328876628615
l2 norm of gradients: 0.83125335466671
l2 norm of weights: 5.412397356012736

Iteration Number: 644
Loss: 72.07123413192913
l2 norm of gradients: 0.830948357332534
l2 norm of weights: 5.410708659344361

Iteration Number: 645
Loss: 72.029204554687
l2 norm of gradients: 0.8306434148156512
l2 norm of weights: 5.409020538690281

Iteration Number: 646
Loss: 71.98720002313118
l2 norm of gradients: 0.8303385270408689
l2 norm of weights: 5.407332994020507

Iteration Number: 647
Loss: 71.94522052699169
l2 norm of gradients: 0.8300336939330576
l2 norm of weights: 5.405646025305441

Iteration Number: 648
Loss: 71.90326605513283
l2 norm of gradients: 0.8297289154171517
l2 norm of weights: 5.403959632515874

Iteration Number: 649
Loss: 71.86133659691482
l2 norm of gradients: 0.8294241914181513
l2 norm of weights: 5.40227381562298

Iteration Number: 650
Loss: 71.81943214158616
l2 norm of gradients: 0.8291195218611235

```

12 norm of weights: 5.400588574598321
-----
Iteration Number: 651
Loss: 71.77755267851784
12 norm of gradients: 0.8288149066712033
12 norm of weights: 5.398903909413838
-----
Iteration Number: 652
Loss: 71.73569819661661
12 norm of gradients: 0.8285103457735951
12 norm of weights: 5.397219820041856
-----
Iteration Number: 653
Loss: 71.69386868556109
12 norm of gradients: 0.8282058390935736
12 norm of weights: 5.3955363064550745
-----
Iteration Number: 654
Loss: 71.65206413447152
12 norm of gradients: 0.8279013865564864
12 norm of weights: 5.393853368626572
-----
Iteration Number: 655
Loss: 71.61028453242324
12 norm of gradients: 0.8275969880877535
12 norm of weights: 5.392171006529797
-----
Iteration Number: 656
Loss: 71.56852986935138
12 norm of gradients: 0.8272926436128705
12 norm of weights: 5.390489220138575
-----
Iteration Number: 657
Loss: 71.52680013401661
12 norm of gradients: 0.8269883530574077
12 norm of weights: 5.388808009427098
-----
Iteration Number: 658
Loss: 71.48509531623506
12 norm of gradients: 0.8266841163470139
12 norm of weights: 5.387127374369928
-----
Iteration Number: 659
Loss: 71.44341540500184
12 norm of gradients: 0.8263799334074157
12 norm of weights: 5.385447314941992
-----
Iteration Number: 660

```

Loss: 71.40176038984288
l2 norm of gradients: 0.8260758041644204
l2 norm of weights: 5.383767831118581

Iteration Number: 661
Loss: 71.36013026054341
l2 norm of gradients: 0.8257717285439157
l2 norm of weights: 5.382088922875346

Iteration Number: 662
Loss: 71.3185250063969
l2 norm of gradients: 0.8254677064718727
l2 norm of weights: 5.380410590188302

Iteration Number: 663
Loss: 71.27694461651214
l2 norm of gradients: 0.8251637378743468
l2 norm of weights: 5.378732833033818

Iteration Number: 664
Loss: 71.23538908070587
l2 norm of gradients: 0.8248598226774779
l2 norm of weights: 5.37705565138862

Iteration Number: 665
Loss: 71.19385838821506
l2 norm of gradients: 0.824555960807493
l2 norm of weights: 5.375379045229786

Iteration Number: 666
Loss: 71.15235252903697
l2 norm of gradients: 0.8242521521907078
l2 norm of weights: 5.373703014534748

Iteration Number: 667
Loss: 71.11087149211056
l2 norm of gradients: 0.823948396753527
l2 norm of weights: 5.372027559281286

Iteration Number: 668
Loss: 71.06941526707568
l2 norm of gradients: 0.8236446944224465
l2 norm of weights: 5.370352679447528

Iteration Number: 669
Loss: 71.02798384398743
l2 norm of gradients: 0.8233410451240539
l2 norm of weights: 5.368678375011944

```
-----  
Iteration Number: 670  
Loss: 70.98657721204748  
l2 norm of gradients: 0.8230374487850318  
l2 norm of weights: 5.367004645953354  
-----  
Iteration Number: 671  
Loss: 70.94519536066439  
l2 norm of gradients: 0.8227339053321568  
l2 norm of weights: 5.36533149225091  
-----  
Iteration Number: 672  
Loss: 70.90383827998157  
l2 norm of gradients: 0.8224304146923026  
l2 norm of weights: 5.363658913884111  
-----  
Iteration Number: 673  
Loss: 70.86250595904689  
l2 norm of gradients: 0.8221269767924405  
l2 norm of weights: 5.361986910832787  
-----  
Iteration Number: 674  
Loss: 70.82119838798731  
l2 norm of gradients: 0.8218235915596417  
l2 norm of weights: 5.360315483077108  
-----  
Iteration Number: 675  
Loss: 70.77991555602786  
l2 norm of gradients: 0.8215202589210778  
l2 norm of weights: 5.358644630597572  
-----  
Iteration Number: 676  
Loss: 70.73865745330279  
l2 norm of gradients: 0.821216978804023  
l2 norm of weights: 5.35697435337501  
-----  
Iteration Number: 677  
Loss: 70.69742406897122  
l2 norm of gradients: 0.8209137511358547  
l2 norm of weights: 5.355304651390581  
-----  
Iteration Number: 678  
Loss: 70.6562153933571  
l2 norm of gradients: 0.820610575844056  
l2 norm of weights: 5.353635524625769  
-----  
Iteration Number: 679  
Loss: 70.61503141565868
```

l2 norm of gradients: 0.8203074528562163
l2 norm of weights: 5.351966973062385

Iteration Number: 680
Loss: 70.57387212572952
l2 norm of gradients: 0.8200043821000331
l2 norm of weights: 5.350298996682559

Iteration Number: 681
Loss: 70.53273751372065
l2 norm of gradients: 0.8197013635033132
l2 norm of weights: 5.348631595468743

Iteration Number: 682
Loss: 70.49162756878046
l2 norm of gradients: 0.8193983969939749
l2 norm of weights: 5.346964769403706

Iteration Number: 683
Loss: 70.45054228112976
l2 norm of gradients: 0.8190954825000482
l2 norm of weights: 5.345298518470533

Iteration Number: 684
Loss: 70.40948164031381
l2 norm of gradients: 0.8187926199496777
l2 norm of weights: 5.343632842652621

Iteration Number: 685
Loss: 70.36844563634271
l2 norm of gradients: 0.8184898092711231
l2 norm of weights: 5.34196774193368

Iteration Number: 686
Loss: 70.32743425902646
l2 norm of gradients: 0.818187050392761
l2 norm of weights: 5.340303216297731

Iteration Number: 687
Loss: 70.2864474979688
l2 norm of gradients: 0.817884343243086
l2 norm of weights: 5.338639265729098

Iteration Number: 688
Loss: 70.2454853432757
l2 norm of gradients: 0.8175816877507133
l2 norm of weights: 5.336975890212413

Iteration Number: 689
Loss: 70.20454778473294
l2 norm of gradients: 0.8172790838443788
l2 norm of weights: 5.33531308973261

Iteration Number: 690
Loss: 70.16363481243162
l2 norm of gradients: 0.8169765314529414
l2 norm of weights: 5.333650864274922

Iteration Number: 691
Loss: 70.12274641592964
l2 norm of gradients: 0.8166740305053846
l2 norm of weights: 5.331989213824885

Iteration Number: 692
Loss: 70.08188258503873
l2 norm of gradients: 0.8163715809308179
l2 norm of weights: 5.330328138368326

Iteration Number: 693
Loss: 70.04104331027503
l2 norm of gradients: 0.8160691826584774
l2 norm of weights: 5.32866763789137

Iteration Number: 694
Loss: 70.000228581181
l2 norm of gradients: 0.815766835617729
l2 norm of weights: 5.327007712380432

Iteration Number: 695
Loss: 69.95943838788013
l2 norm of gradients: 0.8154645397380684
l2 norm of weights: 5.3253483618222175

Iteration Number: 696
Loss: 69.91867272037184
l2 norm of gradients: 0.8151622949491238
l2 norm of weights: 5.323689586203721

Iteration Number: 697
Loss: 69.8779315682928
l2 norm of gradients: 0.8148601011806563
l2 norm of weights: 5.322031385512219

Iteration Number: 698
Loss: 69.83721492200021
l2 norm of gradients: 0.8145579583625625


```
12 norm of weights: 5.3203737597352765
-----
Iteration Number: 699
Loss: 69.79652277143737
12 norm of gradients: 0.8142558664248756
12 norm of weights: 5.318716708860735
-----
Iteration Number: 700
Loss: 69.75585510661085
12 norm of gradients: 0.8139538252977663
12 norm of weights: 5.317060232876716
-----
Iteration Number: 701
Loss: 69.71521191762477
12 norm of gradients: 0.8136518349115454
12 norm of weights: 5.315404331771621
-----
Iteration Number: 702
Loss: 69.67459319454065
12 norm of gradients: 0.813349895196665
12 norm of weights: 5.313749005534122
-----
Iteration Number: 703
Loss: 69.63399892722069
12 norm of gradients: 0.81304800608372
12 norm of weights: 5.312094254153166
-----
Iteration Number: 704
Loss: 69.59342910608493
12 norm of gradients: 0.8127461675034494
12 norm of weights: 5.310440077617969
-----
Iteration Number: 705
Loss: 69.55288372103665
12 norm of gradients: 0.8124443793867381
12 norm of weights: 5.308786475918016
-----
Iteration Number: 706
Loss: 69.51236276218735
12 norm of gradients: 0.8121426416646187
12 norm of weights: 5.3071334490430555
-----
Iteration Number: 707
Loss: 69.47186621957236
12 norm of gradients: 0.8118409542682723
12 norm of weights: 5.305480996983103
-----
Iteration Number: 708
```

```
Loss: 69.43139408377958
12 norm of gradients: 0.8115393171290313
12 norm of weights: 5.303829119728435
-----
Iteration Number: 709
Loss: 69.39094634442317
12 norm of gradients: 0.81123773017838
12 norm of weights: 5.302177817269583
-----
Iteration Number: 710
Loss: 69.35052299206188
12 norm of gradients: 0.8109361933479562
12 norm of weights: 5.300527089597342
-----
Iteration Number: 711
Loss: 69.31012401660071
12 norm of gradients: 0.8106347065695536
12 norm of weights: 5.298876936702755
-----
Iteration Number: 712
Loss: 69.2697494082809
12 norm of gradients: 0.8103332697751221
12 norm of weights: 5.297227358577123
-----
Iteration Number: 713
Loss: 69.22939915762129
12 norm of gradients: 0.8100318828967709
12 norm of weights: 5.295578355211995
-----
Iteration Number: 714
Loss: 69.18907325455524
12 norm of gradients: 0.8097305458667687
12 norm of weights: 5.293929926599169
-----
Iteration Number: 715
Loss: 69.14877168935227
12 norm of gradients: 0.8094292586175464
12 norm of weights: 5.292282072730687
-----
Iteration Number: 716
Loss: 69.10849445228834
12 norm of gradients: 0.8091280210816979
12 norm of weights: 5.290634793598838
-----
Iteration Number: 717
Loss: 69.06824153362437
12 norm of gradients: 0.8088268331919823
12 norm of weights: 5.288988089196149
```

```

-----
Iteration Number: 718
Loss: 69.02801292360815
l2 norm of gradients: 0.8085256948813249
l2 norm of weights: 5.287341959515389
-----
Iteration Number: 719
Loss: 68.98780861256067
l2 norm of gradients: 0.8082246060828193
l2 norm of weights: 5.2856964045495625
-----
Iteration Number: 720
Loss: 68.94762859083211
l2 norm of gradients: 0.8079235667297289
l2 norm of weights: 5.284051424291909
-----
Iteration Number: 721
Loss: 68.90747284873682
l2 norm of gradients: 0.8076225767554885
l2 norm of weights: 5.282407018735902
-----
Iteration Number: 722
Loss: 68.86734137642283
l2 norm of gradients: 0.8073216360937057
l2 norm of weights: 5.280763187875242
-----
Iteration Number: 723
Loss: 68.82723416449316
l2 norm of gradients: 0.807020744678163
l2 norm of weights: 5.27911993170386
-----
Iteration Number: 724
Loss: 68.78715120297801
l2 norm of gradients: 0.8067199024428189
l2 norm of weights: 5.277477250215914
-----
Iteration Number: 725
Loss: 68.74709248257567
l2 norm of gradients: 0.8064191093218097
l2 norm of weights: 5.275835143405783
-----
Iteration Number: 726
Loss: 68.70705799348922
l2 norm of gradients: 0.8061183652494515
l2 norm of weights: 5.274193611268068
-----
Iteration Number: 727
Loss: 68.66704772615142

```

l2 norm of gradients: 0.8058176701602411
l2 norm of weights: 5.27255265379759

Iteration Number: 728
Loss: 68.62706167082325
l2 norm of gradients: 0.8055170239888582
l2 norm of weights: 5.2709122709893865

Iteration Number: 729
Loss: 68.5870998180668
l2 norm of gradients: 0.805216426670167
l2 norm of weights: 5.269272462838708

Iteration Number: 730
Loss: 68.54716215830355
l2 norm of gradients: 0.8049158781392178
l2 norm of weights: 5.26763322934102

Iteration Number: 731
Loss: 68.50724868192516
l2 norm of gradients: 0.8046153783312483
l2 norm of weights: 5.265994570491996

Iteration Number: 732
Loss: 68.46735937932043
l2 norm of gradients: 0.8043149271816854
l2 norm of weights: 5.264356486287518

Iteration Number: 733
Loss: 68.42749424089634
l2 norm of gradients: 0.8040145246261473
l2 norm of weights: 5.262718976723675

Iteration Number: 734
Loss: 68.38765325737435
l2 norm of gradients: 0.8037141706004448
l2 norm of weights: 5.2610820417967545

Iteration Number: 735
Loss: 68.3478364192141
l2 norm of gradients: 0.8034138650405821
l2 norm of weights: 5.2594456815032515

Iteration Number: 736
Loss: 68.30804371645611
l2 norm of gradients: 0.8031136078827604
l2 norm of weights: 5.257809895839855

Iteration Number: 737
Loss: 68.26827514010877
l2 norm of gradients: 0.8028133990633778
l2 norm of weights: 5.256174684803451

Iteration Number: 738
Loss: 68.22853068050097
l2 norm of gradients: 0.8025132385190316
l2 norm of weights: 5.254540048391123

Iteration Number: 739
Loss: 68.1888103282652
l2 norm of gradients: 0.8022131261865201
l2 norm of weights: 5.252905986600143

Iteration Number: 740
Loss: 68.14911407362733
l2 norm of gradients: 0.8019130620028438
l2 norm of weights: 5.251272499427973

Iteration Number: 741
Loss: 68.10944190740985
l2 norm of gradients: 0.8016130459052078
l2 norm of weights: 5.249639586872264

Iteration Number: 742
Loss: 68.06979382023682
l2 norm of gradients: 0.8013130778310223
l2 norm of weights: 5.248007248930851

Iteration Number: 743
Loss: 68.03016980250496
l2 norm of gradients: 0.8010131577179058
l2 norm of weights: 5.246375485601751

Iteration Number: 744
Loss: 67.99056984466041
l2 norm of gradients: 0.8007132855036854
l2 norm of weights: 5.244744296883165

Iteration Number: 745
Loss: 67.95099393747651
l2 norm of gradients: 0.8004134611263986
l2 norm of weights: 5.243113682773466

Iteration Number: 746
Loss: 67.91144207180884
l2 norm of gradients: 0.8001136845242961

```

12 norm of weights: 5.24148364327121
-----
Iteration Number: 747
Loss: 67.8719142377799
12 norm of gradients: 0.7998139556358421
12 norm of weights: 5.239854178375122
-----
Iteration Number: 748
Loss: 67.83241042625049
12 norm of gradients: 0.7995142743997172
12 norm of weights: 5.2382252880841
-----
Iteration Number: 749
Loss: 67.79293062786071
12 norm of gradients: 0.7992146407548187
12 norm of weights: 5.2365969723972094
-----
Iteration Number: 750
Loss: 67.7534748332408
12 norm of gradients: 0.7989150546402632
12 norm of weights: 5.234969231313686
-----
Iteration Number: 751
Loss: 67.71404303296288
12 norm of gradients: 0.7986155159953884
12 norm of weights: 5.233342064832926
-----
Iteration Number: 752
Loss: 67.67463521778113
12 norm of gradients: 0.7983160247597535
12 norm of weights: 5.23171547295449
-----
Iteration Number: 753
Loss: 67.6352513783597
12 norm of gradients: 0.7980165808731429
12 norm of weights: 5.230089455678099
-----
Iteration Number: 754
Loss: 67.59589150543871
12 norm of gradients: 0.7977171842755659
12 norm of weights: 5.228464013003631
-----
Iteration Number: 755
Loss: 67.55655558953507
12 norm of gradients: 0.79741783490726
12 norm of weights: 5.226839144931118
-----
Iteration Number: 756

```

```
Loss: 67.51724362156959
12 norm of gradients: 0.7971185327086907
12 norm of weights: 5.225214851460747
-----
Iteration Number: 757
Loss: 67.47795559217299
12 norm of gradients: 0.7968192776205553
12 norm of weights: 5.223591132592854
-----
Iteration Number: 758
Loss: 67.43869149189861
12 norm of gradients: 0.7965200695837824
12 norm of weights: 5.221967988327925
-----
Iteration Number: 759
Loss: 67.39945131168112
12 norm of gradients: 0.7962209085395359
12 norm of weights: 5.220345418666591
-----
Iteration Number: 760
Loss: 67.36023504212223
12 norm of gradients: 0.7959217944292141
12 norm of weights: 5.218723423609627
-----
Iteration Number: 761
Loss: 67.32104267411171
12 norm of gradients: 0.795622727194454
12 norm of weights: 5.217102003157951
-----
Iteration Number: 762
Loss: 67.28187419844242
12 norm of gradients: 0.7953237067771308
12 norm of weights: 5.215481157312618
-----
Iteration Number: 763
Loss: 67.24272960566866
12 norm of gradients: 0.7950247331193606
12 norm of weights: 5.21386088607482
-----
Iteration Number: 764
Loss: 67.20360888664906
12 norm of gradients: 0.794725806163502
12 norm of weights: 5.212241189445886
-----
Iteration Number: 765
Loss: 67.16451203219926
12 norm of gradients: 0.7944269258521576
12 norm of weights: 5.210622067427275
```

```

-----
Iteration Number: 766
Loss: 67.1254390331829
l2 norm of gradients: 0.7941280921281758
l2 norm of weights: 5.209003520020577
-----
Iteration Number: 767
Loss: 67.0863898803797
l2 norm of gradients: 0.7938293049346521
l2 norm of weights: 5.207385547227508
-----
Iteration Number: 768
Loss: 67.0473645644735
l2 norm of gradients: 0.7935305642149316
l2 norm of weights: 5.205768149049912
-----
Iteration Number: 769
Loss: 67.00836307651674
l2 norm of gradients: 0.7932318699126094
l2 norm of weights: 5.204151325489755
-----
Iteration Number: 770
Loss: 66.96938540715495
l2 norm of gradients: 0.7929332219715333
l2 norm of weights: 5.202535076549123
-----
Iteration Number: 771
Loss: 66.93043154723752
l2 norm of gradients: 0.792634620335805
l2 norm of weights: 5.2009194022302205
-----
Iteration Number: 772
Loss: 66.89150148767922
l2 norm of gradients: 0.7923360649497825
l2 norm of weights: 5.19930430253537
-----
Iteration Number: 773
Loss: 66.85259521941954
l2 norm of gradients: 0.7920375557580801
l2 norm of weights: 5.197689777467003
-----
Iteration Number: 774
Loss: 66.81371273302229
l2 norm of gradients: 0.7917390927055719
l2 norm of weights: 5.196075827027671
-----
Iteration Number: 775
Loss: 66.77485401973411

```


l2 norm of gradients: 0.7914406757373924
l2 norm of weights: 5.194462451220026

Iteration Number: 776
Loss: 66.73601907012736
l2 norm of gradients: 0.791142304798938
l2 norm of weights: 5.192849650046831

Iteration Number: 777
Loss: 66.69720787536015
l2 norm of gradients: 0.7908439798358695
l2 norm of weights: 5.1912374235109535

Iteration Number: 778
Loss: 66.6584204262766
l2 norm of gradients: 0.7905457007941135
l2 norm of weights: 5.189625771615362

Iteration Number: 779
Loss: 66.61965671354902
l2 norm of gradients: 0.7902474676198631
l2 norm of weights: 5.188014694363128

Iteration Number: 780
Loss: 66.5809167283285
l2 norm of gradients: 0.7899492802595808
l2 norm of weights: 5.186404191757416

Iteration Number: 781
Loss: 66.5422004613419
l2 norm of gradients: 0.7896511386599998
l2 norm of weights: 5.1847942638014874

Iteration Number: 782
Loss: 66.50350790380907
l2 norm of gradients: 0.7893530427681248
l2 norm of weights: 5.1831849104987

Iteration Number: 783
Loss: 66.46483904634151
l2 norm of gradients: 0.7890549925312352
l2 norm of weights: 5.181576131852498

Iteration Number: 784
Loss: 66.42619388010796
l2 norm of gradients: 0.7887569878968846
l2 norm of weights: 5.179967927866416

Iteration Number: 785
Loss: 66.38757239602123
l2 norm of gradients: 0.7884590288129054
l2 norm of weights: 5.178360298544075

Iteration Number: 786
Loss: 66.34897458504962
l2 norm of gradients: 0.788161115227407
l2 norm of weights: 5.176753243889179

Iteration Number: 787
Loss: 66.31040043807214
l2 norm of gradients: 0.7878632470887803
l2 norm of weights: 5.175146763905512

Iteration Number: 788
Loss: 66.27184994621754
l2 norm of gradients: 0.7875654243456973
l2 norm of weights: 5.173540858596939

Iteration Number: 789
Loss: 66.2333231002263
l2 norm of gradients: 0.7872676469471146
l2 norm of weights: 5.1719355279674035

Iteration Number: 790
Loss: 66.19481989141133
l2 norm of gradients: 0.7869699148422731
l2 norm of weights: 5.17033077202092

Iteration Number: 791
Loss: 66.15634031043584
l2 norm of gradients: 0.7866722279807011
l2 norm of weights: 5.168726590761577

Iteration Number: 792
Loss: 66.11788434860924
l2 norm of gradients: 0.7863745863122154
l2 norm of weights: 5.167122984193532

Iteration Number: 793
Loss: 66.07945199663844
l2 norm of gradients: 0.7860769897869221
l2 norm of weights: 5.165519952321011

Iteration Number: 794
Loss: 66.04104324589211
l2 norm of gradients: 0.78577943835522

```

12 norm of weights: 5.163917495148306
-----
Iteration Number: 795
Loss: 66.00265808702264
12 norm of gradients: 0.7854819319678006
12 norm of weights: 5.16231561267977
-----
Iteration Number: 796
Loss: 65.96429651135578
12 norm of gradients: 0.7851844705756501
12 norm of weights: 5.160714304919818
-----
Iteration Number: 797
Loss: 65.92595850984858
12 norm of gradients: 0.7848870541300521
12 norm of weights: 5.159113571872923
-----
Iteration Number: 798
Loss: 65.88764407336048
12 norm of gradients: 0.784589682582587
12 norm of weights: 5.157513413543614
-----
Iteration Number: 799
Loss: 65.84935319323863
12 norm of gradients: 0.784292355885136
12 norm of weights: 5.155913829936475
-----
Iteration Number: 800
Loss: 65.8110858602932
12 norm of gradients: 0.7839950739898811
12 norm of weights: 5.154314821056139
-----
Iteration Number: 801
Loss: 65.77284206577856
12 norm of gradients: 0.7836978368493072
12 norm of weights: 5.15271638690729
-----
Iteration Number: 802
Loss: 65.73462180066535
12 norm of gradients: 0.7834006444162033
12 norm of weights: 5.151118527494659
-----
Iteration Number: 803
Loss: 65.69642505614262
12 norm of gradients: 0.7831034966436651
12 norm of weights: 5.14952124282302
-----
Iteration Number: 804

```

```

Loss: 65.65825182296909
12 norm of gradients: 0.7828063934850953
12 norm of weights: 5.147924532897191
-----
Iteration Number: 805
Loss: 65.62010209270234
12 norm of gradients: 0.7825093348942065
12 norm of weights: 5.146328397722029
-----
Iteration Number: 806
Loss: 65.58197585599892
12 norm of gradients: 0.782212320825021
12 norm of weights: 5.14473283730243
-----
Iteration Number: 807
Loss: 65.54387310428058
12 norm of gradients: 0.7819153512318742
12 norm of weights: 5.143137851643322
-----
Iteration Number: 808
Loss: 65.50579382854477
12 norm of gradients: 0.781618426069415
12 norm of weights: 5.141543440749672
-----
Iteration Number: 809
Loss: 65.46773801979803
12 norm of gradients: 0.7813215452926079
12 norm of weights: 5.139949604626471
-----
Iteration Number: 810
Loss: 65.4297056693296
12 norm of gradients: 0.7810247088567343
12 norm of weights: 5.138356343278743
-----
Iteration Number: 811
Loss: 65.39169676822237
12 norm of gradients: 0.7807279167173938
12 norm of weights: 5.136763656711537
-----
Iteration Number: 812
Loss: 65.35371130740586
12 norm of gradients: 0.7804311688305066
12 norm of weights: 5.135171544929928
-----
Iteration Number: 813
Loss: 65.31574927822993
12 norm of gradients: 0.7801344651523141
12 norm of weights: 5.133580007939007

```

```

-----
Iteration Number: 814
Loss: 65.27781067181027
l2 norm of gradients: 0.7798378056393807
l2 norm of weights: 5.131989045743891
-----
Iteration Number: 815
Loss: 65.23989547933164
l2 norm of gradients: 0.7795411902485957
l2 norm of weights: 5.130398658349711
-----
Iteration Number: 816
Loss: 65.20200369174347
l2 norm of gradients: 0.7792446189371743
l2 norm of weights: 5.128808845761614
-----
Iteration Number: 817
Loss: 65.16413530039671
l2 norm of gradients: 0.7789480916626595
l2 norm of weights: 5.127219607984755
-----
Iteration Number: 818
Loss: 65.12629029643568
l2 norm of gradients: 0.7786516083829239
l2 norm of weights: 5.125630945024308
-----
Iteration Number: 819
Loss: 65.0884686708629
l2 norm of gradients: 0.7783551690561699
l2 norm of weights: 5.124042856885446
-----
Iteration Number: 820
Loss: 65.05067041503426
l2 norm of gradients: 0.7780587736409325
l2 norm of weights: 5.122455343573354
-----
Iteration Number: 821
Loss: 65.01289551998991
l2 norm of gradients: 0.7777624220960805
l2 norm of weights: 5.120868405093219
-----
Iteration Number: 822
Loss: 64.97514397700178
l2 norm of gradients: 0.7774661143808179
l2 norm of weights: 5.119282041450227
-----
Iteration Number: 823
Loss: 64.93741577721266

```

```

12 norm of gradients: 0.7771698504546847
12 norm of weights: 5.117696252649566
-----
Iteration Number: 824
Loss: 64.899710911882
12 norm of gradients: 0.7768736302775598
12 norm of weights: 5.116111038696419
-----
Iteration Number: 825
Loss: 64.86202937207129
12 norm of gradients: 0.7765774538096614
12 norm of weights: 5.114526399595962
-----
Iteration Number: 826
Loss: 64.82437114907505
12 norm of gradients: 0.7762813210115486
12 norm of weights: 5.1129423353533685
-----
Iteration Number: 827
Loss: 64.7867362341324
12 norm of gradients: 0.7759852318441227
12 norm of weights: 5.111358845973797
-----
Iteration Number: 828
Loss: 64.74912461827856
12 norm of gradients: 0.7756891862686299
12 norm of weights: 5.109775931462394
-----
Iteration Number: 829
Loss: 64.71153629281416
12 norm of gradients: 0.7753931842466609
12 norm of weights: 5.108193591824295
-----
Iteration Number: 830
Loss: 64.67397124918476
12 norm of gradients: 0.7750972257401535
12 norm of weights: 5.106611827064616
-----
Iteration Number: 831
Loss: 64.63642947815949
12 norm of gradients: 0.7748013107113941
12 norm of weights: 5.105030637188454
-----
Iteration Number: 832
Loss: 64.59891097130107
12 norm of gradients: 0.7745054391230183
12 norm of weights: 5.103450022200887
-----

```

Iteration Number: 833
Loss: 64.56141571978539
l2 norm of gradients: 0.774209610938013
l2 norm of weights: 5.101869982106965

Iteration Number: 834
Loss: 64.52394371477719
l2 norm of gradients: 0.7739138261197176
l2 norm of weights: 5.100290516911718

Iteration Number: 835
Loss: 64.48649494748273
l2 norm of gradients: 0.7736180846318254
l2 norm of weights: 5.098711626620144

Iteration Number: 836
Loss: 64.44906940918254
l2 norm of gradients: 0.773322386438385
l2 norm of weights: 5.097133311237213

Iteration Number: 837
Loss: 64.41166709108406
l2 norm of gradients: 0.7730267315038022
l2 norm of weights: 5.0955555707678615

Iteration Number: 838
Loss: 64.37428798444407
l2 norm of gradients: 0.77273111979284
l2 norm of weights: 5.093978405216992

Iteration Number: 839
Loss: 64.33693208064813
l2 norm of gradients: 0.7724355512706216
l2 norm of weights: 5.09240181458947

Iteration Number: 840
Loss: 64.29959937085555
l2 norm of gradients: 0.7721400259026306
l2 norm of weights: 5.0908257988901235

Iteration Number: 841
Loss: 64.2622898461892
l2 norm of gradients: 0.7718445436547131
l2 norm of weights: 5.089250358123737

Iteration Number: 842
Loss: 64.2250034980718
l2 norm of gradients: 0.7715491044930787

```

12 norm of weights: 5.087675492295051
-----
Iteration Number: 843
Loss: 64.18774031774986
12 norm of gradients: 0.7712537083843014
12 norm of weights: 5.086101201408765
-----
Iteration Number: 844
Loss: 64.15050029634833
12 norm of gradients: 0.7709583552953223
12 norm of weights: 5.084527485469526
-----
Iteration Number: 845
Loss: 64.11328342537124
12 norm of gradients: 0.7706630451934496
12 norm of weights: 5.082954344481934
-----
Iteration Number: 846
Loss: 64.07608969591213
12 norm of gradients: 0.7703677780463605
12 norm of weights: 5.081381778450534
-----
Iteration Number: 847
Loss: 64.03891909926956
12 norm of gradients: 0.7700725538221019
12 norm of weights: 5.079809787379821
-----
Iteration Number: 848
Loss: 64.00177162668176
12 norm of gradients: 0.7697773724890931
12 norm of weights: 5.078238371274229
-----
Iteration Number: 849
Loss: 63.964647269511296
12 norm of gradients: 0.7694822340161255
12 norm of weights: 5.076667530138135
-----
Iteration Number: 850
Loss: 63.92754601910534
12 norm of gradients: 0.769187138372365
12 norm of weights: 5.075097263975857
-----
Iteration Number: 851
Loss: 63.89046786661415
12 norm of gradients: 0.768892085527353
12 norm of weights: 5.073527572791647
-----
Iteration Number: 852

```



```

Loss: 63.85341280334384
l2 norm of gradients: 0.768597075451007
l2 norm of weights: 5.071958456589695
-----
Iteration Number: 853
Loss: 63.8163808206922
l2 norm of gradients: 0.7683021081136228
l2 norm of weights: 5.070389915374122
-----
Iteration Number: 854
Loss: 63.77937190977448
l2 norm of gradients: 0.7680071834858753
l2 norm of weights: 5.068821949148978
-----
Iteration Number: 855
Loss: 63.74238606206494
l2 norm of gradients: 0.7677123015388199
l2 norm of weights: 5.067254557918244
-----
Iteration Number: 856
Loss: 63.70542326876542
l2 norm of gradients: 0.7674174622438936
l2 norm of weights: 5.065687741685826
-----
Iteration Number: 857
Loss: 63.66848352111335
l2 norm of gradients: 0.7671226655729164
l2 norm of weights: 5.064121500455555
-----
Iteration Number: 858
Loss: 63.63156681055416
l2 norm of gradients: 0.7668279114980924
l2 norm of weights: 5.0625558342311825
-----
Iteration Number: 859
Loss: 63.59467312835658
l2 norm of gradients: 0.7665331999920106
l2 norm of weights: 5.060990743016382
-----
Iteration Number: 860
Loss: 63.55780246576631
l2 norm of gradients: 0.7662385310276475
l2 norm of weights: 5.059426226814742
-----
Iteration Number: 861
Loss: 63.520954814028286
l2 norm of gradients: 0.7659439045783665
l2 norm of weights: 5.05786228562977

```

```

-----
Iteration Number: 862
Loss: 63.48413016457589
l2 norm of gradients: 0.7656493206179197
l2 norm of weights: 5.056298919464882
-----
Iteration Number: 863
Loss: 63.44732850875852
l2 norm of gradients: 0.7653547791204504
l2 norm of weights: 5.0547361283234125
-----
Iteration Number: 864
Loss: 63.410549837708075
l2 norm of gradients: 0.7650602800604921
l2 norm of weights: 5.053173912208599
-----
Iteration Number: 865
Loss: 63.3737941429633
l2 norm of gradients: 0.7647658234129713
l2 norm of weights: 5.051612271123589
-----
Iteration Number: 866
Loss: 63.33706141564177
l2 norm of gradients: 0.7644714091532079
l2 norm of weights: 5.050051205071435
-----
Iteration Number: 867
Loss: 63.3003516470961
l2 norm of gradients: 0.7641770372569163
l2 norm of weights: 5.0484907140550925
-----
Iteration Number: 868
Loss: 63.26366482878837
l2 norm of gradients: 0.7638827077002068
l2 norm of weights: 5.046930798077417
-----
Iteration Number: 869
Loss: 63.22700095185727
l2 norm of gradients: 0.7635884204595871
l2 norm of weights: 5.045371457141164
-----
Iteration Number: 870
Loss: 63.19036000770094
l2 norm of gradients: 0.7632941755119621
l2 norm of weights: 5.043812691248987
-----
Iteration Number: 871
Loss: 63.15374198776938

```

l2 norm of gradients: 0.7629999728346367
l2 norm of weights: 5.0422545004034305

Iteration Number: 872
Loss: 63.117146883164665
l2 norm of gradients: 0.7627058124053154
l2 norm of weights: 5.040696884606935

Iteration Number: 873
Loss: 63.08057468526076
l2 norm of gradients: 0.7624116942021043
l2 norm of weights: 5.03913984386183

Iteration Number: 874
Loss: 63.044025385468245
l2 norm of gradients: 0.762117618203512
l2 norm of weights: 5.037583378170334

Iteration Number: 875
Loss: 63.00749897513271
l2 norm of gradients: 0.7618235843884499
l2 norm of weights: 5.036027487534553

Iteration Number: 876
Loss: 62.97099544541536
l2 norm of gradients: 0.7615295927362346
l2 norm of weights: 5.034472171956476

Iteration Number: 877
Loss: 62.93451478783068
l2 norm of gradients: 0.7612356432265878
l2 norm of weights: 5.032917431437975

Iteration Number: 878
Loss: 62.89805699365726
l2 norm of gradients: 0.7609417358396381
l2 norm of weights: 5.031363265980805

Iteration Number: 879
Loss: 62.86162205411982
l2 norm of gradients: 0.7606478705559212
l2 norm of weights: 5.029809675586594

Iteration Number: 880
Loss: 62.82520996062547
l2 norm of gradients: 0.7603540473563813
l2 norm of weights: 5.028256660256851

Iteration Number: 881
Loss: 62.78882070455027
l2 norm of gradients: 0.7600602662223729
l2 norm of weights: 5.026704219992959

Iteration Number: 882
Loss: 62.752454277215406
l2 norm of gradients: 0.7597665271356601
l2 norm of weights: 5.02515235479617

Iteration Number: 883
Loss: 62.71611066986154
l2 norm of gradients: 0.759472830078419
l2 norm of weights: 5.023601064667611

Iteration Number: 884
Loss: 62.67978987382247
l2 norm of gradients: 0.7591791750332381
l2 norm of weights: 5.022050349608275

Iteration Number: 885
Loss: 62.6434918805988
l2 norm of gradients: 0.7588855619831192
l2 norm of weights: 5.020500209619021

Iteration Number: 886
Loss: 62.60721668139576
l2 norm of gradients: 0.7585919909114783
l2 norm of weights: 5.018950644700575

Iteration Number: 887
Loss: 62.57096426746973
l2 norm of gradients: 0.7582984618021473
l2 norm of weights: 5.017401654853522

Iteration Number: 888
Loss: 62.53473463028994
l2 norm of gradients: 0.7580049746393734
l2 norm of weights: 5.015853240078311

Iteration Number: 889
Loss: 62.49852776115251
l2 norm of gradients: 0.7577115294078213
l2 norm of weights: 5.014305400375247

Iteration Number: 890
Loss: 62.46234365135483
l2 norm of gradients: 0.7574181260925739

```
12 norm of weights: 5.012758135744494
-----
Iteration Number: 891
Loss: 62.42618229229419
12 norm of gradients: 0.7571247646791323
12 norm of weights: 5.011211446186069
-----
Iteration Number: 892
Loss: 62.39004367521143
12 norm of gradients: 0.756831445153418
12 norm of weights: 5.009665331699843
-----
Iteration Number: 893
Loss: 62.353927791526225
12 norm of gradients: 0.7565381675017728
12 norm of weights: 5.0081197922855365
-----
Iteration Number: 894
Loss: 62.31783463253746
12 norm of gradients: 0.7562449317109599
12 norm of weights: 5.00657482794272
-----
Iteration Number: 895
Loss: 62.28176418951452
12 norm of gradients: 0.7559517377681645
12 norm of weights: 5.005030438670812
-----
Iteration Number: 896
Loss: 62.24571645390013
12 norm of gradients: 0.755658585660995
12 norm of weights: 5.003486624469074
-----
Iteration Number: 897
Loss: 62.209691417025475
12 norm of gradients: 0.7553654753774843
12 norm of weights: 5.001943385336613
-----
Iteration Number: 898
Loss: 62.17368907000889
12 norm of gradients: 0.755072406906089
12 norm of weights: 5.000400721272374
-----
Iteration Number: 899
Loss: 62.13770940444166
12 norm of gradients: 0.7547793802356915
12 norm of weights: 4.998858632275146
-----
Iteration Number: 900
```

Loss: 62.101752411551715
l2 norm of gradients: 0.754486395355601
l2 norm of weights: 4.9973171183435525

Iteration Number: 901
Loss: 62.06581808260498
l2 norm of gradients: 0.7541934522555529
l2 norm of weights: 4.995776179476053

Iteration Number: 902
Loss: 62.02990640900788
l2 norm of gradients: 0.7539005509257106
l2 norm of weights: 4.994235815670943

Iteration Number: 903
Loss: 61.994017382123324
l2 norm of gradients: 0.753607691356666
l2 norm of weights: 4.992696026926348

Iteration Number: 904
Loss: 61.95815099315161
l2 norm of gradients: 0.7533148735394403
l2 norm of weights: 4.991156813240223

Iteration Number: 905
Loss: 61.922307233471884
l2 norm of gradients: 0.7530220974654844
l2 norm of weights: 4.989618174610355

Iteration Number: 906
Loss: 61.88648609449492
l2 norm of gradients: 0.7527293631266799
l2 norm of weights: 4.9880801110343524

Iteration Number: 907
Loss: 61.850687567408514
l2 norm of gradients: 0.7524366705153399
l2 norm of weights: 4.9865426225096545

Iteration Number: 908
Loss: 61.81491164359202
l2 norm of gradients: 0.752144019624209
l2 norm of weights: 4.985005709033518

Iteration Number: 909
Loss: 61.779158314376545
l2 norm of gradients: 0.7518514104464644
l2 norm of weights: 4.983469370603023

```

-----
Iteration Number: 910
Loss: 61.74342757107216
l2 norm of gradients: 0.7515588429757172
l2 norm of weights: 4.98193360721507
-----
Iteration Number: 911
Loss: 61.70771940495169
l2 norm of gradients: 0.7512663172060118
l2 norm of weights: 4.9803984188663755
-----
Iteration Number: 912
Loss: 61.67203380743381
l2 norm of gradients: 0.7509738331318272
l2 norm of weights: 4.97886380555347
-----
Iteration Number: 913
Loss: 61.63637076969025
l2 norm of gradients: 0.7506813907480777
l2 norm of weights: 4.977329767272702
-----
Iteration Number: 914
Loss: 61.600730283204676
l2 norm of gradients: 0.7503889900501133
l2 norm of weights: 4.97579630402023
-----
Iteration Number: 915
Loss: 61.56511233918655
l2 norm of gradients: 0.7500966310337203
l2 norm of weights: 4.974263415792023
-----
Iteration Number: 916
Loss: 61.52951692888803
l2 norm of gradients: 0.7498043136951213
l2 norm of weights: 4.972731102583859
-----
Iteration Number: 917
Loss: 61.49394404367124
l2 norm of gradients: 0.7495120380309774
l2 norm of weights: 4.971199364391323
-----
Iteration Number: 918
Loss: 61.458393674935664
l2 norm of gradients: 0.7492198040383868
l2 norm of weights: 4.969668201209804
-----
Iteration Number: 919
Loss: 61.422865813841014

```

l2 norm of gradients: 0.7489276117148865
l2 norm of weights: 4.9681376130344965

Iteration Number: 920
Loss: 61.387360451824556
l2 norm of gradients: 0.7486354610584525
l2 norm of weights: 4.966607599860395

Iteration Number: 921
Loss: 61.35187758004535
l2 norm of gradients: 0.7483433520675001
l2 norm of weights: 4.965078161682296

Iteration Number: 922
Loss: 61.31641718993674
l2 norm of gradients: 0.748051284740885
l2 norm of weights: 4.963549298494793

Iteration Number: 923
Loss: 61.28097927266874
l2 norm of gradients: 0.7477592590779029
l2 norm of weights: 4.962021010292276

Iteration Number: 924
Loss: 61.245563819642406
l2 norm of gradients: 0.7474672750782909
l2 norm of weights: 4.960493297068932

Iteration Number: 925
Loss: 61.2101708220581
l2 norm of gradients: 0.7471753327422272
l2 norm of weights: 4.958966158818738

Iteration Number: 926
Loss: 61.17480027132053
l2 norm of gradients: 0.7468834320703319
l2 norm of weights: 4.957439595535464

Iteration Number: 927
Loss: 61.13945215853664
l2 norm of gradients: 0.7465915730636671
l2 norm of weights: 4.955913607212674

Iteration Number: 928
Loss: 61.10412647521586
l2 norm of gradients: 0.7462997557237379
l2 norm of weights: 4.954388193843713

Iteration Number: 929
Loss: 61.06882321246793
l2 norm of gradients: 0.7460079800524924
l2 norm of weights: 4.952863355421719

Iteration Number: 930
Loss: 61.0335423616421
l2 norm of gradients: 0.7457162460523217
l2 norm of weights: 4.951339091939613

Iteration Number: 931
Loss: 60.99828391399268
l2 norm of gradients: 0.7454245537260612
l2 norm of weights: 4.949815403390098

Iteration Number: 932
Loss: 60.96304786081559
l2 norm of gradients: 0.7451329030769902
l2 norm of weights: 4.948292289765661

Iteration Number: 933
Loss: 60.92783419340485
l2 norm of gradients: 0.7448412941088327
l2 norm of weights: 4.946769751058569

Iteration Number: 934
Loss: 60.89264290299789
l2 norm of gradients: 0.7445497268257569
l2 norm of weights: 4.945247787260866

Iteration Number: 935
Loss: 60.85747398086144
l2 norm of gradients: 0.7442582012323766
l2 norm of weights: 4.943726398364374

Iteration Number: 936
Loss: 60.82232741826168
l2 norm of gradients: 0.743966717333751
l2 norm of weights: 4.942205584360693

Iteration Number: 937
Loss: 60.7872032064902
l2 norm of gradients: 0.7436752751353849
l2 norm of weights: 4.940685345241191

Iteration Number: 938
Loss: 60.75210133671712
l2 norm of gradients: 0.7433838746432289

```

12 norm of weights: 4.939165680997017
-----
Iteration Number: 939
Loss: 60.717021800315
12 norm of gradients: 0.74309251586368
12 norm of weights: 4.9376465916190835
-----
Iteration Number: 940
Loss: 60.6819645884633
12 norm of gradients: 0.7428011988035818
12 norm of weights: 4.936128077098075
-----
Iteration Number: 941
Loss: 60.64692969243556
12 norm of gradients: 0.7425099234702242
12 norm of weights: 4.934610137424443
-----
Iteration Number: 942
Loss: 60.61191710349817
12 norm of gradients: 0.7422186898713438
12 norm of weights: 4.933092772588408
-----
Iteration Number: 943
Loss: 60.576926812902286
12 norm of gradients: 0.7419274980151253
12 norm of weights: 4.931575982579953
-----
Iteration Number: 944
Loss: 60.541958811817636
12 norm of gradients: 0.7416363479101995
12 norm of weights: 4.930059767388824
-----
Iteration Number: 945
Loss: 60.507013091488695
12 norm of gradients: 0.7413452395656454
12 norm of weights: 4.928544127004528
-----
Iteration Number: 946
Loss: 60.47208964322812
12 norm of gradients: 0.7410541729909891
12 norm of weights: 4.927029061416337
-----
Iteration Number: 947
Loss: 60.437188458238104
12 norm of gradients: 0.7407631481962049
12 norm of weights: 4.925514570613277
-----
Iteration Number: 948

```

Loss: 60.40230952768594
l2 norm of gradients: 0.7404721651917145
l2 norm of weights: 4.924000654584133

Iteration Number: 949
Loss: 60.36745284287339
l2 norm of gradients: 0.7401812239883881
l2 norm of weights: 4.922487313317446

Iteration Number: 950
Loss: 60.33261839498602
l2 norm of gradients: 0.7398903245975432
l2 norm of weights: 4.920974546801511

Iteration Number: 951
Loss: 60.297806175236296
l2 norm of gradients: 0.7395994670309464
l2 norm of weights: 4.919462355024378

Iteration Number: 952
Loss: 60.26301617488369
l2 norm of gradients: 0.7393086513008116
l2 norm of weights: 4.917950737973847

Iteration Number: 953
Loss: 60.22824838506056
l2 norm of gradients: 0.7390178774198016
l2 norm of weights: 4.916439695637467

Iteration Number: 954
Loss: 60.19350279705835
l2 norm of gradients: 0.7387271454010272
l2 norm of weights: 4.914929228002539

Iteration Number: 955
Loss: 60.15877940196203
l2 norm of gradients: 0.7384364552580478
l2 norm of weights: 4.913419335056108

Iteration Number: 956
Loss: 60.124078191121704
l2 norm of gradients: 0.7381458070048709
l2 norm of weights: 4.911910016784969

Iteration Number: 957
Loss: 60.08939915561557
l2 norm of gradients: 0.7378552006559524
l2 norm of weights: 4.910401273175658

```

-----
Iteration Number: 958
Loss: 60.05474228667125
l2 norm of gradients: 0.7375646362261967
l2 norm of weights: 4.908893104214457
-----
Iteration Number: 959
Loss: 60.02010757548208
l2 norm of gradients: 0.7372741137309564
l2 norm of weights: 4.907385509887388
-----
Iteration Number: 960
Loss: 59.985495013257
l2 norm of gradients: 0.7369836331860322
l2 norm of weights: 4.905878490180214
-----
Iteration Number: 961
Loss: 59.95090459111809
l2 norm of gradients: 0.7366931946076732
l2 norm of weights: 4.90437204507844
-----
Iteration Number: 962
Loss: 59.916336300285806
l2 norm of gradients: 0.7364027980125766
l2 norm of weights: 4.902866174567305
-----
Iteration Number: 963
Loss: 59.88179013195144
l2 norm of gradients: 0.7361124434178874
l2 norm of weights: 4.901360878631786
-----
Iteration Number: 964
Loss: 59.847266077155595
l2 norm of gradients: 0.7358221308411987
l2 norm of weights: 4.899856157256599
-----
Iteration Number: 965
Loss: 59.812764127209256
l2 norm of gradients: 0.7355318603005516
l2 norm of weights: 4.898352010426188
-----
Iteration Number: 966
Loss: 59.778284273250506
l2 norm of gradients: 0.7352416318144345
l2 norm of weights: 4.896848438124734
-----
Iteration Number: 967
Loss: 59.7438265063767

```

l2 norm of gradients: 0.7349514454017834
l2 norm of weights: 4.895345440336148

Iteration Number: 968
Loss: 59.70939081771311
l2 norm of gradients: 0.7346613010819818
l2 norm of weights: 4.893843017044072

Iteration Number: 969
Loss: 59.67497719850758
l2 norm of gradients: 0.7343711988748604
l2 norm of weights: 4.8923411682318765

Iteration Number: 970
Loss: 59.64058563985223
l2 norm of gradients: 0.7340811388006968
l2 norm of weights: 4.890839893882661

Iteration Number: 971
Loss: 59.606216132855884
l2 norm of gradients: 0.7337911208802153
l2 norm of weights: 4.889339193979249

Iteration Number: 972
Loss: 59.57186866870107
l2 norm of gradients: 0.7335011451345866
l2 norm of weights: 4.887839068504192

Iteration Number: 973
Loss: 59.53754323846293
l2 norm of gradients: 0.7332112115854286
l2 norm of weights: 4.886339517439764

Iteration Number: 974
Loss: 59.50323983330314
l2 norm of gradients: 0.7329213202548044
l2 norm of weights: 4.884840540767963

Iteration Number: 975
Loss: 59.46895844433666
l2 norm of gradients: 0.7326314711652229
l2 norm of weights: 4.883342138470509

Iteration Number: 976
Loss: 59.43469906270569
l2 norm of gradients: 0.7323416643396393
l2 norm of weights: 4.881844310528839

Iteration Number: 977
Loss: 59.40046167945577
l2 norm of gradients: 0.7320518998014531
l2 norm of weights: 4.880347056924115

Iteration Number: 978
Loss: 59.36624628573725
l2 norm of gradients: 0.7317621775745096
l2 norm of weights: 4.878850377637213

Iteration Number: 979
Loss: 59.332052872684564
l2 norm of gradients: 0.7314724976830981
l2 norm of weights: 4.877354272648729

Iteration Number: 980
Loss: 59.297881431260855
l2 norm of gradients: 0.7311828601519524
l2 norm of weights: 4.875858741938972

Iteration Number: 981
Loss: 59.26373195272467
l2 norm of gradients: 0.7308932650062504
l2 norm of weights: 4.874363785487969

Iteration Number: 982
Loss: 59.22960442804233
l2 norm of gradients: 0.7306037122716131
l2 norm of weights: 4.872869403275458

Iteration Number: 983
Loss: 59.19549884838041
l2 norm of gradients: 0.7303142019741055
l2 norm of weights: 4.871375595280892

Iteration Number: 984
Loss: 59.16141520474444
l2 norm of gradients: 0.7300247341402344
l2 norm of weights: 4.869882361483433

Iteration Number: 985
Loss: 59.12735348822063
l2 norm of gradients: 0.7297353087969499
l2 norm of weights: 4.868389701861956

Iteration Number: 986
Loss: 59.093313689913
l2 norm of gradients: 0.729445925971643

```

12 norm of weights: 4.866897616395043
-----
Iteration Number: 987
Loss: 59.059295800847934
12 norm of gradients: 0.7291565856921473
12 norm of weights: 4.865406105060987
-----
Iteration Number: 988
Loss: 59.025299812084185
12 norm of gradients: 0.7288672879867372
12 norm of weights: 4.863915167837786
-----
Iteration Number: 989
Loss: 58.99132571468374
12 norm of gradients: 0.728578032884127
12 norm of weights: 4.862424804703145
-----
Iteration Number: 990
Loss: 58.95737349972826
12 norm of gradients: 0.7282888204134722
12 norm of weights: 4.860935015634475
-----
Iteration Number: 991
Loss: 58.923443158199184
12 norm of gradients: 0.7279996506043672
12 norm of weights: 4.859445800608889
-----
Iteration Number: 992
Loss: 58.88953468116008
12 norm of gradients: 0.7277105234868457
12 norm of weights: 4.857957159603206
-----
Iteration Number: 993
Loss: 58.855648059655465
12 norm of gradients: 0.7274214390913802
12 norm of weights: 4.8564690925939455
-----
Iteration Number: 994
Loss: 58.82178328466672
12 norm of gradients: 0.7271323974488813
12 norm of weights: 4.854981599557328
-----
Iteration Number: 995
Loss: 58.78794034726232
12 norm of gradients: 0.7268433985906965
12 norm of weights: 4.853494680469275
-----
Iteration Number: 996

```

```
Loss: 58.75411923842684
l2 norm of gradients: 0.7265544425486111
l2 norm of weights: 4.852008335305407
-----
Iteration Number: 997
Loss: 58.72031994917738
l2 norm of gradients: 0.7262655293548462
l2 norm of weights: 4.850522564041043
-----
Iteration Number: 998
Loss: 58.686542470528615
l2 norm of gradients: 0.7259766590420591
l2 norm of weights: 4.849037366651198
-----
Iteration Number: 999
Loss: 58.652786793442715
l2 norm of gradients: 0.725687831643342
l2 norm of weights: 4.847552743110587
-----
Iteration Number: 1000
Loss: 58.61905290896117
l2 norm of gradients: 0.7253990471922218
l2 norm of weights: 4.846068693393616
-----
Iteration Number: 1001
Loss: 58.58534080807193
l2 norm of gradients: 0.7251103057226592
l2 norm of weights: 4.844585217474387
-----
Iteration Number: 1002
Loss: 58.55165048171466
l2 norm of gradients: 0.7248216072690483
l2 norm of weights: 4.8431023153267
-----
Iteration Number: 1003
Loss: 58.517981920921386
l2 norm of gradients: 0.7245329518662159
l2 norm of weights: 4.84161998692404
-----
Iteration Number: 1004
Loss: 58.48433511660723
l2 norm of gradients: 0.7242443395494205
l2 norm of weights: 4.84013823223959
-----
Iteration Number: 1005
Loss: 58.45071005974456
l2 norm of gradients: 0.7239557703543523
l2 norm of weights: 4.838657051246222
```



```

-----
Iteration Number: 1006
Loss: 58.41710674135757
l2 norm of gradients: 0.7236672443171316
l2 norm of weights: 4.837176443916498
-----
Iteration Number: 1007
Loss: 58.38352515228897
l2 norm of gradients: 0.7233787614743086
l2 norm of weights: 4.835696410222669
-----
Iteration Number: 1008
Loss: 58.34996528361278
l2 norm of gradients: 0.7230903218628628
l2 norm of weights: 4.834216950136676
-----
Iteration Number: 1009
Loss: 58.31642712617558
l2 norm of gradients: 0.7228019255202021
l2 norm of weights: 4.8327380636301465
-----
Iteration Number: 1010
Loss: 58.282910670953896
l2 norm of gradients: 0.7225135724841619
l2 norm of weights: 4.831259750674396
-----
Iteration Number: 1011
Loss: 58.24941590885297
l2 norm of gradients: 0.7222252627930045
l2 norm of weights: 4.8297820112404235
-----
Iteration Number: 1012
Loss: 58.21594283086329
l2 norm of gradients: 0.7219369964854181
l2 norm of weights: 4.8283048452989155
-----
Iteration Number: 1013
Loss: 58.18249142782604
l2 norm of gradients: 0.7216487736005162
l2 norm of weights: 4.826828252820244
-----
Iteration Number: 1014
Loss: 58.14906169070316
l2 norm of gradients: 0.721360594177837
l2 norm of weights: 4.825352233774461
-----
Iteration Number: 1015
Loss: 58.115653610419294

```

l2 norm of gradients: 0.7210724582573422
l2 norm of weights: 4.8238767881313045

Iteration Number: 1016
Loss: 58.08226717781561
l2 norm of gradients: 0.7207843658794162
l2 norm of weights: 4.822401915860194

Iteration Number: 1017
Loss: 58.04890238381207
l2 norm of gradients: 0.7204963170848653
l2 norm of weights: 4.82092761693023

Iteration Number: 1018
Loss: 58.01555921930815
l2 norm of gradients: 0.7202083119149171
l2 norm of weights: 4.819453891310193

Iteration Number: 1019
Loss: 57.982237675182475
l2 norm of gradients: 0.7199203504112193
l2 norm of weights: 4.817980738968546

Iteration Number: 1020
Loss: 57.94893774228816
l2 norm of gradients: 0.7196324326158385
l2 norm of weights: 4.81650815987343

Iteration Number: 1021
Loss: 57.91565941156258
l2 norm of gradients: 0.7193445585712606
l2 norm of weights: 4.815036153992662

Iteration Number: 1022
Loss: 57.88240267381571
l2 norm of gradients: 0.7190567283203884
l2 norm of weights: 4.813564721293741

Iteration Number: 1023
Loss: 57.84916751992095
l2 norm of gradients: 0.7187689419065411
l2 norm of weights: 4.812093861743841

Iteration Number: 1024
Loss: 57.8159539407493
l2 norm of gradients: 0.718481199373454
l2 norm of weights: 4.810623575309813

Iteration Number: 1025
Loss: 57.78276192713305
l2 norm of gradients: 0.718193500765277
l2 norm of weights: 4.809153861958182

Iteration Number: 1026
Loss: 57.74959146990218
l2 norm of gradients: 0.7179058461265732
l2 norm of weights: 4.8076847216551535

Iteration Number: 1027
Loss: 57.71644255989503
l2 norm of gradients: 0.717618235502319
l2 norm of weights: 4.806216154366602

Iteration Number: 1028
Loss: 57.68331518795892
l2 norm of gradients: 0.7173306689379022
l2 norm of weights: 4.80474816005808

Iteration Number: 1029
Loss: 57.65020934490296
l2 norm of gradients: 0.7170431464791217
l2 norm of weights: 4.803280738694809

Iteration Number: 1030
Loss: 57.61712502154639
l2 norm of gradients: 0.7167556681721862
l2 norm of weights: 4.8018138902416885

Iteration Number: 1031
Loss: 57.58406220871928
l2 norm of gradients: 0.7164682340637117
l2 norm of weights: 4.800347614663286

Iteration Number: 1032
Loss: 57.55102089714797
l2 norm of gradients: 0.7161808442007238
l2 norm of weights: 4.798881911923843

Iteration Number: 1033
Loss: 57.51800107771376
l2 norm of gradients: 0.7158934986306535
l2 norm of weights: 4.797416781987271

Iteration Number: 1034
Loss: 57.48500274121174
l2 norm of gradients: 0.7156061974013379

```

12 norm of weights: 4.795952224817151
-----
Iteration Number: 1035
Loss: 57.45202587835363
12 norm of gradients: 0.7153189405610181
12 norm of weights: 4.7944882403767375
-----
Iteration Number: 1036
Loss: 57.41907047999322
12 norm of gradients: 0.7150317281583386
12 norm of weights: 4.793024828628949
-----
Iteration Number: 1037
Loss: 57.386136536831856
12 norm of gradients: 0.7147445602423464
12 norm of weights: 4.791561989536379
-----
Iteration Number: 1038
Loss: 57.3532240396663
12 norm of gradients: 0.7144574368624896
12 norm of weights: 4.790099723061283
-----
Iteration Number: 1039
Loss: 57.3203329793007
12 norm of gradients: 0.7141703580686164
12 norm of weights: 4.78863802916559
-----
Iteration Number: 1040
Loss: 57.28746334641867
12 norm of gradients: 0.7138833239109736
12 norm of weights: 4.787176907810892
-----
Iteration Number: 1041
Loss: 57.25461513179783
12 norm of gradients: 0.713596334440206
12 norm of weights: 4.78571635895845
-----
Iteration Number: 1042
Loss: 57.2217883261643
12 norm of gradients: 0.7133093897073547
12 norm of weights: 4.7842563825691915
-----
Iteration Number: 1043
Loss: 57.18898292028744
12 norm of gradients: 0.7130224897638564
12 norm of weights: 4.7827969786037094
-----
Iteration Number: 1044

```

Loss: 57.15619890485739
l2 norm of gradients: 0.712735634661542
l2 norm of weights: 4.781338147022261

Iteration Number: 1045
Loss: 57.123436270613716
l2 norm of gradients: 0.7124488244526356
l2 norm of weights: 4.77987988778477

Iteration Number: 1046
Loss: 57.090695008206715
l2 norm of gradients: 0.7121620591897525
l2 norm of weights: 4.778422200850825

Iteration Number: 1047
Loss: 57.05797510847646
l2 norm of gradients: 0.7118753389258996
l2 norm of weights: 4.776965086179675

Iteration Number: 1048
Loss: 57.025276562004514
l2 norm of gradients: 0.711588663714472
l2 norm of weights: 4.775508543730238

Iteration Number: 1049
Loss: 56.99259935950198
l2 norm of gradients: 0.711302033609254
l2 norm of weights: 4.774052573461093

Iteration Number: 1050
Loss: 56.9599434917122
l2 norm of gradients: 0.7110154486644161
l2 norm of weights: 4.77259717533048

Iteration Number: 1051
Loss: 56.92730894925228
l2 norm of gradients: 0.7107289089345146
l2 norm of weights: 4.771142349296304

Iteration Number: 1052
Loss: 56.89469572282486
l2 norm of gradients: 0.7104424144744901
l2 norm of weights: 4.769688095316131

Iteration Number: 1053
Loss: 56.86210380309968
l2 norm of gradients: 0.7101559653396664
l2 norm of weights: 4.768234413347189

```

-----
Iteration Number: 1054
Loss: 56.8295331807297
l2 norm of gradients: 0.709869561585749
l2 norm of weights: 4.766781303346368
-----
Iteration Number: 1055
Loss: 56.79698384638657
l2 norm of gradients: 0.7095832032688241
l2 norm of weights: 4.765328765270218
-----
Iteration Number: 1056
Loss: 56.76445579069835
l2 norm of gradients: 0.7092968904453565
l2 norm of weights: 4.76387679907495
-----
Iteration Number: 1057
Loss: 56.73194900430233
l2 norm of gradients: 0.7090106231721892
l2 norm of weights: 4.762425404716435
-----
Iteration Number: 1058
Loss: 56.69946347783145
l2 norm of gradients: 0.7087244015065417
l2 norm of weights: 4.760974582150205
-----
Iteration Number: 1059
Loss: 56.66699920190998
l2 norm of gradients: 0.7084382255060085
l2 norm of weights: 4.759524331331453
-----
Iteration Number: 1060
Loss: 56.634556167212075
l2 norm of gradients: 0.7081520952285583
l2 norm of weights: 4.758074652215027
-----
Iteration Number: 1061
Loss: 56.602134364260245
l2 norm of gradients: 0.7078660107325312
l2 norm of weights: 4.756625544755439
-----
Iteration Number: 1062
Loss: 56.56973378372778
l2 norm of gradients: 0.7075799720766394
l2 norm of weights: 4.7551770089068555
-----
Iteration Number: 1063
Loss: 56.53735441619676

```

l2 norm of gradients: 0.7072939793199646
l2 norm of weights: 4.753729044623107

Iteration Number: 1064
Loss: 56.50499625224359
l2 norm of gradients: 0.707008032521956
l2 norm of weights: 4.752281651857676

Iteration Number: 1065
Loss: 56.472659282472506
l2 norm of gradients: 0.7067221317424306
l2 norm of weights: 4.75083483056371

Iteration Number: 1066
Loss: 56.440343497451714
l2 norm of gradients: 0.7064362770415701
l2 norm of weights: 4.749388580694008

Iteration Number: 1067
Loss: 56.408048887743384
l2 norm of gradients: 0.7061504684799209
l2 norm of weights: 4.7479429022010295

Iteration Number: 1068
Loss: 56.375775443954765
l2 norm of gradients: 0.7058647061183911
l2 norm of weights: 4.746497795036893

Iteration Number: 1069
Loss: 56.34352315660792
l2 norm of gradients: 0.7055789900182509
l2 norm of weights: 4.745053259153372

Iteration Number: 1070
Loss: 56.31129201626363
l2 norm of gradients: 0.7052933202411296
l2 norm of weights: 4.743609294501897

Iteration Number: 1071
Loss: 56.27908201347575
l2 norm of gradients: 0.7050076968490143
l2 norm of weights: 4.742165901033555

Iteration Number: 1072
Loss: 56.24689313874876
l2 norm of gradients: 0.70472211990425
l2 norm of weights: 4.740723078699093

Iteration Number: 1073
Loss: 56.21472538267164
l2 norm of gradients: 0.704436589469536
l2 norm of weights: 4.739280827448908

Iteration Number: 1074
Loss: 56.18257873572831
l2 norm of gradients: 0.7041511056079259
l2 norm of weights: 4.737839147233061

Iteration Number: 1075
Loss: 56.15045318841602
l2 norm of gradients: 0.7038656683828252
l2 norm of weights: 4.736398038001263

Iteration Number: 1076
Loss: 56.11834873129534
l2 norm of gradients: 0.7035802778579903
l2 norm of weights: 4.734957499702885

Iteration Number: 1077
Loss: 56.08626535485238
l2 norm of gradients: 0.703294934097527
l2 norm of weights: 4.7335175322869505

Iteration Number: 1078
Loss: 56.054203049592495
l2 norm of gradients: 0.7030096371658884
l2 norm of weights: 4.7320781357021415

Iteration Number: 1079
Loss: 56.02216180597368
l2 norm of gradients: 0.7027243871278744
l2 norm of weights: 4.7306393098967945

Iteration Number: 1080
Loss: 55.990141614490994
l2 norm of gradients: 0.702439184048629
l2 norm of weights: 4.729201054818901

Iteration Number: 1081
Loss: 55.95814246562746
l2 norm of gradients: 0.7021540279936396
l2 norm of weights: 4.727763370416109

Iteration Number: 1082
Loss: 55.9261643498316
l2 norm of gradients: 0.7018689190287352


```

12 norm of weights: 4.726326256635719
-----
Iteration Number: 1083
Loss: 55.89420725761029
12 norm of gradients: 0.7015838572200842
12 norm of weights: 4.724889713424692
-----
Iteration Number: 1084
Loss: 55.86227117939911
12 norm of gradients: 0.7012988426341943
12 norm of weights: 4.7234537407296395
-----
Iteration Number: 1085
Loss: 55.83035610561225
12 norm of gradients: 0.7010138753379088
12 norm of weights: 4.72201833849683
-----
Iteration Number: 1086
Loss: 55.79846202672908
12 norm of gradients: 0.7007289553984078
12 norm of weights: 4.720583506672186
-----
Iteration Number: 1087
Loss: 55.76658893321069
12 norm of gradients: 0.7004440828832036
12 norm of weights: 4.7191492452012875
-----
Iteration Number: 1088
Loss: 55.73473681539972
12 norm of gradients: 0.700159257860141
12 norm of weights: 4.717715554029366
-----
Iteration Number: 1089
Loss: 55.70290566378782
12 norm of gradients: 0.6998744803973959
12 norm of weights: 4.716282433101309
-----
Iteration Number: 1090
Loss: 55.67109546873425
12 norm of gradients: 0.6995897505634715
12 norm of weights: 4.71484988236166
-----
Iteration Number: 1091
Loss: 55.63930622068347
12 norm of gradients: 0.6993050684271993
12 norm of weights: 4.713417901754618
-----
Iteration Number: 1092

```

Loss: 55.60753791005926
l2 norm of gradients: 0.6990204340577363
l2 norm of weights: 4.711986491224035

Iteration Number: 1093
Loss: 55.57579052718192
l2 norm of gradients: 0.6987358475245627
l2 norm of weights: 4.7105556507134185

Iteration Number: 1094
Loss: 55.544064062504624
l2 norm of gradients: 0.6984513088974815
l2 norm of weights: 4.709125380165931

Iteration Number: 1095
Loss: 55.51235850636315
l2 norm of gradients: 0.6981668182466161
l2 norm of weights: 4.707695679524389

Iteration Number: 1096
Loss: 55.48067384913196
l2 norm of gradients: 0.6978823756424084
l2 norm of weights: 4.706266548731266

Iteration Number: 1097
Loss: 55.44901008120931
l2 norm of gradients: 0.6975979811556182
l2 norm of weights: 4.704837987728689

Iteration Number: 1098
Loss: 55.417367192915926
l2 norm of gradients: 0.6973136348573206
l2 norm of weights: 4.703409996458442

Iteration Number: 1099
Loss: 55.38574517462912
l2 norm of gradients: 0.6970293368189041
l2 norm of weights: 4.70198257486196

Iteration Number: 1100
Loss: 55.354144016657536
l2 norm of gradients: 0.6967450871120702
l2 norm of weights: 4.700555722880337

Iteration Number: 1101
Loss: 55.322563709365944
l2 norm of gradients: 0.69646088580883
l2 norm of weights: 4.699129440454322

```

-----
Iteration Number: 1102
Loss: 55.29100424308881
l2 norm of gradients: 0.6961767329815042
l2 norm of weights: 4.697703727524316
-----
Iteration Number: 1103
Loss: 55.25946560811497
l2 norm of gradients: 0.6958926287027201
l2 norm of weights: 4.696278584030379
-----
Iteration Number: 1104
Loss: 55.227947794804315
l2 norm of gradients: 0.6956085730454101
l2 norm of weights: 4.694854009912225
-----
Iteration Number: 1105
Loss: 55.1964507934457
l2 norm of gradients: 0.6953245660828109
l2 norm of weights: 4.6934300051092235
-----
Iteration Number: 1106
Loss: 55.16497459432913
l2 norm of gradients: 0.6950406078884604
l2 norm of weights: 4.692006569560401
-----
Iteration Number: 1107
Loss: 55.13351918777087
l2 norm of gradients: 0.6947566985361977
l2 norm of weights: 4.690583703204437
-----
Iteration Number: 1108
Loss: 55.102084564027436
l2 norm of gradients: 0.6944728381001591
l2 norm of weights: 4.689161405979671
-----
Iteration Number: 1109
Loss: 55.07067071341754
l2 norm of gradients: 0.694189026654778
l2 norm of weights: 4.687739677824095
-----
Iteration Number: 1110
Loss: 55.0392776261826
l2 norm of gradients: 0.6939052642747829
l2 norm of weights: 4.6863185186753595
-----
Iteration Number: 1111
Loss: 55.00790529261652

```

l2 norm of gradients: 0.6936215510351958
l2 norm of weights: 4.684897928470772

Iteration Number: 1112
Loss: 54.97655370297506
l2 norm of gradients: 0.6933378870113295
l2 norm of weights: 4.683477907147292

Iteration Number: 1113
Loss: 54.945222847496154
l2 norm of gradients: 0.6930542722787866
l2 norm of weights: 4.682058454641542

Iteration Number: 1114
Loss: 54.91391271642568
l2 norm of gradients: 0.6927707069134574
l2 norm of weights: 4.680639570889797

Iteration Number: 1115
Loss: 54.882623300019944
l2 norm of gradients: 0.692487190991519
l2 norm of weights: 4.67922125582799

Iteration Number: 1116
Loss: 54.851354588502176
l2 norm of gradients: 0.6922037245894324
l2 norm of weights: 4.677803509391714

Iteration Number: 1117
Loss: 54.82010657210837
l2 norm of gradients: 0.6919203077839408
l2 norm of weights: 4.676386331516214

Iteration Number: 1118
Loss: 54.788879241013845
l2 norm of gradients: 0.6916369406520682
l2 norm of weights: 4.674969722136399

Iteration Number: 1119
Loss: 54.75767258549067
l2 norm of gradients: 0.691353623271118
l2 norm of weights: 4.673553681186832

Iteration Number: 1120
Loss: 54.72648659570648
l2 norm of gradients: 0.6910703557186708
l2 norm of weights: 4.6721382086017345

Iteration Number: 1121
Loss: 54.69532126185192
l2 norm of gradients: 0.6907871380725817
l2 norm of weights: 4.670723304314987

Iteration Number: 1122
Loss: 54.664176574165104
l2 norm of gradients: 0.6905039704109802
l2 norm of weights: 4.66930896826013

Iteration Number: 1123
Loss: 54.63305252278381
l2 norm of gradients: 0.6902208528122672
l2 norm of weights: 4.667895200370359

Iteration Number: 1124
Loss: 54.6019490978953
l2 norm of gradients: 0.6899377853551132
l2 norm of weights: 4.666482000578533

Iteration Number: 1125
Loss: 54.570866289675436
l2 norm of gradients: 0.6896547681184573
l2 norm of weights: 4.665069368817168

Iteration Number: 1126
Loss: 54.53980408830583
l2 norm of gradients: 0.6893718011815044
l2 norm of weights: 4.66365730501844

Iteration Number: 1127
Loss: 54.508762483893925
l2 norm of gradients: 0.689088884623724
l2 norm of weights: 4.662245809114182

Iteration Number: 1128
Loss: 54.477741466643835
l2 norm of gradients: 0.6888060185248485
l2 norm of weights: 4.660834881035895

Iteration Number: 1129
Loss: 54.446741026672285
l2 norm of gradients: 0.6885232029648705
l2 norm of weights: 4.659424520714732

Iteration Number: 1130
Loss: 54.41576115411277
l2 norm of gradients: 0.6882404380240417

```

12 norm of weights: 4.658014728081511
-----
Iteration Number: 1131
Loss: 54.38480183910428
12 norm of gradients: 0.6879577237828709
12 norm of weights: 4.65660550306671
-----
Iteration Number: 1132
Loss: 54.35386307173811
12 norm of gradients: 0.6876750603221221
12 norm of weights: 4.65519684560047
-----
Iteration Number: 1133
Loss: 54.3229448421753
12 norm of gradients: 0.6873924477228132
12 norm of weights: 4.65378875561259
-----
Iteration Number: 1134
Loss: 54.29204714051369
12 norm of gradients: 0.6871098860662125
12 norm of weights: 4.652381233032535
-----
Iteration Number: 1135
Loss: 54.26116995683182
12 norm of gradients: 0.6868273754338385
12 norm of weights: 4.65097427778943
-----
Iteration Number: 1136
Loss: 54.23031328124985
12 norm of gradients: 0.6865449159074583
12 norm of weights: 4.649567889812063
-----
Iteration Number: 1137
Loss: 54.199477103833374
12 norm of gradients: 0.6862625075690835
12 norm of weights: 4.648162069028885
-----
Iteration Number: 1138
Loss: 54.16866141469328
12 norm of gradients: 0.6859801505009705
12 norm of weights: 4.646756815368011
-----
Iteration Number: 1139
Loss: 54.13786620386013
12 norm of gradients: 0.6856978447856182
12 norm of weights: 4.645352128757219
-----
Iteration Number: 1140

```

```
Loss: 54.10709146143668
12 norm of gradients: 0.6854155905057653
12 norm of weights: 4.643948009123951
-----
Iteration Number: 1141
Loss: 54.07633717749896
12 norm of gradients: 0.6851333877443891
12 norm of weights: 4.642544456395314
-----
Iteration Number: 1142
Loss: 54.04560334206871
12 norm of gradients: 0.6848512365847037
12 norm of weights: 4.641141470498079
-----
Iteration Number: 1143
Loss: 54.0148899452077
12 norm of gradients: 0.6845691371101574
12 norm of weights: 4.639739051358682
-----
Iteration Number: 1144
Loss: 53.98419697695026
12 norm of gradients: 0.684287089404432
12 norm of weights: 4.638337198903227
-----
Iteration Number: 1145
Loss: 53.95352442734809
12 norm of gradients: 0.6840050935514396
12 norm of weights: 4.63693591305748
-----
Iteration Number: 1146
Loss: 53.9228722863894
12 norm of gradients: 0.6837231496353217
12 norm of weights: 4.635535193746875
-----
Iteration Number: 1147
Loss: 53.89224054414501
12 norm of gradients: 0.683441257740447
12 norm of weights: 4.634135040896515
-----
Iteration Number: 1148
Loss: 53.86162919059036
12 norm of gradients: 0.6831594179514092
12 norm of weights: 4.632735454431168
-----
Iteration Number: 1149
Loss: 53.83103821575587
12 norm of gradients: 0.6828776303530256
12 norm of weights: 4.6313364342752665
```

```

-----
Iteration Number: 1150
Loss: 53.800467609649374
l2 norm of gradients: 0.6825958950303354
l2 norm of weights: 4.629937980352917
-----
Iteration Number: 1151
Loss: 53.76991736225152
l2 norm of gradients: 0.6823142120685962
l2 norm of weights: 4.628540092587892
-----
Iteration Number: 1152
Loss: 53.73938746353294
l2 norm of gradients: 0.6820325815532846
l2 norm of weights: 4.627142770903631
-----
Iteration Number: 1153
Loss: 53.708877903522065
l2 norm of gradients: 0.6817510035700923
l2 norm of weights: 4.6257460152232435
-----
Iteration Number: 1154
Loss: 53.67838867215054
l2 norm of gradients: 0.6814694782049255
l2 norm of weights: 4.624349825469511
-----
Iteration Number: 1155
Loss: 53.64791975940159
l2 norm of gradients: 0.6811880055439014
l2 norm of weights: 4.622954201564882
-----
Iteration Number: 1156
Loss: 53.617471155243834
l2 norm of gradients: 0.6809065856733482
l2 norm of weights: 4.621559143431478
-----
Iteration Number: 1157
Loss: 53.587042849624716
l2 norm of gradients: 0.680625218679802
l2 norm of weights: 4.62016465099109
-----
Iteration Number: 1158
Loss: 53.55663483250061
l2 norm of gradients: 0.6803439046500055
l2 norm of weights: 4.618770724165182
-----
Iteration Number: 1159
Loss: 53.52624709380859

```


l2 norm of gradients: 0.6800626436709053
l2 norm of weights: 4.617377362874888

Iteration Number: 1160
Loss: 53.495879623491994
l2 norm of gradients: 0.6797814358296508
l2 norm of weights: 4.615984567041016

Iteration Number: 1161
Loss: 53.46553241145239
l2 norm of gradients: 0.679500281213592
l2 norm of weights: 4.614592336584047

Iteration Number: 1162
Loss: 53.43520544764466
l2 norm of gradients: 0.6792191799102778
l2 norm of weights: 4.613200671424134

Iteration Number: 1163
Loss: 53.40489872196521
l2 norm of gradients: 0.6789381320074535
l2 norm of weights: 4.611809571481108

Iteration Number: 1164
Loss: 53.3746122243558
l2 norm of gradients: 0.6786571375930593
l2 norm of weights: 4.610419036674467

Iteration Number: 1165
Loss: 53.34434594469913
l2 norm of gradients: 0.6783761967552289
l2 norm of weights: 4.609029066923391

Iteration Number: 1166
Loss: 53.31409987286773
l2 norm of gradients: 0.6780953095822865
l2 norm of weights: 4.607639662146733

Iteration Number: 1167
Loss: 53.2838739987831
l2 norm of gradients: 0.6778144761627456
l2 norm of weights: 4.6062508222630205

Iteration Number: 1168
Loss: 53.25366831234179
l2 norm of gradients: 0.6775336965853072
l2 norm of weights: 4.604862547190459

Iteration Number: 1169
Loss: 53.22348280339554
l2 norm of gradients: 0.6772529709388573
l2 norm of weights: 4.603474836846932

Iteration Number: 1170
Loss: 53.19331746181066
l2 norm of gradients: 0.6769722993124658
l2 norm of weights: 4.602087691149996

Iteration Number: 1171
Loss: 53.16317227747234
l2 norm of gradients: 0.6766916817953836
l2 norm of weights: 4.600701110016892

Iteration Number: 1172
Loss: 53.13304724023681
l2 norm of gradients: 0.6764111184770419
l2 norm of weights: 4.5993150933645355

Iteration Number: 1173
Loss: 53.10294233995867
l2 norm of gradients: 0.6761306094470492
l2 norm of weights: 4.5979296411095225

Iteration Number: 1174
Loss: 53.07285756646605
l2 norm of gradients: 0.6758501547951895
l2 norm of weights: 4.596544753168127

Iteration Number: 1175
Loss: 53.04279290963791
l2 norm of gradients: 0.6755697546114213
l2 norm of weights: 4.595160429456305

Iteration Number: 1176
Loss: 53.01274835926504
l2 norm of gradients: 0.675289408985875
l2 norm of weights: 4.593776669889693

Iteration Number: 1177
Loss: 52.982723905186916
l2 norm of gradients: 0.6750091180088509
l2 norm of weights: 4.592393474383608

Iteration Number: 1178
Loss: 52.95271953725046
l2 norm of gradients: 0.6747288817708178

```

12 norm of weights: 4.591010842853052
-----
Iteration Number: 1179
Loss: 52.92273524526165
12 norm of gradients: 0.6744487003624103
12 norm of weights: 4.589628775212703
-----
Iteration Number: 1180
Loss: 52.89277101901188
12 norm of gradients: 0.674168573874428
12 norm of weights: 4.588247271376931
-----
Iteration Number: 1181
Loss: 52.86282684832187
12 norm of gradients: 0.6738885023978326
12 norm of weights: 4.58686633125978
-----
Iteration Number: 1182
Loss: 52.83290272295961
12 norm of gradients: 0.6736084860237467
12 norm of weights: 4.5854859547749856
-----
Iteration Number: 1183
Loss: 52.802998632770496
12 norm of gradients: 0.6733285248434512
12 norm of weights: 4.584106141835965
-----
Iteration Number: 1184
Loss: 52.77311456750329
12 norm of gradients: 0.6730486189483842
12 norm of weights: 4.582726892355823
-----
Iteration Number: 1185
Loss: 52.74325051693349
12 norm of gradients: 0.6727687684301386
12 norm of weights: 4.5813482062473465
-----
Iteration Number: 1186
Loss: 52.71340647085145
12 norm of gradients: 0.6724889733804602
12 norm of weights: 4.579970083423013
-----
Iteration Number: 1187
Loss: 52.68358241901336
12 norm of gradients: 0.6722092338912461
12 norm of weights: 4.578592523794985
-----
Iteration Number: 1188

```

Loss: 52.653778351186595
l2 norm of gradients: 0.6719295500545427
l2 norm of weights: 4.577215527275113

Iteration Number: 1189
Loss: 52.62399425712317
l2 norm of gradients: 0.6716499219625431
l2 norm of weights: 4.575839093774937

Iteration Number: 1190
Loss: 52.59423012657134
l2 norm of gradients: 0.6713703497075868
l2 norm of weights: 4.574463223205686

Iteration Number: 1191
Loss: 52.564485949269226
l2 norm of gradients: 0.6710908333821561
l2 norm of weights: 4.573087915478277

Iteration Number: 1192
Loss: 52.53476171496318
l2 norm of gradients: 0.6708113730788756
l2 norm of weights: 4.57171317050332

Iteration Number: 1193
Loss: 52.50505741339883
l2 norm of gradients: 0.6705319688905091
l2 norm of weights: 4.570338988191113

Iteration Number: 1194
Loss: 52.47537303426809
l2 norm of gradients: 0.670252620909959
l2 norm of weights: 4.568965368451648

Iteration Number: 1195
Loss: 52.445708567326115
l2 norm of gradients: 0.6699733292302629
l2 norm of weights: 4.567592311194607

Iteration Number: 1196
Loss: 52.41606400224313
l2 norm of gradients: 0.6696940939445932
l2 norm of weights: 4.566219816329367

Iteration Number: 1197
Loss: 52.3864393287768
l2 norm of gradients: 0.6694149151462547
l2 norm of weights: 4.5648478837649975

```

-----
Iteration Number: 1198
Loss: 52.35683453659744
l2 norm of gradients: 0.669135792928682
l2 norm of weights: 4.5634765134102615
-----
Iteration Number: 1199
Loss: 52.32724961540601
l2 norm of gradients: 0.6688567273854384
l2 norm of weights: 4.562105705173618
-----
Iteration Number: 1200
Loss: 52.297684554890324
l2 norm of gradients: 0.6685777186102144
l2 norm of weights: 4.5607354589632205
-----
Iteration Number: 1201
Loss: 52.26813934475086
l2 norm of gradients: 0.6682987666968249
l2 norm of weights: 4.5593657746869205
-----
Iteration Number: 1202
Loss: 52.23861397465954
l2 norm of gradients: 0.6680198717392077
l2 norm of weights: 4.557996652252262
-----
Iteration Number: 1203
Loss: 52.20910843428689
l2 norm of gradients: 0.6677410338314221
l2 norm of weights: 4.556628091566492
-----
Iteration Number: 1204
Loss: 52.17962271329371
l2 norm of gradients: 0.6674622530676458
l2 norm of weights: 4.555260092536551
-----
Iteration Number: 1205
Loss: 52.150156801354086
l2 norm of gradients: 0.6671835295421749
l2 norm of weights: 4.553892655069082
-----
Iteration Number: 1206
Loss: 52.120710688131865
l2 norm of gradients: 0.6669048633494201
l2 norm of weights: 4.552525779070424
-----
Iteration Number: 1207
Loss: 52.09128436325821

```

l2 norm of gradients: 0.6666262545839066
l2 norm of weights: 4.551159464446618

Iteration Number: 1208
Loss: 52.0618778163932
l2 norm of gradients: 0.6663477033402708
l2 norm of weights: 4.549793711103407

Iteration Number: 1209
Loss: 52.0324910371655
l2 norm of gradients: 0.6660692097132593
l2 norm of weights: 4.548428518946231

Iteration Number: 1210
Loss: 52.00312401523796
l2 norm of gradients: 0.665790773797727
l2 norm of weights: 4.547063887880238

Iteration Number: 1211
Loss: 51.97377674019152
l2 norm of gradients: 0.6655123956886346
l2 norm of weights: 4.545699817810276

Iteration Number: 1212
Loss: 51.944449201693224
l2 norm of gradients: 0.665234075481048
l2 norm of weights: 4.544336308640896

Iteration Number: 1213
Loss: 51.915141389332646
l2 norm of gradients: 0.6649558132701351
l2 norm of weights: 4.542973360276354

Iteration Number: 1214
Loss: 51.88585329273203
l2 norm of gradients: 0.664677609151165
l2 norm of weights: 4.541610972620611

Iteration Number: 1215
Loss: 51.85658490151411
l2 norm of gradients: 0.6643994632195054
l2 norm of weights: 4.540249145577334

Iteration Number: 1216
Loss: 51.82733620525844
l2 norm of gradients: 0.6641213755706216
l2 norm of weights: 4.538887879049897

Iteration Number: 1217
Loss: 51.79810719356748
l2 norm of gradients: 0.663843346300074
l2 norm of weights: 4.537527172941378

Iteration Number: 1218
Loss: 51.768897856036546
l2 norm of gradients: 0.6635653755035166
l2 norm of weights: 4.536167027154567

Iteration Number: 1219
Loss: 51.739708182251995
l2 norm of gradients: 0.6632874632766949
l2 norm of weights: 4.534807441591959

Iteration Number: 1220
Loss: 51.71053816178872
l2 norm of gradients: 0.6630096097154445
l2 norm of weights: 4.533448416155762

Iteration Number: 1221
Loss: 51.681387784227596
l2 norm of gradients: 0.6627318149156892
l2 norm of weights: 4.532089950747891

Iteration Number: 1222
Loss: 51.6522570391376
l2 norm of gradients: 0.6624540789734388
l2 norm of weights: 4.530732045269971

Iteration Number: 1223
Loss: 51.62314591607777
l2 norm of gradients: 0.6621764019847881
l2 norm of weights: 4.529374699623343

Iteration Number: 1224
Loss: 51.59405440461156
l2 norm of gradients: 0.6618987840459137
l2 norm of weights: 4.528017913709054

Iteration Number: 1225
Loss: 51.564982494282134
l2 norm of gradients: 0.6616212252530743
l2 norm of weights: 4.526661687427869

Iteration Number: 1226
Loss: 51.53593017467588
l2 norm of gradients: 0.6613437257026067

```

12 norm of weights: 4.525306020680264
-----
Iteration Number: 1227
Loss: 51.50689743529661
12 norm of gradients: 0.6610662854909257
12 norm of weights: 4.52395091336643
-----
Iteration Number: 1228
Loss: 51.47788426569252
12 norm of gradients: 0.6607889047145213
12 norm of weights: 4.522596365386273
-----
Iteration Number: 1229
Loss: 51.44889065540587
12 norm of gradients: 0.6605115834699576
12 norm of weights: 4.5212423766394165
-----
Iteration Number: 1230
Loss: 51.419916593957765
12 norm of gradients: 0.6602343218538705
12 norm of weights: 4.519888947025198
-----
Iteration Number: 1231
Loss: 51.39096207088006
12 norm of gradients: 0.6599571199629662
12 norm of weights: 4.518536076442673
-----
Iteration Number: 1232
Loss: 51.362027075693504
12 norm of gradients: 0.6596799778940198
12 norm of weights: 4.517183764790617
-----
Iteration Number: 1233
Loss: 51.333111597897236
12 norm of gradients: 0.659402895743872
12 norm of weights: 4.515832011967521
-----
Iteration Number: 1234
Loss: 51.30421562700937
12 norm of gradients: 0.6591258736094302
12 norm of weights: 4.514480817871599
-----
Iteration Number: 1235
Loss: 51.27533915253656
12 norm of gradients: 0.6588489115876635
12 norm of weights: 4.513130182400784
-----
Iteration Number: 1236

```


Loss: 51.24648216397783
l2 norm of gradients: 0.6585720097756032
l2 norm of weights: 4.511780105452728

Iteration Number: 1237
Loss: 51.217644650823516
l2 norm of gradients: 0.6582951682703405
l2 norm of weights: 4.5104305869248105

Iteration Number: 1238
Loss: 51.18882660254985
l2 norm of gradients: 0.6580183871690243
l2 norm of weights: 4.509081626714127

Iteration Number: 1239
Loss: 51.16002800867122
l2 norm of gradients: 0.6577416665688599
l2 norm of weights: 4.507733224717501

Iteration Number: 1240
Loss: 51.13124885863503
l2 norm of gradients: 0.657465006567107
l2 norm of weights: 4.506385380831477

Iteration Number: 1241
Loss: 51.10248914194428
l2 norm of gradients: 0.6571884072610785
l2 norm of weights: 4.505038094952328

Iteration Number: 1242
Loss: 51.07374884804831
l2 norm of gradients: 0.6569118687481382
l2 norm of weights: 4.50369136697605

Iteration Number: 1243
Loss: 51.04502796642586
l2 norm of gradients: 0.6566353911256994
l2 norm of weights: 4.502345196798366

Iteration Number: 1244
Loss: 51.01632648653523
l2 norm of gradients: 0.6563589744912227
l2 norm of weights: 4.5009995843147275

Iteration Number: 1245
Loss: 50.987644397818116
l2 norm of gradients: 0.6560826189422154
l2 norm of weights: 4.4996545294203125

```

-----
Iteration Number: 1246
Loss: 50.95898168975646
l2 norm of gradients: 0.6558063245762287
l2 norm of weights: 4.498310032010028
-----
Iteration Number: 1247
Loss: 50.93033835175861
l2 norm of gradients: 0.6555300914908563
l2 norm of weights: 4.4969660919785115
-----
Iteration Number: 1248
Loss: 50.90171437329358
l2 norm of gradients: 0.6552539197837328
l2 norm of weights: 4.4956227092201315
-----
Iteration Number: 1249
Loss: 50.87310974379548
l2 norm of gradients: 0.6549778095525328
l2 norm of weights: 4.494279883628986
-----
Iteration Number: 1250
Loss: 50.844524452682336
l2 norm of gradients: 0.6547017608949672
l2 norm of weights: 4.492937615098904
-----
Iteration Number: 1251
Loss: 50.815958489411955
l2 norm of gradients: 0.6544257739087839
l2 norm of weights: 4.491595903523451
-----
Iteration Number: 1252
Loss: 50.78741184338125
l2 norm of gradients: 0.6541498486917645
l2 norm of weights: 4.490254748795923
-----
Iteration Number: 1253
Loss: 50.75888450400797
l2 norm of gradients: 0.6538739853417231
l2 norm of weights: 4.488914150809352
-----
Iteration Number: 1254
Loss: 50.730376460724166
l2 norm of gradients: 0.6535981839565047
l2 norm of weights: 4.487574109456503
-----
Iteration Number: 1255
Loss: 50.7018877029376

```

l2 norm of gradients: 0.6533224446339837
l2 norm of weights: 4.486234624629879

Iteration Number: 1256
Loss: 50.67341822005176
l2 norm of gradients: 0.653046767472062
l2 norm of weights: 4.48489569622172

Iteration Number: 1257
Loss: 50.64496800147364
l2 norm of gradients: 0.6527711525686676
l2 norm of weights: 4.483557324124002

Iteration Number: 1258
Loss: 50.61653703658951
l2 norm of gradients: 0.6524956000217524
l2 norm of weights: 4.482219508228439

Iteration Number: 1259
Loss: 50.58812531479265
l2 norm of gradients: 0.6522201099292915
l2 norm of weights: 4.480882248426484

Iteration Number: 1260
Loss: 50.55973282547953
l2 norm of gradients: 0.6519446823892805
l2 norm of weights: 4.479545544609334

Iteration Number: 1261
Loss: 50.531359558034396
l2 norm of gradients: 0.651669317499735
l2 norm of weights: 4.478209396667922

Iteration Number: 1262
Loss: 50.50300550184501
l2 norm of gradients: 0.6513940153586878
l2 norm of weights: 4.476873804492925

Iteration Number: 1263
Loss: 50.47467064627925
l2 norm of gradients: 0.6511187760641882
l2 norm of weights: 4.47553876797476

Iteration Number: 1264
Loss: 50.44635498070424
l2 norm of gradients: 0.6508435997143002
l2 norm of weights: 4.47420428700359

Iteration Number: 1265
Loss: 50.41805849449518
l2 norm of gradients: 0.6505684864071005
l2 norm of weights: 4.47287036146932

Iteration Number: 1266
Loss: 50.3897811770231
l2 norm of gradients: 0.6502934362406776
l2 norm of weights: 4.4715369912616

Iteration Number: 1267
Loss: 50.36152301763242
l2 norm of gradients: 0.6500184493131291
l2 norm of weights: 4.470204176269827

Iteration Number: 1268
Loss: 50.33328400568161
l2 norm of gradients: 0.6497435257225614
l2 norm of weights: 4.4688719163831445

Iteration Number: 1269
Loss: 50.305064130528365
l2 norm of gradients: 0.6494686655670874
l2 norm of weights: 4.467540211490439

Iteration Number: 1270
Loss: 50.27686338152171
l2 norm of gradients: 0.649193868944825
l2 norm of weights: 4.466209061480351

Iteration Number: 1271
Loss: 50.24868174800107
l2 norm of gradients: 0.6489191359538955
l2 norm of weights: 4.464878466241265

Iteration Number: 1272
Loss: 50.220519219321304
l2 norm of gradients: 0.6486444666924222
l2 norm of weights: 4.463548425661317

Iteration Number: 1273
Loss: 50.1923757847842
l2 norm of gradients: 0.6483698612585289
l2 norm of weights: 4.4622189396283956

Iteration Number: 1274
Loss: 50.16425143375775
l2 norm of gradients: 0.648095319750338

```

12 norm of weights: 4.4608900080301375
-----
Iteration Number: 1275
Loss: 50.13614615555317
12 norm of gradients: 0.6478208422659691
12 norm of weights: 4.459561630753934
-----
Iteration Number: 1276
Loss: 50.10805993948797
12 norm of gradients: 0.6475464289035379
12 norm of weights: 4.458233807686927
-----
Iteration Number: 1277
Loss: 50.07999277490991
12 norm of gradients: 0.6472720797611539
12 norm of weights: 4.456906538716013
-----
Iteration Number: 1278
Loss: 50.05194465109745
12 norm of gradients: 0.6469977949369196
12 norm of weights: 4.455579823727844
-----
Iteration Number: 1279
Loss: 50.023915557400336
12 norm of gradients: 0.6467235745289281
12 norm of weights: 4.454253662608827
-----
Iteration Number: 1280
Loss: 49.995905483100685
12 norm of gradients: 0.6464494186352632
12 norm of weights: 4.452928055245125
-----
Iteration Number: 1281
Loss: 49.96791441752181
12 norm of gradients: 0.6461753273539953
12 norm of weights: 4.451603001522657
-----
Iteration Number: 1282
Loss: 49.93994234995751
12 norm of gradients: 0.6459013007831826
12 norm of weights: 4.4502785013271025
-----
Iteration Number: 1283
Loss: 49.91198926970268
12 norm of gradients: 0.6456273390208679
12 norm of weights: 4.448954554543898
-----
Iteration Number: 1284

```

Loss: 49.88405516606534
l2 norm of gradients: 0.6453534421650777
l2 norm of weights: 4.44763116105824

Iteration Number: 1285
Loss: 49.85614002832303
l2 norm of gradients: 0.6450796103138208
l2 norm of weights: 4.446308320755085

Iteration Number: 1286
Loss: 49.82824384577393
l2 norm of gradients: 0.6448058435650863
l2 norm of weights: 4.444986033519151

Iteration Number: 1287
Loss: 49.80036660769056
l2 norm of gradients: 0.6445321420168424
l2 norm of weights: 4.443664299234919

Iteration Number: 1288
Loss: 49.77250830337089
l2 norm of gradients: 0.6442585057670356
l2 norm of weights: 4.4423431177866295

Iteration Number: 1289
Loss: 49.74466892206355
l2 norm of gradients: 0.6439849349135878
l2 norm of weights: 4.441022489058291

Iteration Number: 1290
Loss: 49.71684845308314
l2 norm of gradients: 0.6437114295543964
l2 norm of weights: 4.439702412933676

Iteration Number: 1291
Loss: 49.68904688565405
l2 norm of gradients: 0.6434379897873317
l2 norm of weights: 4.438382889296319

Iteration Number: 1292
Loss: 49.661264209075235
l2 norm of gradients: 0.6431646157102359
l2 norm of weights: 4.437063918029523

Iteration Number: 1293
Loss: 49.633500412598686
l2 norm of gradients: 0.6428913074209218
l2 norm of weights: 4.43574549901636

```

-----
Iteration Number: 1294
Loss: 49.60575548548978
l2 norm of gradients: 0.6426180650171709
l2 norm of weights: 4.434427632139665
-----
Iteration Number: 1295
Loss: 49.57802941700266
l2 norm of gradients: 0.6423448885967323
l2 norm of weights: 4.433110317282047
-----
Iteration Number: 1296
Loss: 49.55032219638155
l2 norm of gradients: 0.6420717782573216
l2 norm of weights: 4.431793554325882
-----
Iteration Number: 1297
Loss: 49.522633812890575
l2 norm of gradients: 0.6417987340966187
l2 norm of weights: 4.430477343153315
-----
Iteration Number: 1298
Loss: 49.49496425578514
l2 norm of gradients: 0.6415257562122672
l2 norm of weights: 4.429161683646266
-----
Iteration Number: 1299
Loss: 49.467313514271034
l2 norm of gradients: 0.6412528447018719
l2 norm of weights: 4.4278465756864245
-----
Iteration Number: 1300
Loss: 49.439681577633685
l2 norm of gradients: 0.640979999662999
l2 norm of weights: 4.426532019155252
-----
Iteration Number: 1301
Loss: 49.412068435086006
l2 norm of gradients: 0.640707221193173
l2 norm of weights: 4.425218013933988
-----
Iteration Number: 1302
Loss: 49.38447407586857
l2 norm of gradients: 0.6404345093898767
l2 norm of weights: 4.423904559903642
-----
Iteration Number: 1303
Loss: 49.35689848921382

```

l2 norm of gradients: 0.6401618643505487
l2 norm of weights: 4.422591656945

Iteration Number: 1304
Loss: 49.3293416643375
l2 norm of gradients: 0.6398892861725831
l2 norm of weights: 4.421279304938627

Iteration Number: 1305
Loss: 49.30180359048734
l2 norm of gradients: 0.6396167749533275
l2 norm of weights: 4.419967503764864

Iteration Number: 1306
Loss: 49.27428425686409
l2 norm of gradients: 0.6393443307900811
l2 norm of weights: 4.418656253303828

Iteration Number: 1307
Loss: 49.24678365271054
l2 norm of gradients: 0.6390719537800948
l2 norm of weights: 4.417345553435417

Iteration Number: 1308
Loss: 49.219301767217104
l2 norm of gradients: 0.6387996440205687
l2 norm of weights: 4.416035404039306

Iteration Number: 1309
Loss: 49.191838589622066
l2 norm of gradients: 0.6385274016086507
l2 norm of weights: 4.414725804994956

Iteration Number: 1310
Loss: 49.16439410912101
l2 norm of gradients: 0.6382552266414362
l2 norm of weights: 4.4134167561816025

Iteration Number: 1311
Loss: 49.13696831492731
l2 norm of gradients: 0.6379831192159656
l2 norm of weights: 4.412108257478269

Iteration Number: 1312
Loss: 49.10956119624672
l2 norm of gradients: 0.6377110794292238
l2 norm of weights: 4.410800308763759

Iteration Number: 1313
Loss: 49.082172742274395
l2 norm of gradients: 0.6374391073781381
l2 norm of weights: 4.409492909916659

Iteration Number: 1314
Loss: 49.05480294221969
l2 norm of gradients: 0.6371672031595779
l2 norm of weights: 4.408186060815344

Iteration Number: 1315
Loss: 49.027451785276284
l2 norm of gradients: 0.6368953668703525
l2 norm of weights: 4.406879761337971

Iteration Number: 1316
Loss: 49.000119260639465
l2 norm of gradients: 0.6366235986072101
l2 norm of weights: 4.4055740113624875

Iteration Number: 1317
Loss: 48.97280535749646
l2 norm of gradients: 0.6363518984668369
l2 norm of weights: 4.404268810766623

Iteration Number: 1318
Loss: 48.945510065031286
l2 norm of gradients: 0.636080266545855
l2 norm of weights: 4.402964159427899

Iteration Number: 1319
Loss: 48.91823337244787
l2 norm of gradients: 0.6358087029408218
l2 norm of weights: 4.4016600572236255

Iteration Number: 1320
Loss: 48.890975268916485
l2 norm of gradients: 0.6355372077482285
l2 norm of weights: 4.400356504030901

Iteration Number: 1321
Loss: 48.86373574362375
l2 norm of gradients: 0.6352657810644986
l2 norm of weights: 4.399053499726615

Iteration Number: 1322
Loss: 48.83651478573998
l2 norm of gradients: 0.6349944229859871

```

12 norm of weights: 4.3977510441874506
-----
Iteration Number: 1323
Loss: 48.809312384446955
12 norm of gradients: 0.6347231336089789
12 norm of weights: 4.39644913728988
-----
Iteration Number: 1324
Loss: 48.78212852891921
12 norm of gradients: 0.6344519130296875
12 norm of weights: 4.3951477789101725
-----
Iteration Number: 1325
Loss: 48.754963208339596
12 norm of gradients: 0.6341807613442543
12 norm of weights: 4.393846968924388
-----
Iteration Number: 1326
Loss: 48.72781641185472
12 norm of gradients: 0.6339096786487466
12 norm of weights: 4.392546707208382
-----
Iteration Number: 1327
Loss: 48.70068812864148
12 norm of gradients: 0.6336386650391566
12 norm of weights: 4.391246993637808
-----
Iteration Number: 1328
Loss: 48.673578347861984
12 norm of gradients: 0.6333677206114006
12 norm of weights: 4.389947828088114
-----
Iteration Number: 1329
Loss: 48.646487058677614
12 norm of gradients: 0.6330968454613173
12 norm of weights: 4.388649210434546
-----
Iteration Number: 1330
Loss: 48.61941425025786
12 norm of gradients: 0.6328260396846668
12 norm of weights: 4.387351140552148
-----
Iteration Number: 1331
Loss: 48.592359911752304
12 norm of gradients: 0.6325553033771294
12 norm of weights: 4.386053618315765
-----
Iteration Number: 1332

```

```
Loss: 48.56532403230899
12 norm of gradients: 0.6322846366343043
12 norm of weights: 4.384756643600041
-----
Iteration Number: 1333
Loss: 48.53830660109798
12 norm of gradients: 0.6320140395517085
12 norm of weights: 4.383460216279418
-----
Iteration Number: 1334
Loss: 48.51130760725298
12 norm of gradients: 0.6317435122247751
12 norm of weights: 4.3821643362281435
-----
Iteration Number: 1335
Loss: 48.48432703993063
12 norm of gradients: 0.6314730547488535
12 norm of weights: 4.380869003320267
-----
Iteration Number: 1336
Loss: 48.457364888270675
12 norm of gradients: 0.6312026672192061
12 norm of weights: 4.379574217429641
-----
Iteration Number: 1337
Loss: 48.43042114143549
12 norm of gradients: 0.6309323497310095
12 norm of weights: 4.378279978429921
-----
Iteration Number: 1338
Loss: 48.40349578853991
12 norm of gradients: 0.6306621023793513
12 norm of weights: 4.37698628619457
-----
Iteration Number: 1339
Loss: 48.37658881874614
12 norm of gradients: 0.6303919252592299
12 norm of weights: 4.375693140596853
-----
Iteration Number: 1340
Loss: 48.3497002211842
12 norm of gradients: 0.630121818465554
12 norm of weights: 4.374400541509848
-----
Iteration Number: 1341
Loss: 48.322829984988275
12 norm of gradients: 0.6298517820931395
12 norm of weights: 4.373108488806435
```

```

-----
Iteration Number: 1342
Loss: 48.295978099294004
l2 norm of gradients: 0.6295818162367104
l2 norm of weights: 4.371816982359304
-----
Iteration Number: 1343
Loss: 48.269144553237076
l2 norm of gradients: 0.6293119209908965
l2 norm of weights: 4.3705260220409565
-----
Iteration Number: 1344
Loss: 48.24232933595067
l2 norm of gradients: 0.6290420964502328
l2 norm of weights: 4.369235607723702
-----
Iteration Number: 1345
Loss: 48.21553243655211
l2 norm of gradients: 0.6287723427091575
l2 norm of weights: 4.367945739279661
-----
Iteration Number: 1346
Loss: 48.18875384418026
l2 norm of gradients: 0.6285026598620126
l2 norm of weights: 4.366656416580769
-----
Iteration Number: 1347
Loss: 48.16199354795331
l2 norm of gradients: 0.6282330480030407
l2 norm of weights: 4.3653676394987695
-----
Iteration Number: 1348
Loss: 48.13525153699461
l2 norm of gradients: 0.6279635072263857
l2 norm of weights: 4.364079407905223
-----
Iteration Number: 1349
Loss: 48.10852780041711
l2 norm of gradients: 0.6276940376260903
l2 norm of weights: 4.362791721671503
-----
Iteration Number: 1350
Loss: 48.08182232736618
l2 norm of gradients: 0.6274246392960959
l2 norm of weights: 4.361504580668798
-----
Iteration Number: 1351
Loss: 48.05513510693167

```

l2 norm of gradients: 0.6271553123302411
l2 norm of weights: 4.3602179847681155

Iteration Number: 1352
Loss: 48.02846612824977
l2 norm of gradients: 0.6268860568222607
l2 norm of weights: 4.358931933840276

Iteration Number: 1353
Loss: 48.001815380428575
l2 norm of gradients: 0.6266168728657844
l2 norm of weights: 4.35764642775592

Iteration Number: 1354
Loss: 47.97518285256901
l2 norm of gradients: 0.6263477605543359
l2 norm of weights: 4.356361466385507

Iteration Number: 1355
Loss: 47.94856853380588
l2 norm of gradients: 0.6260787199813321
l2 norm of weights: 4.3550770495993145

Iteration Number: 1356
Loss: 47.92197241323525
l2 norm of gradients: 0.6258097512400816
l2 norm of weights: 4.35379317726744

Iteration Number: 1357
Loss: 47.89539447996051
l2 norm of gradients: 0.625540854423784
l2 norm of weights: 4.352509849259806

Iteration Number: 1358
Loss: 47.868834723101756
l2 norm of gradients: 0.6252720296255285
l2 norm of weights: 4.351227065446152

Iteration Number: 1359
Loss: 47.842293131758325
l2 norm of gradients: 0.6250032769382932
l2 norm of weights: 4.349944825696044

Iteration Number: 1360
Loss: 47.81576969503503
l2 norm of gradients: 0.6247345964549439
l2 norm of weights: 4.3486631298788705

Iteration Number: 1361
Loss: 47.78926440202914
l2 norm of gradients: 0.6244659882682333
l2 norm of weights: 4.347381977863844

Iteration Number: 1362
Loss: 47.76277724185998
l2 norm of gradients: 0.6241974524707994
l2 norm of weights: 4.346101369520002

Iteration Number: 1363
Loss: 47.7363082036028
l2 norm of gradients: 0.623928989155165
l2 norm of weights: 4.344821304716211

Iteration Number: 1364
Loss: 47.709857276372915
l2 norm of gradients: 0.6236605984137368
l2 norm of weights: 4.343541783321162

Iteration Number: 1365
Loss: 47.68342444926584
l2 norm of gradients: 0.6233922803388038
l2 norm of weights: 4.342262805203375

Iteration Number: 1366
Loss: 47.65700971136835
l2 norm of gradients: 0.6231240350225368
l2 norm of weights: 4.340984370231198

Iteration Number: 1367
Loss: 47.63061305179238
l2 norm of gradients: 0.6228558625569875
l2 norm of weights: 4.339706478272809

Iteration Number: 1368
Loss: 47.60423445961595
l2 norm of gradients: 0.622587763034087
l2 norm of weights: 4.338429129196218

Iteration Number: 1369
Loss: 47.57787392393611
l2 norm of gradients: 0.6223197365456451
l2 norm of weights: 4.337152322869263

Iteration Number: 1370
Loss: 47.551531433845994
l2 norm of gradients: 0.6220517831833493

```
12 norm of weights: 4.335876059159617
-----
Iteration Number: 1371
Loss: 47.5252069784354
12 norm of gradients: 0.6217839030387643
12 norm of weights: 4.334600337934785
-----
Iteration Number: 1372
Loss: 47.49890054679571
12 norm of gradients: 0.6215160962033303
12 norm of weights: 4.333325159062104
-----
Iteration Number: 1373
Loss: 47.47261212801375
12 norm of gradients: 0.6212483627683617
12 norm of weights: 4.332050522408751
-----
Iteration Number: 1374
Loss: 47.446341711174966
12 norm of gradients: 0.6209807028250481
12 norm of weights: 4.33077642784173
-----
Iteration Number: 1375
Loss: 47.42008928536553
12 norm of gradients: 0.6207131164644513
12 norm of weights: 4.329502875227891
-----
Iteration Number: 1376
Loss: 47.3938548396783
12 norm of gradients: 0.6204456037775046
12 norm of weights: 4.328229864433914
-----
Iteration Number: 1377
Loss: 47.36763836318496
12 norm of gradients: 0.6201781648550135
12 norm of weights: 4.3269573953263185
-----
Iteration Number: 1378
Loss: 47.341439844972925
12 norm of gradients: 0.6199107997876532
12 norm of weights: 4.325685467771466
-----
Iteration Number: 1379
Loss: 47.31525927413899
12 norm of gradients: 0.6196435086659675
12 norm of weights: 4.324414081635553
-----
Iteration Number: 1380
```

Loss: 47.2890966397373
l2 norm of gradients: 0.6193762915803696
l2 norm of weights: 4.323143236784621

Iteration Number: 1381
Loss: 47.26295193087475
l2 norm of gradients: 0.6191091486211393
l2 norm of weights: 4.321872933084549

Iteration Number: 1382
Loss: 47.2368251366139
l2 norm of gradients: 0.6188420798784233
l2 norm of weights: 4.320603170401061

Iteration Number: 1383
Loss: 47.210716246043305
l2 norm of gradients: 0.6185750854422338
l2 norm of weights: 4.319333948599722

Iteration Number: 1384
Loss: 47.18462524823825
l2 norm of gradients: 0.6183081654024478
l2 norm of weights: 4.318065267545941

Iteration Number: 1385
Loss: 47.15855213227858
l2 norm of gradients: 0.6180413198488057
l2 norm of weights: 4.316797127104973

Iteration Number: 1386
Loss: 47.13249688722981
l2 norm of gradients: 0.6177745488709115
l2 norm of weights: 4.315529527141919

Iteration Number: 1387
Loss: 47.106459502190276
l2 norm of gradients: 0.6175078525582312
l2 norm of weights: 4.314262467521723

Iteration Number: 1388
Loss: 47.080439966217796
l2 norm of gradients: 0.6172412310000912
l2 norm of weights: 4.3129959481091795

Iteration Number: 1389
Loss: 47.05443826839816
l2 norm of gradients: 0.616974684285679
l2 norm of weights: 4.311729968768929


```

-----
Iteration Number: 1390
Loss: 47.02845439779312
l2 norm of gradients: 0.6167082125040417
l2 norm of weights: 4.310464529365461
-----
Iteration Number: 1391
Loss: 47.002488343489645
l2 norm of gradients: 0.6164418157440845
l2 norm of weights: 4.309199629763118
-----
Iteration Number: 1392
Loss: 46.976540094550174
l2 norm of gradients: 0.6161754940945707
l2 norm of weights: 4.3079352698260855
-----
Iteration Number: 1393
Loss: 46.95060964005993
l2 norm of gradients: 0.6159092476441204
l2 norm of weights: 4.306671449418409
-----
Iteration Number: 1394
Loss: 46.924696969082774
l2 norm of gradients: 0.61564307648121
l2 norm of weights: 4.3054081684039796
-----
Iteration Number: 1395
Loss: 46.89880207069255
l2 norm of gradients: 0.6153769806941709
l2 norm of weights: 4.304145426646544
-----
Iteration Number: 1396
Loss: 46.87292493396078
l2 norm of gradients: 0.6151109603711893
l2 norm of weights: 4.302883224009704
-----
Iteration Number: 1397
Loss: 46.84706554796181
l2 norm of gradients: 0.6148450156003047
l2 norm of weights: 4.301621560356912
-----
Iteration Number: 1398
Loss: 46.82122390176074
l2 norm of gradients: 0.6145791464694097
l2 norm of weights: 4.300360435551479
-----
Iteration Number: 1399
Loss: 46.79539998442937

```

l2 norm of gradients: 0.6143133530662488
l2 norm of weights: 4.299099849456571

Iteration Number: 1400
Loss: 46.76959378503992
l2 norm of gradients: 0.6140476354784177
l2 norm of weights: 4.297839801935213

Iteration Number: 1401
Loss: 46.743805292668476
l2 norm of gradients: 0.6137819937933623
l2 norm of weights: 4.296580292850284

Iteration Number: 1402
Loss: 46.71803449637718
l2 norm of gradients: 0.6135164280983787
l2 norm of weights: 4.2953213220645265

Iteration Number: 1403
Loss: 46.69228138523321
l2 norm of gradients: 0.6132509384806113
l2 norm of weights: 4.2940628894405375

Iteration Number: 1404
Loss: 46.66654594831585
l2 norm of gradients: 0.612985525027053
l2 norm of weights: 4.292804994840778

Iteration Number: 1405
Loss: 46.64082817468936
l2 norm of gradients: 0.6127201878245433
l2 norm of weights: 4.291547638127569

Iteration Number: 1406
Loss: 46.61512805341415
l2 norm of gradients: 0.6124549269597691
l2 norm of weights: 4.2902908191630935

Iteration Number: 1407
Loss: 46.58944557356827
l2 norm of gradients: 0.6121897425192623
l2 norm of weights: 4.289034537809396

Iteration Number: 1408
Loss: 46.56378072422114
l2 norm of gradients: 0.6119246345894003
l2 norm of weights: 4.287778793928387

Iteration Number: 1409
Loss: 46.53813349443554
l2 norm of gradients: 0.6116596032564047
l2 norm of weights: 4.286523587381841

Iteration Number: 1410
Loss: 46.512503873283926
l2 norm of gradients: 0.6113946486063401
l2 norm of weights: 4.285268918031394

Iteration Number: 1411
Loss: 46.486891849841896
l2 norm of gradients: 0.6111297707251143
l2 norm of weights: 4.284014785738553

Iteration Number: 1412
Loss: 46.46129741316954
l2 norm of gradients: 0.6108649696984773
l2 norm of weights: 4.28276119036469

Iteration Number: 1413
Loss: 46.4357205523345
l2 norm of gradients: 0.6106002456120196
l2 norm of weights: 4.2815081317710435

Iteration Number: 1414
Loss: 46.4101612564113
l2 norm of gradients: 0.6103355985511733
l2 norm of weights: 4.280255609818722

Iteration Number: 1415
Loss: 46.38461951446736
l2 norm of gradients: 0.6100710286012092
l2 norm of weights: 4.279003624368703

Iteration Number: 1416
Loss: 46.359095315567544
l2 norm of gradients: 0.6098065358472382
l2 norm of weights: 4.277752175281833

Iteration Number: 1417
Loss: 46.33358864878509
l2 norm of gradients: 0.6095421203742092
l2 norm of weights: 4.276501262418831

Iteration Number: 1418
Loss: 46.308099503193915
l2 norm of gradients: 0.6092777822669087

```

12 norm of weights: 4.275250885640286
-----
Iteration Number: 1419
Loss: 46.2826278678552
12 norm of gradients: 0.6090135216099601
12 norm of weights: 4.274001044806662
-----
Iteration Number: 1420
Loss: 46.25717373184111
12 norm of gradients: 0.6087493384878239
12 norm of weights: 4.272751739778293
-----
Iteration Number: 1421
Loss: 46.231737084226125
12 norm of gradients: 0.6084852329847951
12 norm of weights: 4.271502970415391
-----
Iteration Number: 1422
Loss: 46.20631791407718
12 norm of gradients: 0.6082212051850043
12 norm of weights: 4.270254736578038
-----
Iteration Number: 1423
Loss: 46.180916210470244
12 norm of gradients: 0.6079572551724164
12 norm of weights: 4.2690070381261975
-----
Iteration Number: 1424
Loss: 46.155531962464174
12 norm of gradients: 0.6076933830308294
12 norm of weights: 4.267759874919705
-----
Iteration Number: 1425
Loss: 46.13016515914267
12 norm of gradients: 0.6074295888438748
12 norm of weights: 4.266513246818276
-----
Iteration Number: 1426
Loss: 46.104815789578055
12 norm of gradients: 0.6071658726950158
12 norm of weights: 4.265267153681503
-----
Iteration Number: 1427
Loss: 46.07948384283184
12 norm of gradients: 0.6069022346675474
12 norm of weights: 4.264021595368857
-----
Iteration Number: 1428

```

```
Loss: 46.05416930798681
12 norm of gradients: 0.6066386748445959
12 norm of weights: 4.262776571739689
-----
Iteration Number: 1429
Loss: 46.02887217410996
12 norm of gradients: 0.6063751933091172
12 norm of weights: 4.261532082653233
-----
Iteration Number: 1430
Loss: 46.0035924302817
12 norm of gradients: 0.6061117901438973
12 norm of weights: 4.2602881279686
-----
Iteration Number: 1431
Loss: 45.97833006557108
12 norm of gradients: 0.605848465431551
12 norm of weights: 4.2590447075447875
-----
Iteration Number: 1432
Loss: 45.95308506905639
12 norm of gradients: 0.6055852192545218
12 norm of weights: 4.257801821240672
-----
Iteration Number: 1433
Loss: 45.927857429809144
12 norm of gradients: 0.6053220516950802
12 norm of weights: 4.256559468915016
-----
Iteration Number: 1434
Loss: 45.902647136913735
12 norm of gradients: 0.6050589628353247
12 norm of weights: 4.255317650426468
-----
Iteration Number: 1435
Loss: 45.87745417943804
12 norm of gradients: 0.6047959527571797
12 norm of weights: 4.254076365633558
-----
Iteration Number: 1436
Loss: 45.852278546460525
12 norm of gradients: 0.6045330215423955
12 norm of weights: 4.252835614394705
-----
Iteration Number: 1437
Loss: 45.82712022706394
12 norm of gradients: 0.604270169272548
12 norm of weights: 4.251595396568215
```

```
-----  
Iteration Number: 1438  
Loss: 45.801979210323104  
l2 norm of gradients: 0.6040073960290375  
l2 norm of weights: 4.250355712012279  
-----  
Iteration Number: 1439  
Loss: 45.77685548532247  
l2 norm of gradients: 0.6037447018930886  
l2 norm of weights: 4.249116560584979  
-----  
Iteration Number: 1440  
Loss: 45.751749041135206  
l2 norm of gradients: 0.603482086945749  
l2 norm of weights: 4.247877942144286  
-----  
Iteration Number: 1441  
Loss: 45.72665986685025  
l2 norm of gradients: 0.6032195512678896  
l2 norm of weights: 4.246639856548061  
-----  
Iteration Number: 1442  
Loss: 45.701587951543615  
l2 norm of gradients: 0.6029570949402034  
l2 norm of weights: 4.245402303654054  
-----  
Iteration Number: 1443  
Loss: 45.67653328430128  
l2 norm of gradients: 0.6026947180432052  
l2 norm of weights: 4.24416528331991  
-----  
Iteration Number: 1444  
Loss: 45.651495854207674  
l2 norm of gradients: 0.6024324206572309  
l2 norm of weights: 4.242928795403164  
-----  
Iteration Number: 1445  
Loss: 45.62647565034365  
l2 norm of gradients: 0.6021702028624372  
l2 norm of weights: 4.241692839761244  
-----  
Iteration Number: 1446  
Loss: 45.601472661800614  
l2 norm of gradients: 0.6019080647388001  
l2 norm of weights: 4.240457416251474  
-----  
Iteration Number: 1447  
Loss: 45.576486877655974
```

l2 norm of gradients: 0.6016460063661161
l2 norm of weights: 4.239222524731072

Iteration Number: 1448
Loss: 45.55151828700676
l2 norm of gradients: 0.6013840278239997
l2 norm of weights: 4.237988165057149

Iteration Number: 1449
Loss: 45.52656687893751
l2 norm of gradients: 0.6011221291918842
l2 norm of weights: 4.236754337086717

Iteration Number: 1450
Loss: 45.50163264254061
l2 norm of gradients: 0.6008603105490202
l2 norm of weights: 4.235521040676683

Iteration Number: 1451
Loss: 45.47671556689788
l2 norm of gradients: 0.6005985719744763
l2 norm of weights: 4.23428827568385

Iteration Number: 1452
Loss: 45.45181564111113
l2 norm of gradients: 0.6003369135471371
l2 norm of weights: 4.233056041964923

Iteration Number: 1453
Loss: 45.426932854267164
l2 norm of gradients: 0.6000753353457039
l2 norm of weights: 4.231824339376503

Iteration Number: 1454
Loss: 45.40206719546007
l2 norm of gradients: 0.5998138374486931
l2 norm of weights: 4.230593167775096

Iteration Number: 1455
Loss: 45.37721865379162
l2 norm of gradients: 0.5995524199344366
l2 norm of weights: 4.229362527017104

Iteration Number: 1456
Loss: 45.352387218348525
l2 norm of gradients: 0.599291082881081
l2 norm of weights: 4.228132416958833

Iteration Number: 1457
Loss: 45.32757287823317
l2 norm of gradients: 0.5990298263665865
l2 norm of weights: 4.226902837456493

Iteration Number: 1458
Loss: 45.302775622543365
l2 norm of gradients: 0.5987686504687274
l2 norm of weights: 4.225673788366196

Iteration Number: 1459
Loss: 45.27799544038295
l2 norm of gradients: 0.5985075552650907
l2 norm of weights: 4.224445269543956

Iteration Number: 1460
Loss: 45.253232320846564
l2 norm of gradients: 0.5982465408330763
l2 norm of weights: 4.223217280845696

Iteration Number: 1461
Loss: 45.22848625304167
l2 norm of gradients: 0.5979856072498956
l2 norm of weights: 4.221989822127242

Iteration Number: 1462
Loss: 45.20375722606801
l2 norm of gradients: 0.5977247545925723
l2 norm of weights: 4.220762893244326

Iteration Number: 1463
Loss: 45.179045229031985
l2 norm of gradients: 0.5974639829379407
l2 norm of weights: 4.219536494052591

Iteration Number: 1464
Loss: 45.15435025104766
l2 norm of gradients: 0.5972032923626459
l2 norm of weights: 4.218310624407583

Iteration Number: 1465
Loss: 45.12967228121039
l2 norm of gradients: 0.5969426829431431
l2 norm of weights: 4.21708528416476

Iteration Number: 1466
Loss: 45.10501130863545
l2 norm of gradients: 0.5966821547556972


```

12 norm of weights: 4.215860473179489
-----
Iteration Number: 1467
Loss: 45.08036732243511
12 norm of gradients: 0.5964217078763822
12 norm of weights: 4.214636191307045
-----
Iteration Number: 1468
Loss: 45.05574031172241
12 norm of gradients: 0.5961613423810807
12 norm of weights: 4.213412438402618
-----
Iteration Number: 1469
Loss: 45.031130265609676
12 norm of gradients: 0.5959010583454839
12 norm of weights: 4.212189214321307
-----
Iteration Number: 1470
Loss: 45.00653717320873
12 norm of gradients: 0.5956408558450904
12 norm of weights: 4.2109665189181245
-----
Iteration Number: 1471
Loss: 44.98196102364332
12 norm of gradients: 0.5953807349552067
12 norm of weights: 4.209744352047995
-----
Iteration Number: 1472
Loss: 44.957401806027455
12 norm of gradients: 0.5951206957509456
12 norm of weights: 4.20852271356576
-----
Iteration Number: 1473
Loss: 44.93285950948381
12 norm of gradients: 0.5948607383072267
12 norm of weights: 4.2073016033261705
-----
Iteration Number: 1474
Loss: 44.90833412312864
12 norm of gradients: 0.5946008626987755
12 norm of weights: 4.2060810211839
-----
Iteration Number: 1475
Loss: 44.883825636094734
12 norm of gradients: 0.5943410690001233
12 norm of weights: 4.204860966993534
-----
Iteration Number: 1476

```

Loss: 44.85933403749618
l2 norm of gradients: 0.5940813572856063
l2 norm of weights: 4.203641440609577

Iteration Number: 1477
Loss: 44.83485931646979
l2 norm of gradients: 0.5938217276293652
l2 norm of weights: 4.202422441886448

Iteration Number: 1478
Loss: 44.810401462140476
l2 norm of gradients: 0.593562180105346
l2 norm of weights: 4.201203970678487

Iteration Number: 1479
Loss: 44.78596046363569
l2 norm of gradients: 0.5933027147872972
l2 norm of weights: 4.199986026839958

Iteration Number: 1480
Loss: 44.76153631009358
l2 norm of gradients: 0.5930433317487719
l2 norm of weights: 4.198768610225035

Iteration Number: 1481
Loss: 44.73712899063755
l2 norm of gradients: 0.5927840310631257
l2 norm of weights: 4.197551720687821

Iteration Number: 1482
Loss: 44.7127384944167
l2 norm of gradients: 0.592524812803517
l2 norm of weights: 4.196335358082338

Iteration Number: 1483
Loss: 44.688364810555846
l2 norm of gradients: 0.5922656770429063
l2 norm of weights: 4.19511952226253

Iteration Number: 1484
Loss: 44.66400792820434
l2 norm of gradients: 0.5920066238540563
l2 norm of weights: 4.1939042130822655

Iteration Number: 1485
Loss: 44.63966783649788
l2 norm of gradients: 0.5917476533095309
l2 norm of weights: 4.192689430395333

```

-----
Iteration Number: 1486
Loss: 44.61534452457959
l2 norm of gradients: 0.5914887654816954
l2 norm of weights: 4.191475174055452
-----
Iteration Number: 1487
Loss: 44.59103798159989
l2 norm of gradients: 0.5912299604427149
l2 norm of weights: 4.190261443916262
-----
Iteration Number: 1488
Loss: 44.56674819669869
l2 norm of gradients: 0.590971238264556
l2 norm of weights: 4.1890482398313305
-----
Iteration Number: 1489
Loss: 44.54247515902975
l2 norm of gradients: 0.5907125990189843
l2 norm of weights: 4.187835561654152
-----
Iteration Number: 1490
Loss: 44.51821885774597
l2 norm of gradients: 0.5904540427775656
l2 norm of weights: 4.186623409238148
-----
Iteration Number: 1491
Loss: 44.49397928199356
l2 norm of gradients: 0.5901955696116645
l2 norm of weights: 4.185411782436669
-----
Iteration Number: 1492
Loss: 44.46975642093523
l2 norm of gradients: 0.5899371795924445
l2 norm of weights: 4.184200681102993
-----
Iteration Number: 1493
Loss: 44.445550263723526
l2 norm of gradients: 0.5896788727908675
l2 norm of weights: 4.182990105090331
-----
Iteration Number: 1494
Loss: 44.42136079951544
l2 norm of gradients: 0.5894206492776939
l2 norm of weights: 4.1817800542518215
-----
Iteration Number: 1495
Loss: 44.397188017483614

```

l2 norm of gradients: 0.5891625091234812
l2 norm of weights: 4.180570528440535

Iteration Number: 1496
Loss: 44.37303190678071
l2 norm of gradients: 0.588904452398585
l2 norm of weights: 4.1793615275094735

Iteration Number: 1497
Loss: 44.34889245657413
l2 norm of gradients: 0.5886464791731573
l2 norm of weights: 4.1781530513115746

Iteration Number: 1498
Loss: 44.324769656036395
l2 norm of gradients: 0.5883885895171473
l2 norm of weights: 4.176945099699705

Iteration Number: 1499
Loss: 44.300663494333634
l2 norm of gradients: 0.5881307835003003
l2 norm of weights: 4.175737672526669

Iteration Number: 1500
Loss: 44.276573960643915
l2 norm of gradients: 0.5878730611921581
l2 norm of weights: 4.174530769645204

Iteration Number: 1501
Loss: 44.25250104413344
l2 norm of gradients: 0.5876154226620574
l2 norm of weights: 4.173324390907986

Iteration Number: 1502
Loss: 44.22844473399131
l2 norm of gradients: 0.5873578679791308
l2 norm of weights: 4.172118536167621

Iteration Number: 1503
Loss: 44.20440501938403
l2 norm of gradients: 0.5871003972123062
l2 norm of weights: 4.170913205276659

Iteration Number: 1504
Loss: 44.180381889497596
l2 norm of gradients: 0.5868430104303057
l2 norm of weights: 4.169708398087585

Iteration Number: 1505
Loss: 44.15637533352191
l2 norm of gradients: 0.5865857077016462
l2 norm of weights: 4.168504114452821

Iteration Number: 1506
Loss: 44.132385340637754
l2 norm of gradients: 0.5863284890946386
l2 norm of weights: 4.16730035422473

Iteration Number: 1507
Loss: 44.10841190003568
l2 norm of gradients: 0.5860713546773876
l2 norm of weights: 4.166097117255617

Iteration Number: 1508
Loss: 44.08445500091012
l2 norm of gradients: 0.5858143045177916
l2 norm of weights: 4.164894403397723

Iteration Number: 1509
Loss: 44.060514632444956
l2 norm of gradients: 0.5855573386835417
l2 norm of weights: 4.163692212503234

Iteration Number: 1510
Loss: 44.03659078385094
l2 norm of gradients: 0.5853004572421225
l2 norm of weights: 4.162490544424276

Iteration Number: 1511
Loss: 44.01268344431158
l2 norm of gradients: 0.5850436602608111
l2 norm of weights: 4.1612893990129205

Iteration Number: 1512
Loss: 43.98879260303834
l2 norm of gradients: 0.5847869478066765
l2 norm of weights: 4.160088776121181

Iteration Number: 1513
Loss: 43.96491824923272
l2 norm of gradients: 0.5845303199465806
l2 norm of weights: 4.1588886756010135

Iteration Number: 1514
Loss: 43.941060372103145
l2 norm of gradients: 0.5842737767471756

```

12 norm of weights: 4.157689097304322
-----
Iteration Number: 1515
Loss: 43.91721896085294
12 norm of gradients: 0.5840173182749071
12 norm of weights: 4.1564900410829555
-----
Iteration Number: 1516
Loss: 43.89339400469844
12 norm of gradients: 0.5837609445960104
12 norm of weights: 4.155291506788709
-----
Iteration Number: 1517
Loss: 43.86958549285196
12 norm of gradients: 0.5835046557765123
12 norm of weights: 4.154093494273323
-----
Iteration Number: 1518
Loss: 43.845793414527975
12 norm of gradients: 0.5832484518822304
12 norm of weights: 4.15289600338849
-----
Iteration Number: 1519
Loss: 43.8220177589522
12 norm of gradients: 0.5829923329787726
12 norm of weights: 4.151699033985847
-----
Iteration Number: 1520
Loss: 43.798258515342546
12 norm of gradients: 0.5827362991315368
12 norm of weights: 4.150502585916981
-----
Iteration Number: 1521
Loss: 43.77451567292512
12 norm of gradients: 0.5824803504057109
12 norm of weights: 4.149306659033432
-----
Iteration Number: 1522
Loss: 43.750789220925434
12 norm of gradients: 0.5822244868662728
12 norm of weights: 4.148111253186687
-----
Iteration Number: 1523
Loss: 43.72707914857903
12 norm of gradients: 0.581968708577989
12 norm of weights: 4.146916368228186
-----
Iteration Number: 1524

```

Loss: 43.70338544511673
l2 norm of gradients: 0.5817130156054158
l2 norm of weights: 4.145722004009321

Iteration Number: 1525
Loss: 43.67970809976825
l2 norm of gradients: 0.5814574080128982
l2 norm of weights: 4.144528160381437

Iteration Number: 1526
Loss: 43.65604710178002
l2 norm of gradients: 0.5812018858645701
l2 norm of weights: 4.143334837195832

Iteration Number: 1527
Loss: 43.632402440394834
l2 norm of gradients: 0.5809464492243531
l2 norm of weights: 4.142142034303759

Iteration Number: 1528
Loss: 43.608774104853005
l2 norm of gradients: 0.580691098155958
l2 norm of weights: 4.140949751556424

Iteration Number: 1529
Loss: 43.5851620844044
l2 norm of gradients: 0.5804358327228831
l2 norm of weights: 4.13975798880499

Iteration Number: 1530
Loss: 43.56156636829446
l2 norm of gradients: 0.5801806529884143
l2 norm of weights: 4.138566745900578

Iteration Number: 1531
Loss: 43.53798694578116
l2 norm of gradients: 0.5799255590156248
l2 norm of weights: 4.137376022694262

Iteration Number: 1532
Loss: 43.5144238061228
l2 norm of gradients: 0.579670550867376
l2 norm of weights: 4.136185819037077

Iteration Number: 1533
Loss: 43.49087693857516
l2 norm of gradients: 0.5794156286063155
l2 norm of weights: 4.134996134780014

```

-----
Iteration Number: 1534
Loss: 43.46734633240214
l2 norm of gradients: 0.5791607922948783
l2 norm of weights: 4.1338069697740245
-----
Iteration Number: 1535
Loss: 43.44383197686578
l2 norm of gradients: 0.5789060419952858
l2 norm of weights: 4.132618323870019
-----
Iteration Number: 1536
Loss: 43.420333861239534
l2 norm of gradients: 0.5786513777695462
l2 norm of weights: 4.13143019691887
-----
Iteration Number: 1537
Loss: 43.39685197479135
l2 norm of gradients: 0.5783967996794536
l2 norm of weights: 4.130242588771408
-----
Iteration Number: 1538
Loss: 43.37338630679582
l2 norm of gradients: 0.578142307786588
l2 norm of weights: 4.129055499278428
-----
Iteration Number: 1539
Loss: 43.34993684652903
l2 norm of gradients: 0.5778879021523161
l2 norm of weights: 4.127868928290686
-----
Iteration Number: 1540
Loss: 43.3265035832785
l2 norm of gradients: 0.5776335828377893
l2 norm of weights: 4.126682875658903
-----
Iteration Number: 1541
Loss: 43.30308650632008
l2 norm of gradients: 0.5773793499039451
l2 norm of weights: 4.125497341233762
-----
Iteration Number: 1542
Loss: 43.27968560494453
l2 norm of gradients: 0.577125203411506
l2 norm of weights: 4.124312324865911
-----
Iteration Number: 1543
Loss: 43.256300868443574

```


l2 norm of gradients: 0.5768711434209799
l2 norm of weights: 4.1231278264059625

Iteration Number: 1544
Loss: 43.2329322861073
l2 norm of gradients: 0.5766171699926594
l2 norm of weights: 4.121943845704497

Iteration Number: 1545
Loss: 43.209579847236
l2 norm of gradients: 0.576363283186622
l2 norm of weights: 4.120760382612058

Iteration Number: 1546
Loss: 43.186243541125165
l2 norm of gradients: 0.5761094830627298
l2 norm of weights: 4.11957743697916

Iteration Number: 1547
Loss: 43.162923357082015
l2 norm of gradients: 0.5758557696806289
l2 norm of weights: 4.118395008656283

Iteration Number: 1548
Loss: 43.13961928440985
l2 norm of gradients: 0.5756021430997504
l2 norm of weights: 4.117213097493877

Iteration Number: 1549
Loss: 43.11633131242163
l2 norm of gradients: 0.575348603379309
l2 norm of weights: 4.116031703342361

Iteration Number: 1550
Loss: 43.093059430427374
l2 norm of gradients: 0.5750951505783034
l2 norm of weights: 4.114850826052121

Iteration Number: 1551
Loss: 43.069803627747476
l2 norm of gradients: 0.5748417847555162
l2 norm of weights: 4.113670465473518

Iteration Number: 1552
Loss: 43.04656389369887
l2 norm of gradients: 0.5745885059695133
l2 norm of weights: 4.112490621456882

Iteration Number: 1553
Loss: 43.023340217606965
l2 norm of gradients: 0.5743353142786443
l2 norm of weights: 4.111311293852514

Iteration Number: 1554
Loss: 43.00013258879293
l2 norm of gradients: 0.5740822097410424
l2 norm of weights: 4.11013248251069

Iteration Number: 1555
Loss: 42.97694099659333
l2 norm of gradients: 0.5738291924146233
l2 norm of weights: 4.1089541872816575

Iteration Number: 1556
Loss: 42.95376543033841
l2 norm of gradients: 0.5735762623570865
l2 norm of weights: 4.107776408015638

Iteration Number: 1557
Loss: 42.93060587936574
l2 norm of gradients: 0.5733234196259136
l2 norm of weights: 4.106599144562828

Iteration Number: 1558
Loss: 42.90746233302002
l2 norm of gradients: 0.5730706642783696
l2 norm of weights: 4.105422396773399

Iteration Number: 1559
Loss: 42.88433478063607
l2 norm of gradients: 0.5728179963715019
l2 norm of weights: 4.104246164497499

Iteration Number: 1560
Loss: 42.861223211567676
l2 norm of gradients: 0.5725654159621402
l2 norm of weights: 4.103070447585251

Iteration Number: 1561
Loss: 42.838127615166734
l2 norm of gradients: 0.5723129231068969
l2 norm of weights: 4.101895245886757

Iteration Number: 1562
Loss: 42.8150479807878
l2 norm of gradients: 0.5720605178621663

```

12 norm of weights: 4.100720559252094
-----
Iteration Number: 1563
Loss: 42.79198429778588
12 norm of gradients: 0.5718082002841252
12 norm of weights: 4.099546387531324
-----
Iteration Number: 1564
Loss: 42.76893655552284
12 norm of gradients: 0.571555970428732
12 norm of weights: 4.0983727305744795
-----
Iteration Number: 1565
Loss: 42.74590474336792
12 norm of gradients: 0.5713038283517272
12 norm of weights: 4.097199588231577
-----
Iteration Number: 1566
Loss: 42.72288885068992
12 norm of gradients: 0.5710517741086331
12 norm of weights: 4.096026960352615
-----
Iteration Number: 1567
Loss: 42.6998888668571
12 norm of gradients: 0.5707998077547535
12 norm of weights: 4.094854846787572
-----
Iteration Number: 1568
Loss: 42.67690478125036
12 norm of gradients: 0.5705479293451742
12 norm of weights: 4.093683247386405
-----
Iteration Number: 1569
Loss: 42.65393658324913
12 norm of gradients: 0.5702961389347618
12 norm of weights: 4.092512161999059
-----
Iteration Number: 1570
Loss: 42.63098426224021
12 norm of gradients: 0.5700444365781645
12 norm of weights: 4.091341590475457
-----
Iteration Number: 1571
Loss: 42.608047807602965
12 norm of gradients: 0.5697928223298121
12 norm of weights: 4.090171532665509
-----
Iteration Number: 1572

```

```

Loss: 42.585127208735585
l2 norm of gradients: 0.5695412962439149
l2 norm of weights: 4.089001988419105
-----
Iteration Number: 1573
Loss: 42.56222245503312
l2 norm of gradients: 0.5692898583744649
l2 norm of weights: 4.087832957586127
-----
Iteration Number: 1574
Loss: 42.53933353589447
l2 norm of gradients: 0.5690385087752348
l2 norm of weights: 4.086664440016438
-----
Iteration Number: 1575
Loss: 42.51646044071818
l2 norm of gradients: 0.568787247499778
l2 norm of weights: 4.085496435559886
-----
Iteration Number: 1576
Loss: 42.493603158916045
l2 norm of gradients: 0.5685360746014292
l2 norm of weights: 4.084328944066309
-----
Iteration Number: 1577
Loss: 42.47076167989937
l2 norm of gradients: 0.5682849901333032
l2 norm of weights: 4.083161965385529
-----
Iteration Number: 1578
Loss: 42.44793599307835
l2 norm of gradients: 0.5680339941482957
l2 norm of weights: 4.081995499367361
-----
Iteration Number: 1579
Loss: 42.4251260878719
l2 norm of gradients: 0.5677830866990833
l2 norm of weights: 4.080829545861604
-----
Iteration Number: 1580
Loss: 42.40233195370587
l2 norm of gradients: 0.5675322678381225
l2 norm of weights: 4.07966410471805
-----
Iteration Number: 1581
Loss: 42.379553580002494
l2 norm of gradients: 0.5672815376176508
l2 norm of weights: 4.078499175786479

```

```

-----
Iteration Number: 1582
Loss: 42.35679095619755
l2 norm of gradients: 0.5670308960896854
l2 norm of weights: 4.077334758916661
-----
Iteration Number: 1583
Loss: 42.3340440717189
l2 norm of gradients: 0.5667803433060243
l2 norm of weights: 4.076170853958359
-----
Iteration Number: 1584
Loss: 42.31131291600788
l2 norm of gradients: 0.5665298793182457
l2 norm of weights: 4.075007460761328
-----
Iteration Number: 1585
Loss: 42.28859747850806
l2 norm of gradients: 0.5662795041777077
l2 norm of weights: 4.073844579175313
-----
Iteration Number: 1586
Loss: 42.26589774866165
l2 norm of gradients: 0.5660292179355484
l2 norm of weights: 4.072682209050057
-----
Iteration Number: 1587
Loss: 42.243213715921875
l2 norm of gradients: 0.5657790206426866
l2 norm of weights: 4.071520350235289
-----
Iteration Number: 1588
Loss: 42.22054536974365
l2 norm of gradients: 0.5655289123498202
l2 norm of weights: 4.070359002580742
-----
Iteration Number: 1589
Loss: 42.197892699584955
l2 norm of gradients: 0.5652788931074277
l2 norm of weights: 4.069198165936135
-----
Iteration Number: 1590
Loss: 42.17525569490547
l2 norm of gradients: 0.5650289629657674
l2 norm of weights: 4.068037840151189
-----
Iteration Number: 1591
Loss: 42.15263434517356

```

l2 norm of gradients: 0.5647791219748772
l2 norm of weights: 4.066878025075618

Iteration Number: 1592
Loss: 42.13002863986021
l2 norm of gradients: 0.564529370184575
l2 norm of weights: 4.0657187205591345

Iteration Number: 1593
Loss: 42.10743856844355
l2 norm of gradients: 0.5642797076444586
l2 norm of weights: 4.064559926451446

Iteration Number: 1594
Loss: 42.08486412039592
l2 norm of gradients: 0.564030134403905
l2 norm of weights: 4.063401642602262

Iteration Number: 1595
Loss: 42.06230528520815
l2 norm of gradients: 0.5637806505120715
l2 norm of weights: 4.0622438688612865

Iteration Number: 1596
Loss: 42.03976205236091
l2 norm of gradients: 0.5635312560178946
l2 norm of weights: 4.061086605078226

Iteration Number: 1597
Loss: 42.01723441135049
l2 norm of gradients: 0.5632819509700909
l2 norm of weights: 4.059929851102785

Iteration Number: 1598
Loss: 41.994722351671555
l2 norm of gradients: 0.5630327354171561
l2 norm of weights: 4.058773606784669

Iteration Number: 1599
Loss: 41.972225862823805
l2 norm of gradients: 0.5627836094073659
l2 norm of weights: 4.0576178719735845

Iteration Number: 1600
Loss: 41.94974493431257
l2 norm of gradients: 0.5625345729887755
l2 norm of weights: 4.056462646519241

Iteration Number: 1601
Loss: 41.92727955564423
l2 norm of gradients: 0.5622856262092194
l2 norm of weights: 4.055307930271347

Iteration Number: 1602
Loss: 41.90482971633534
l2 norm of gradients: 0.5620367691163116
l2 norm of weights: 4.054153723079619

Iteration Number: 1603
Loss: 41.88239540590371
l2 norm of gradients: 0.561788001757446
l2 norm of weights: 4.053000024793772

Iteration Number: 1604
Loss: 41.85997661386803
l2 norm of gradients: 0.5615393241797958
l2 norm of weights: 4.051846835263529

Iteration Number: 1605
Loss: 41.83757332975867
l2 norm of gradients: 0.5612907364303134
l2 norm of weights: 4.050694154338614

Iteration Number: 1606
Loss: 41.81518554310203
l2 norm of gradients: 0.5610422385557311
l2 norm of weights: 4.049541981868758

Iteration Number: 1607
Loss: 41.7928132434331
l2 norm of gradients: 0.5607938306025603
l2 norm of weights: 4.048390317703699

Iteration Number: 1608
Loss: 41.770456420291744
l2 norm of gradients: 0.5605455126170923
l2 norm of weights: 4.047239161693183

Iteration Number: 1609
Loss: 41.74811506322543
l2 norm of gradients: 0.5602972846453975
l2 norm of weights: 4.046088513686956

Iteration Number: 1610
Loss: 41.72578916178005
l2 norm of gradients: 0.5600491467333258

```
12 norm of weights: 4.044938373534778
-----
Iteration Number: 1611
Loss: 41.70347870550817
12 norm of gradients: 0.559801098926507
12 norm of weights: 4.043788741086418
-----
Iteration Number: 1612
Loss: 41.68118368396665
12 norm of gradients: 0.5595531412703497
12 norm of weights: 4.042639616191648
-----
Iteration Number: 1613
Loss: 41.65890408671647
12 norm of gradients: 0.5593052738100426
12 norm of weights: 4.0414909987002545
-----
Iteration Number: 1614
Loss: 41.636639903326625
12 norm of gradients: 0.5590574965905533
12 norm of weights: 4.040342888462031
-----
Iteration Number: 1615
Loss: 41.61439112336723
12 norm of gradients: 0.5588098096566295
12 norm of weights: 4.039195285326784
-----
Iteration Number: 1616
Loss: 41.59215773641115
12 norm of gradients: 0.5585622130527982
12 norm of weights: 4.038048189144329
-----
Iteration Number: 1617
Loss: 41.56993973204
12 norm of gradients: 0.5583147068233656
12 norm of weights: 4.036901599764495
-----
Iteration Number: 1618
Loss: 41.54773709983928
12 norm of gradients: 0.558067291012418
12 norm of weights: 4.0357555170371215
-----
Iteration Number: 1619
Loss: 41.52554982939644
12 norm of gradients: 0.5578199656638211
12 norm of weights: 4.0346099408120635
-----
Iteration Number: 1620
```


Loss: 41.50337791030642
l2 norm of gradients: 0.55757273082122
l2 norm of weights: 4.033464870939185

Iteration Number: 1621
Loss: 41.48122133216701
l2 norm of gradients: 0.5573255865280398
l2 norm of weights: 4.03232030726837

Iteration Number: 1622
Loss: 41.45908008458328
l2 norm of gradients: 0.5570785328274849
l2 norm of weights: 4.031176249649511

Iteration Number: 1623
Loss: 41.436954157156784
l2 norm of gradients: 0.5568315697625396
l2 norm of weights: 4.030032697932521

Iteration Number: 1624
Loss: 41.414843539507665
l2 norm of gradients: 0.5565846973759678
l2 norm of weights: 4.028889651967325

Iteration Number: 1625
Loss: 41.39274822124739
l2 norm of gradients: 0.5563379157103134
l2 norm of weights: 4.0277471116038654

Iteration Number: 1626
Loss: 41.370668192002086
l2 norm of gradients: 0.5560912248079
l2 norm of weights: 4.026605076692103

Iteration Number: 1627
Loss: 41.348603441392754
l2 norm of gradients: 0.555844624710831
l2 norm of weights: 4.025463547082014

Iteration Number: 1628
Loss: 41.326553959055865
l2 norm of gradients: 0.5555981154609894
l2 norm of weights: 4.024322522623592

Iteration Number: 1629
Loss: 41.30451973462495
l2 norm of gradients: 0.5553516971000386
l2 norm of weights: 4.023182003166853

```

-----
Iteration Number: 1630
Loss: 41.282500757742135
l2 norm of gradients: 0.5551053696694219
l2 norm of weights: 4.022041988561827
-----
Iteration Number: 1631
Loss: 41.26049701805282
l2 norm of gradients: 0.5548591332103623
l2 norm of weights: 4.020902478658567
-----
Iteration Number: 1632
Loss: 41.23850850520642
l2 norm of gradients: 0.5546129877638631
l2 norm of weights: 4.0197634733071474
-----
Iteration Number: 1633
Loss: 41.216535208859284
l2 norm of gradients: 0.5543669333707079
l2 norm of weights: 4.018624972357659
-----
Iteration Number: 1634
Loss: 41.194577118668654
l2 norm of gradients: 0.5541209700714599
l2 norm of weights: 4.0174869756602165
-----
Iteration Number: 1635
Loss: 41.17263422430452
l2 norm of gradients: 0.5538750979064633
l2 norm of weights: 4.016349483064957
-----
Iteration Number: 1636
Loss: 41.150706515433214
l2 norm of gradients: 0.5536293169158417
l2 norm of weights: 4.01521249442204
-----
Iteration Number: 1637
Loss: 41.128793981728585
l2 norm of gradients: 0.5533836271395001
l2 norm of weights: 4.014076009581646
-----
Iteration Number: 1638
Loss: 41.106896612872454
l2 norm of gradients: 0.5531380286171231
l2 norm of weights: 4.012940028393981
-----
Iteration Number: 1639
Loss: 41.0850143985502

```

l2 norm of gradients: 0.5528925213881761
l2 norm of weights: 4.011804550709274

Iteration Number: 1640
Loss: 41.0631473284502
l2 norm of gradients: 0.552647105491905
l2 norm of weights: 4.010669576377779

Iteration Number: 1641
Loss: 41.04129539226585
l2 norm of gradients: 0.5524017809673362
l2 norm of weights: 4.0095351052497765

Iteration Number: 1642
Loss: 41.01945857969717
l2 norm of gradients: 0.5521565478532768
l2 norm of weights: 4.00840113717557

Iteration Number: 1643
Loss: 40.99763688044918
l2 norm of gradients: 0.5519114061883151
l2 norm of weights: 4.00726767200549

Iteration Number: 1644
Loss: 40.97583028423058
l2 norm of gradients: 0.5516663560108197
l2 norm of weights: 4.006134709589896

Iteration Number: 1645
Loss: 40.95403878075668
l2 norm of gradients: 0.5514213973589402
l2 norm of weights: 4.005002249779172

Iteration Number: 1646
Loss: 40.932262359746275
l2 norm of gradients: 0.5511765302706075
l2 norm of weights: 4.0038702924237315

Iteration Number: 1647
Loss: 40.910501010921216
l2 norm of gradients: 0.5509317547835333
l2 norm of weights: 4.002738837374015

Iteration Number: 1648
Loss: 40.88875472401789
l2 norm of gradients: 0.5506870709352104
l2 norm of weights: 4.001607884480495

Iteration Number: 1649
Loss: 40.867023488763394
l2 norm of gradients: 0.5504424787629127
l2 norm of weights: 4.000477433593669

Iteration Number: 1650
Loss: 40.84530729490247
l2 norm of gradients: 0.5501979783036964
l2 norm of weights: 3.999347484564067

Iteration Number: 1651
Loss: 40.823606132176835
l2 norm of gradients: 0.5499535695943977
l2 norm of weights: 3.9982180372422502

Iteration Number: 1652
Loss: 40.80191999033682
l2 norm of gradients: 0.5497092526716353
l2 norm of weights: 3.997089091478809

Iteration Number: 1653
Loss: 40.780248859139554
l2 norm of gradients: 0.549465027571809
l2 norm of weights: 3.995960647124366

Iteration Number: 1654
Loss: 40.75859272834333
l2 norm of gradients: 0.5492208943311006
l2 norm of weights: 3.9948327040295757

Iteration Number: 1655
Loss: 40.736951587714664
l2 norm of gradients: 0.5489768529854734
l2 norm of weights: 3.9937052620451263

Iteration Number: 1656
Loss: 40.71532542702284
l2 norm of gradients: 0.5487329035706728
l2 norm of weights: 3.9925783210217385

Iteration Number: 1657
Loss: 40.693714236043164
l2 norm of gradients: 0.5484890461222262
l2 norm of weights: 3.9914518808101644

Iteration Number: 1658
Loss: 40.67211800455854
l2 norm of gradients: 0.5482452806754426

```

12 norm of weights: 3.990325941261194
-----
Iteration Number: 1659
Loss: 40.65053672235258
12 norm of gradients: 0.5480016072654141
12 norm of weights: 3.9892005022256503
-----
Iteration Number: 1660
Loss: 40.62897037921944
12 norm of gradients: 0.5477580259270143
12 norm of weights: 3.98807556355439
-----
Iteration Number: 1661
Loss: 40.60741896495358
12 norm of gradients: 0.5475145366948991
12 norm of weights: 3.986951125098307
-----
Iteration Number: 1662
Loss: 40.585882469358296
12 norm of gradients: 0.5472711396035079
12 norm of weights: 3.9858271867083315
-----
Iteration Number: 1663
Loss: 40.56436088223855
12 norm of gradients: 0.5470278346870615
12 norm of weights: 3.98470374823543
-----
Iteration Number: 1664
Loss: 40.54285419340884
12 norm of gradients: 0.5467846219795646
12 norm of weights: 3.9835808095306064
-----
Iteration Number: 1665
Loss: 40.521362392685035
12 norm of gradients: 0.5465415015148035
12 norm of weights: 3.982458370444901
-----
Iteration Number: 1666
Loss: 40.49988546989111
12 norm of gradients: 0.5462984733263483
12 norm of weights: 3.981336430829394
-----
Iteration Number: 1667
Loss: 40.478423414855506
12 norm of gradients: 0.5460555374475521
12 norm of weights: 3.9802149905352047
-----
Iteration Number: 1668

```

Loss: 40.45697621741163
l2 norm of gradients: 0.5458126939115511
l2 norm of weights: 3.9790940494134883

Iteration Number: 1669
Loss: 40.43554386739938
l2 norm of gradients: 0.5455699427512646
l2 norm of weights: 3.9779736073154437

Iteration Number: 1670
Loss: 40.41412635466218
l2 norm of gradients: 0.5453272839993956
l2 norm of weights: 3.9768536640923076

Iteration Number: 1671
Loss: 40.392723669048486
l2 norm of gradients: 0.5450847176884307
l2 norm of weights: 3.975734219595358

Iteration Number: 1672
Loss: 40.37133580041697
l2 norm of gradients: 0.5448422438506398
l2 norm of weights: 3.9746152736759144

Iteration Number: 1673
Loss: 40.34996273862519
l2 norm of gradients: 0.5445998625180771
l2 norm of weights: 3.9734968261853374

Iteration Number: 1674
Loss: 40.3286044735397
l2 norm of gradients: 0.5443575737225804
l2 norm of weights: 3.97237887697503

Iteration Number: 1675
Loss: 40.30726099503269
l2 norm of gradients: 0.544115377495772
l2 norm of weights: 3.971261425896438

Iteration Number: 1676
Loss: 40.28593229297876
l2 norm of gradients: 0.5438732738690578
l2 norm of weights: 3.9701444728010493

Iteration Number: 1677
Loss: 40.26461835726441
l2 norm of gradients: 0.5436312628736286
l2 norm of weights: 3.9690280175403982

```

-----
Iteration Number: 1678
Loss: 40.243319177773856
l2 norm of gradients: 0.5433893445404592
l2 norm of weights: 3.967912059966059
-----
Iteration Number: 1679
Loss: 40.22203474440244
l2 norm of gradients: 0.5431475189003095
l2 norm of weights: 3.9667965999296544
-----
Iteration Number: 1680
Loss: 40.200765047047604
l2 norm of gradients: 0.542905785983724
l2 norm of weights: 3.9656816372828496
-----
Iteration Number: 1681
Loss: 40.17951007561496
l2 norm of gradients: 0.5426641458210317
l2 norm of weights: 3.9645671718773565
-----
Iteration Number: 1682
Loss: 40.1582698200128
l2 norm of gradients: 0.5424225984423473
l2 norm of weights: 3.9634532035649332
-----
Iteration Number: 1683
Loss: 40.13704427015633
l2 norm of gradients: 0.5421811438775703
l2 norm of weights: 3.962339732197383
-----
Iteration Number: 1684
Loss: 40.11583341596936
l2 norm of gradients: 0.5419397821563856
l2 norm of weights: 3.961226757626558
-----
Iteration Number: 1685
Loss: 40.094637247374514
l2 norm of gradients: 0.5416985133082636
l2 norm of weights: 3.960114279704356
-----
Iteration Number: 1686
Loss: 40.07345575430711
l2 norm of gradients: 0.5414573373624603
l2 norm of weights: 3.9590022982827238
-----
Iteration Number: 1687
Loss: 40.052288926702225

```

l2 norm of gradients: 0.5412162543480177
l2 norm of weights: 3.957890813213656

Iteration Number: 1688
Loss: 40.031136754504324
l2 norm of gradients: 0.5409752642937632
l2 norm of weights: 3.9567798243491965

Iteration Number: 1689
Loss: 40.00999922766297
l2 norm of gradients: 0.540734367228311
l2 norm of weights: 3.9556693315414377

Iteration Number: 1690
Loss: 39.9888763361319
l2 norm of gradients: 0.5404935631800611
l2 norm of weights: 3.954559334642523

Iteration Number: 1691
Loss: 39.96776806986964
l2 norm of gradients: 0.5402528521772
l2 norm of weights: 3.9534498335046453

Iteration Number: 1692
Loss: 39.94667441884559
l2 norm of gradients: 0.5400122342477008
l2 norm of weights: 3.9523408279800485

Iteration Number: 1693
Loss: 39.925595373027264
l2 norm of gradients: 0.5397717094193233
l2 norm of weights: 3.9512323179210265

Iteration Number: 1694
Loss: 39.9045309223941
l2 norm of gradients: 0.5395312777196142
l2 norm of weights: 3.950124303179926

Iteration Number: 1695
Loss: 39.88348105692821
l2 norm of gradients: 0.5392909391759072
l2 norm of weights: 3.949016783609145

Iteration Number: 1696
Loss: 39.86244576661998
l2 norm of gradients: 0.5390506938153233
l2 norm of weights: 3.947909759061136

Iteration Number: 1697
Loss: 39.841425041460056
l2 norm of gradients: 0.538810541664771
l2 norm of weights: 3.946803229388401

Iteration Number: 1698
Loss: 39.820418871450464
l2 norm of gradients: 0.538570482750946
l2 norm of weights: 3.9456971944434986

Iteration Number: 1699
Loss: 39.79942724659745
l2 norm of gradients: 0.538330517100332
l2 norm of weights: 3.9445916540790393

Iteration Number: 1700
Loss: 39.77845015691162
l2 norm of gradients: 0.5380906447392005
l2 norm of weights: 3.9434866081476883

Iteration Number: 1701
Loss: 39.75748759240995
l2 norm of gradients: 0.5378508656936112
l2 norm of weights: 3.9423820565021668

Iteration Number: 1702
Loss: 39.73653954311454
l2 norm of gradients: 0.5376111799894119
l2 norm of weights: 3.9412779989952496

Iteration Number: 1703
Loss: 39.71560599905614
l2 norm of gradients: 0.5373715876522387
l2 norm of weights: 3.940174435479767

Iteration Number: 1704
Loss: 39.69468695026843
l2 norm of gradients: 0.5371320887075167
l2 norm of weights: 3.939071365808607

Iteration Number: 1705
Loss: 39.673782386790386
l2 norm of gradients: 0.5368926831804595
l2 norm of weights: 3.9379687898347138

Iteration Number: 1706
Loss: 39.652892298672185
l2 norm of gradients: 0.5366533710960694

```

12 norm of weights: 3.936866707411086
-----
Iteration Number: 1707
Loss: 39.632016675961
12 norm of gradients: 0.5364141524791386
12 norm of weights: 3.9357651183907825
-----
Iteration Number: 1708
Loss: 39.6111555087168
12 norm of gradients: 0.5361750273542477
12 norm of weights: 3.9346640226269196
-----
Iteration Number: 1709
Loss: 39.59030878700407
12 norm of gradients: 0.5359359957457674
12 norm of weights: 3.9335634199726703
-----
Iteration Number: 1710
Loss: 39.56947650089085
12 norm of gradients: 0.5356970576778582
12 norm of weights: 3.9324633102812676
-----
Iteration Number: 1711
Loss: 39.54865864045305
12 norm of gradients: 0.53545821317447
12 norm of weights: 3.931363693406003
-----
Iteration Number: 1712
Loss: 39.52785519577292
12 norm of gradients: 0.5352194622593428
12 norm of weights: 3.930264569200228
-----
Iteration Number: 1713
Loss: 39.50706615693651
12 norm of gradients: 0.5349808049560073
12 norm of weights: 3.9291659375173533
-----
Iteration Number: 1714
Loss: 39.48629151403586
12 norm of gradients: 0.5347422412877844
12 norm of weights: 3.9280677982108507
-----
Iteration Number: 1715
Loss: 39.46553125717221
12 norm of gradients: 0.5345037712777857
12 norm of weights: 3.9269701511342525
-----
Iteration Number: 1716

```

```
Loss: 39.444785376450454
l2 norm of gradients: 0.5342653949489132
l2 norm of weights: 3.9258729961411523
-----
Iteration Number: 1717
Loss: 39.42405386198115
l2 norm of gradients: 0.5340271123238608
l2 norm of weights: 3.924776333085205
-----
Iteration Number: 1718
Loss: 39.40333670387981
l2 norm of gradients: 0.5337889234251129
l2 norm of weights: 3.923680161820127
-----
Iteration Number: 1719
Loss: 39.382633892269986
l2 norm of gradients: 0.5335508282749456
l2 norm of weights: 3.922584482199699
-----
Iteration Number: 1720
Loss: 39.361945417280644
l2 norm of gradients: 0.5333128268954267
l2 norm of weights: 3.9214892940777633
-----
Iteration Number: 1721
Loss: 39.34127126904641
l2 norm of gradients: 0.5330749193084156
l2 norm of weights: 3.9203945973082246
-----
Iteration Number: 1722
Loss: 39.32061143770889
l2 norm of gradients: 0.532837105535564
l2 norm of weights: 3.9193003917450535
-----
Iteration Number: 1723
Loss: 39.29996591341353
l2 norm of gradients: 0.5325993855983157
l2 norm of weights: 3.9182066772422828
-----
Iteration Number: 1724
Loss: 39.27933468631396
l2 norm of gradients: 0.532361759517907
l2 norm of weights: 3.9171134536540104
-----
Iteration Number: 1725
Loss: 39.25871774656937
l2 norm of gradients: 0.5321242273153666
l2 norm of weights: 3.9160207208343993
```

```

-----
Iteration Number: 1726
Loss: 39.23811508434246
l2 norm of gradients: 0.5318867890115165
l2 norm of weights: 3.9149284786376772
-----
Iteration Number: 1727
Loss: 39.217526689806824
l2 norm of gradients: 0.5316494446269715
l2 norm of weights: 3.9138367269181384
-----
Iteration Number: 1728
Loss: 39.19695255313831
l2 norm of gradients: 0.5314121941821397
l2 norm of weights: 3.912745465530142
-----
Iteration Number: 1729
Loss: 39.176392664518914
l2 norm of gradients: 0.5311750376972227
l2 norm of weights: 3.9116546943281154
-----
Iteration Number: 1730
Loss: 39.15584701414038
l2 norm of gradients: 0.5309379751922161
l2 norm of weights: 3.9105644131665507
-----
Iteration Number: 1731
Loss: 39.13531559219491
l2 norm of gradients: 0.5307010066869088
l2 norm of weights: 3.9094746219000087
-----
Iteration Number: 1732
Loss: 39.11479838888506
l2 norm of gradients: 0.5304641322008843
l2 norm of weights: 3.9083853203831183
-----
Iteration Number: 1733
Loss: 39.094295394418076
l2 norm of gradients: 0.5302273517535204
l2 norm of weights: 3.907296508470574
-----
Iteration Number: 1734
Loss: 39.07380659900736
l2 norm of gradients: 0.5299906653639894
l2 norm of weights: 3.9062081860171416
-----
Iteration Number: 1735
Loss: 39.053331992874575

```

l2 norm of gradients: 0.5297540730512585
l2 norm of weights: 3.9051203528776544

Iteration Number: 1736
Loss: 39.0328715662435
l2 norm of gradients: 0.5295175748340897
l2 norm of weights: 3.9040330089070143

Iteration Number: 1737
Loss: 39.01242530934549
l2 norm of gradients: 0.5292811707310405
l2 norm of weights: 3.902946153960194

Iteration Number: 1738
Loss: 38.99199321241896
l2 norm of gradients: 0.5290448607604635
l2 norm of weights: 3.9018597878922345

Iteration Number: 1739
Loss: 38.97157526570837
l2 norm of gradients: 0.5288086449405075
l2 norm of weights: 3.9007739105582493

Iteration Number: 1740
Loss: 38.95117145946675
l2 norm of gradients: 0.5285725232891167
l2 norm of weights: 3.899688521813421

Iteration Number: 1741
Loss: 38.93078178394676
l2 norm of gradients: 0.528336495824032
l2 norm of weights: 3.898603621513004

Iteration Number: 1742
Loss: 38.910406229413645
l2 norm of gradients: 0.5281005625627899
l2 norm of weights: 3.8975192095123234

Iteration Number: 1743
Loss: 38.890044786133785
l2 norm of gradients: 0.5278647235227245
l2 norm of weights: 3.896435285666777

Iteration Number: 1744
Loss: 38.86969744438582
l2 norm of gradients: 0.5276289787209657
l2 norm of weights: 3.8953518498318345

Iteration Number: 1745
Loss: 38.849364194449635
l2 norm of gradients: 0.5273933281744411
l2 norm of weights: 3.8942689018630374

Iteration Number: 1746
Loss: 38.82904502661307
l2 norm of gradients: 0.5271577718998756
l2 norm of weights: 3.893186441616001

Iteration Number: 1747
Loss: 38.80873993116891
l2 norm of gradients: 0.5269223099137912
l2 norm of weights: 3.892104468946414

Iteration Number: 1748
Loss: 38.788448898419574
l2 norm of gradients: 0.5266869422325082
l2 norm of weights: 3.8910229837100374

Iteration Number: 1749
Loss: 38.76817191866994
l2 norm of gradients: 0.5264516688721447
l2 norm of weights: 3.8899419857627078

Iteration Number: 1750
Loss: 38.7479089822324
l2 norm of gradients: 0.5262164898486167
l2 norm of weights: 3.888861474960336

Iteration Number: 1751
Loss: 38.72766007942705
l2 norm of gradients: 0.5259814051776392
l2 norm of weights: 3.8877814511589057

Iteration Number: 1752
Loss: 38.70742520057731
l2 norm of gradients: 0.5257464148747257
l2 norm of weights: 3.8867019142144774

Iteration Number: 1753
Loss: 38.68720433601763
l2 norm of gradients: 0.5255115189551887
l2 norm of weights: 3.885622863983187

Iteration Number: 1754
Loss: 38.666997476084134
l2 norm of gradients: 0.5252767174341397

```

12 norm of weights: 3.884544300321245
-----
Iteration Number: 1755
Loss: 38.64680461112045
12 norm of gradients: 0.5250420103264901
12 norm of weights: 3.8834662230849397
-----
Iteration Number: 1756
Loss: 38.62662573147907
12 norm of gradients: 0.5248073976469506
12 norm of weights: 3.882388632130633
-----
Iteration Number: 1757
Loss: 38.6064608275151
12 norm of gradients: 0.5245728794100322
12 norm of weights: 3.881311527314767
-----
Iteration Number: 1758
Loss: 38.58630988959354
12 norm of gradients: 0.5243384556300456
12 norm of weights: 3.880234908493859
-----
Iteration Number: 1759
Loss: 38.56617290808151
12 norm of gradients: 0.5241041263211025
12 norm of weights: 3.8791587755245036
-----
Iteration Number: 1760
Loss: 38.54604987335656
12 norm of gradients: 0.523869891497115
12 norm of weights: 3.8780831282633743
-----
Iteration Number: 1761
Loss: 38.52594077580147
12 norm of gradients: 0.5236357511717962
12 norm of weights: 3.8770079665672217
-----
Iteration Number: 1762
Loss: 38.50584560580222
12 norm of gradients: 0.5234017053586603
12 norm of weights: 3.875933290292876
-----
Iteration Number: 1763
Loss: 38.48576435375601
12 norm of gradients: 0.523167754071023
12 norm of weights: 3.8748590992972445
-----
Iteration Number: 1764

```

Loss: 38.46569701006428
l2 norm of gradients: 0.522933897322002
l2 norm of weights: 3.8737853934373154

Iteration Number: 1765
Loss: 38.4456435651345
l2 norm of gradients: 0.5227001351245166
l2 norm of weights: 3.872712172570156

Iteration Number: 1766
Loss: 38.425604009379676
l2 norm of gradients: 0.5224664674912884
l2 norm of weights: 3.871639436552912

Iteration Number: 1767
Loss: 38.4055783332231
l2 norm of gradients: 0.5222328944348417
l2 norm of weights: 3.870567185242811

Iteration Number: 1768
Loss: 38.38556652709007
l2 norm of gradients: 0.5219994159675031
l2 norm of weights: 3.86949541849716

Iteration Number: 1769
Loss: 38.36556858141284
l2 norm of gradients: 0.5217660321014027
l2 norm of weights: 3.868424136173348

Iteration Number: 1770
Loss: 38.345584486633925
l2 norm of gradients: 0.5215327428484736
l2 norm of weights: 3.8673533381288427

Iteration Number: 1771
Loss: 38.32561423319832
l2 norm of gradients: 0.5212995482204524
l2 norm of weights: 3.8662830242211954

Iteration Number: 1772
Loss: 38.305657811558824
l2 norm of gradients: 0.5210664482288797
l2 norm of weights: 3.865213194308039

Iteration Number: 1773
Loss: 38.28571521217539
l2 norm of gradients: 0.5208334428850997
l2 norm of weights: 3.864143848247086


```

-----
Iteration Number: 1774
Loss: 38.26578642551356
l2 norm of gradients: 0.5206005322002616
l2 norm of weights: 3.8630749858961346
-----
Iteration Number: 1775
Loss: 38.245871442045626
l2 norm of gradients: 0.5203677161853185
l2 norm of weights: 3.862006607113064
-----
Iteration Number: 1776
Loss: 38.22597025224918
l2 norm of gradients: 0.5201349948510291
l2 norm of weights: 3.860938711755835
-----
Iteration Number: 1777
Loss: 38.20608284661138
l2 norm of gradients: 0.5199023682079563
l2 norm of weights: 3.859871299682495
-----
Iteration Number: 1778
Loss: 38.186209215623286
l2 norm of gradients: 0.519669836266469
l2 norm of weights: 3.8588043707511717
-----
Iteration Number: 1779
Loss: 38.16634934978198
l2 norm of gradients: 0.5194373990367419
l2 norm of weights: 3.8577379248200785
-----
Iteration Number: 1780
Loss: 38.14650323959375
l2 norm of gradients: 0.5192050565287553
l2 norm of weights: 3.8566719617475127
-----
Iteration Number: 1781
Loss: 38.126670875568976
l2 norm of gradients: 0.5189728087522956
l2 norm of weights: 3.8556064813918556
-----
Iteration Number: 1782
Loss: 38.10685224822609
l2 norm of gradients: 0.518740655716956
l2 norm of weights: 3.8545414836115732
-----
Iteration Number: 1783
Loss: 38.087047348089605

```

l2 norm of gradients: 0.5185085974321366
l2 norm of weights: 3.8534769682652175

Iteration Number: 1784
Loss: 38.067256165689216
l2 norm of gradients: 0.518276633907044
l2 norm of weights: 3.8524129352114245

Iteration Number: 1785
Loss: 38.04747869156293
l2 norm of gradients: 0.5180447651506925
l2 norm of weights: 3.851349384308917

Iteration Number: 1786
Loss: 38.02771491625665
l2 norm of gradients: 0.517812991171904
l2 norm of weights: 3.850286315416503

Iteration Number: 1787
Loss: 38.007964830317775
l2 norm of gradients: 0.517581311979308
l2 norm of weights: 3.849223728393077

Iteration Number: 1788
Loss: 37.98822842430476
l2 norm of gradients: 0.5173497275813425
l2 norm of weights: 3.8481616230976203

Iteration Number: 1789
Loss: 37.968505688781725
l2 norm of gradients: 0.5171182379862537
l2 norm of weights: 3.84709999993892

Iteration Number: 1790
Loss: 37.948796614319356
l2 norm of gradients: 0.5168868432020963
l2 norm of weights: 3.846038857126971

Iteration Number: 1791
Loss: 37.92910119149356
l2 norm of gradients: 0.516655543236735
l2 norm of weights: 3.8449781961701754

Iteration Number: 1792
Loss: 37.90941941088833
l2 norm of gradients: 0.5164243380978425
l2 norm of weights: 3.8439180163781432

Iteration Number: 1793
Loss: 37.88975126309426
l2 norm of gradients: 0.5161932277929017
l2 norm of weights: 3.8428583176102915

Iteration Number: 1794
Loss: 37.87009673870829
l2 norm of gradients: 0.5159622123292056
l2 norm of weights: 3.841799099726127

Iteration Number: 1795
Loss: 37.85045582833263
l2 norm of gradients: 0.5157312917138568
l2 norm of weights: 3.8407403625852425

Iteration Number: 1796
Loss: 37.83082852257688
l2 norm of gradients: 0.5155004659537686
l2 norm of weights: 3.839682106047322

Iteration Number: 1797
Loss: 37.811214812060356
l2 norm of gradients: 0.5152697350556653
l2 norm of weights: 3.838624329972136

Iteration Number: 1798
Loss: 37.79161468740272
l2 norm of gradients: 0.5150390990260818
l2 norm of weights: 3.837567034219547

Iteration Number: 1799
Loss: 37.77202813923754
l2 norm of gradients: 0.5148085578713645
l2 norm of weights: 3.836510218649505

Iteration Number: 1800
Loss: 37.75245515819908
l2 norm of gradients: 0.5145781115976711
l2 norm of weights: 3.8354538831220495

Iteration Number: 1801
Loss: 37.732895734931354
l2 norm of gradients: 0.5143477602109715
l2 norm of weights: 3.8343980274973113

Iteration Number: 1802
Loss: 37.713349860083156
l2 norm of gradients: 0.5141175037170482

```
12 norm of weights: 3.8333426516355114
-----
Iteration Number: 1803
Loss: 37.69381752431372
12 norm of gradients: 0.5138873421214952
12 norm of weights: 3.8322877553969596
-----
Iteration Number: 1804
Loss: 37.6742987182825
12 norm of gradients: 0.5136572754297201
12 norm of weights: 3.8312333386420594
-----
Iteration Number: 1805
Loss: 37.654793432662835
12 norm of gradients: 0.5134273036469431
12 norm of weights: 3.830179401231302
-----
Iteration Number: 1806
Loss: 37.635301658129535
12 norm of gradients: 0.5131974267781982
12 norm of weights: 3.829125943025272
-----
Iteration Number: 1807
Loss: 37.61582338536733
12 norm of gradients: 0.5129676448283328
12 norm of weights: 3.8280729638846456
-----
Iteration Number: 1808
Loss: 37.59635860506427
12 norm of gradients: 0.512737957802008
12 norm of weights: 3.827020463670189
-----
Iteration Number: 1809
Loss: 37.57690730791928
12 norm of gradients: 0.5125083657036998
12 norm of weights: 3.825968442242763
-----
Iteration Number: 1810
Loss: 37.55746948463439
12 norm of gradients: 0.5122788685376987
12 norm of weights: 3.824916899463318
-----
Iteration Number: 1811
Loss: 37.538045125919815
12 norm of gradients: 0.5120494663081092
12 norm of weights: 3.823865835192898
-----
Iteration Number: 1812
```

Loss: 37.51863422249309
l2 norm of gradients: 0.5118201590188522
l2 norm of weights: 3.8228152492926406

Iteration Number: 1813
Loss: 37.499236765077214
l2 norm of gradients: 0.5115909466736634
l2 norm of weights: 3.8217651416237746

Iteration Number: 1814
Loss: 37.47985274440277
l2 norm of gradients: 0.5113618292760945
l2 norm of weights: 3.8207155120476233

Iteration Number: 1815
Loss: 37.46048215120809
l2 norm of gradients: 0.5111328068295131
l2 norm of weights: 3.8196663604256034

Iteration Number: 1816
Loss: 37.44112497623462
l2 norm of gradients: 0.5109038793371036
l2 norm of weights: 3.8186176866192247

Iteration Number: 1817
Loss: 37.4217812102354
l2 norm of gradients: 0.5106750468018666
l2 norm of weights: 3.8175694904900905

Iteration Number: 1818
Loss: 37.40245084396589
l2 norm of gradients: 0.5104463092266203
l2 norm of weights: 3.816521771899899

Iteration Number: 1819
Loss: 37.383133868192
l2 norm of gradients: 0.5102176666139998
l2 norm of weights: 3.815474530710443

Iteration Number: 1820
Loss: 37.36383027368318
l2 norm of gradients: 0.5099891189664582
l2 norm of weights: 3.814427766783608

Iteration Number: 1821
Loss: 37.344540051218225
l2 norm of gradients: 0.509760666286266
l2 norm of weights: 3.8133814799813766

```

-----
Iteration Number: 1822
Loss: 37.3252631915804
l2 norm of gradients: 0.5095323085755128
l2 norm of weights: 3.8123356701658246
-----
Iteration Number: 1823
Loss: 37.305999685562625
l2 norm of gradients: 0.5093040458361062
l2 norm of weights: 3.8112903371991242
-----
Iteration Number: 1824
Loss: 37.28674952396104
l2 norm of gradients: 0.509075878069773
l2 norm of weights: 3.8102454809435424
-----
Iteration Number: 1825
Loss: 37.267512697582106
l2 norm of gradients: 0.5088478052780587
l2 norm of weights: 3.8092011012614417
-----
Iteration Number: 1826
Loss: 37.24828919723608
l2 norm of gradients: 0.5086198274623293
l2 norm of weights: 3.8081571980152797
-----
Iteration Number: 1827
Loss: 37.22907901374259
l2 norm of gradients: 0.5083919446237697
l2 norm of weights: 3.8071137710676126
-----
Iteration Number: 1828
Loss: 37.20988213792603
l2 norm of gradients: 0.5081641567633856
l2 norm of weights: 3.8060708202810902
-----
Iteration Number: 1829
Loss: 37.19069856061961
l2 norm of gradients: 0.5079364638820028
l2 norm of weights: 3.8050283455184597
-----
Iteration Number: 1830
Loss: 37.171528272661526
l2 norm of gradients: 0.5077088659802681
l2 norm of weights: 3.8039863466425645
-----
Iteration Number: 1831
Loss: 37.15237126489657

```

l2 norm of gradients: 0.5074813630586494
l2 norm of weights: 3.802944823516346

Iteration Number: 1832
Loss: 37.13322752817858
l2 norm of gradients: 0.5072539551174364
l2 norm of weights: 3.8019037760028414

Iteration Number: 1833
Loss: 37.11409705336672
l2 norm of gradients: 0.5070266421567401
l2 norm of weights: 3.8008632039651853

Iteration Number: 1834
Loss: 37.09497983132649
l2 norm of gradients: 0.5067994241764936
l2 norm of weights: 3.7998231072666107

Iteration Number: 1835
Loss: 37.07587585293158
l2 norm of gradients: 0.5065723011764529
l2 norm of weights: 3.7987834857704463

Iteration Number: 1836
Loss: 37.05678510906257
l2 norm of gradients: 0.5063452731561964
l2 norm of weights: 3.797744339340121

Iteration Number: 1837
Loss: 37.03770759060478
l2 norm of gradients: 0.5061183401151255
l2 norm of weights: 3.7967056678391597

Iteration Number: 1838
Loss: 37.01864328845297
l2 norm of gradients: 0.5058915020524654
l2 norm of weights: 3.7956674711311864

Iteration Number: 1839
Loss: 36.99959219350745
l2 norm of gradients: 0.5056647589672646
l2 norm of weights: 3.7946297490799226

Iteration Number: 1840
Loss: 36.98055429667505
l2 norm of gradients: 0.5054381108583961
l2 norm of weights: 3.7935925015491896

Iteration Number: 1841
Loss: 36.961529588869446
l2 norm of gradients: 0.505211557724557
l2 norm of weights: 3.792555728402907

Iteration Number: 1842
Loss: 36.94251806101313
l2 norm of gradients: 0.5049850995642691
l2 norm of weights: 3.7915194295050925

Iteration Number: 1843
Loss: 36.92351970403381
l2 norm of gradients: 0.5047587363758795
l2 norm of weights: 3.7904836047198627

Iteration Number: 1844
Loss: 36.90453450886561
l2 norm of gradients: 0.5045324681575605
l2 norm of weights: 3.7894482539114356

Iteration Number: 1845
Loss: 36.88556246645016
l2 norm of gradients: 0.50430629490731
l2 norm of weights: 3.788413376944126

Iteration Number: 1846
Loss: 36.86660356773723
l2 norm of gradients: 0.5040802166229525
l2 norm of weights: 3.7873789736823498

Iteration Number: 1847
Loss: 36.84765780368112
l2 norm of gradients: 0.5038542333021385
l2 norm of weights: 3.786345043990621

Iteration Number: 1848
Loss: 36.828725165245174
l2 norm of gradients: 0.503628344942345
l2 norm of weights: 3.785311587733556

Iteration Number: 1849
Loss: 36.80980564339802
l2 norm of gradients: 0.5034025515408763
l2 norm of weights: 3.7842786047758694

Iteration Number: 1850
Loss: 36.7908992291164
l2 norm of gradients: 0.5031768530948645


```

12 norm of weights: 3.7832460949823767
-----
Iteration Number: 1851
Loss: 36.772005913382586
12 norm of gradients: 0.5029512496012689
12 norm of weights: 3.7822140582179933
-----
Iteration Number: 1852
Loss: 36.75312568718779
12 norm of gradients: 0.5027257410568768
12 norm of weights: 3.781182494347735
-----
Iteration Number: 1853
Loss: 36.734258541528
12 norm of gradients: 0.5025003274583044
12 norm of weights: 3.780151403236719
-----
Iteration Number: 1854
Loss: 36.715404467407716
12 norm of gradients: 0.5022750088019966
12 norm of weights: 3.7791207847501633
-----
Iteration Number: 1855
Loss: 36.696563455836305
12 norm of gradients: 0.5020497850842269
12 norm of weights: 3.778090638753386
-----
Iteration Number: 1856
Loss: 36.67773549783416
12 norm of gradients: 0.5018246563010987
12 norm of weights: 3.777060965111807
-----
Iteration Number: 1857
Loss: 36.658920584423015
12 norm of gradients: 0.5015996224485452
12 norm of weights: 3.776031763690947
-----
Iteration Number: 1858
Loss: 36.640118706636116
12 norm of gradients: 0.5013746835223292
12 norm of weights: 3.775003034356429
-----
Iteration Number: 1859
Loss: 36.62132985551151
12 norm of gradients: 0.5011498395180446
12 norm of weights: 3.7739747769739767
-----
Iteration Number: 1860

```

```

Loss: 36.6025540220951
12 norm of gradients: 0.5009250904311161
12 norm of weights: 3.7729469914094165
-----
Iteration Number: 1861
Loss: 36.58379119743898
12 norm of gradients: 0.5007004362567987
12 norm of weights: 3.771919677528675
-----
Iteration Number: 1862
Loss: 36.56504137260192
12 norm of gradients: 0.5004758769901801
12 norm of weights: 3.7708928351977815
-----
Iteration Number: 1863
Loss: 36.54630453865026
12 norm of gradients: 0.5002514126261791
12 norm of weights: 3.769866464282868
-----
Iteration Number: 1864
Loss: 36.52758068665838
12 norm of gradients: 0.500027043159547
12 norm of weights: 3.768840564650169
-----
Iteration Number: 1865
Loss: 36.50886980770573
12 norm of gradients: 0.49980276858486733
12 norm of weights: 3.767815136166019
-----
Iteration Number: 1866
Loss: 36.490171892878664
12 norm of gradients: 0.4995785888965568
12 norm of weights: 3.766790178696858
-----
Iteration Number: 1867
Loss: 36.47148693327254
12 norm of gradients: 0.4993545040888655
12 norm of weights: 3.7657656921092273
-----
Iteration Number: 1868
Loss: 36.45281491998801
12 norm of gradients: 0.4991305141558768
12 norm of weights: 3.76474167626977
-----
Iteration Number: 1869
Loss: 36.43415584413189
12 norm of gradients: 0.49890661909150813
12 norm of weights: 3.7637181310452332

```

```

-----
Iteration Number: 1870
Loss: 36.41550969682124
l2 norm of gradients: 0.4986828188895114
l2 norm of weights: 3.7626950563024666
-----
Iteration Number: 1871
Loss: 36.396876469176696
l2 norm of gradients: 0.4984591135434728
l2 norm of weights: 3.7616724519084235
-----
Iteration Number: 1872
Loss: 36.37825615232778
l2 norm of gradients: 0.49823550304681397
l2 norm of weights: 3.76065031773016
-----
Iteration Number: 1873
Loss: 36.35964873740946
l2 norm of gradients: 0.4980119873927918
l2 norm of weights: 3.759628653634835
-----
Iteration Number: 1874
Loss: 36.34105421556462
l2 norm of gradients: 0.49778856657449855
l2 norm of weights: 3.7586074594897116
-----
Iteration Number: 1875
Loss: 36.322472577944986
l2 norm of gradients: 0.4975652405848632
l2 norm of weights: 3.7575867351621564
-----
Iteration Number: 1876
Loss: 36.303903815705056
l2 norm of gradients: 0.49734200941665047
l2 norm of weights: 3.7565664805196386
-----
Iteration Number: 1877
Loss: 36.28534792001031
l2 norm of gradients: 0.49711887306246255
l2 norm of weights: 3.7555466954297327
-----
Iteration Number: 1878
Loss: 36.26680488203126
l2 norm of gradients: 0.4968958315147382
l2 norm of weights: 3.754527379760116
-----
Iteration Number: 1879
Loss: 36.248274692944676

```

l2 norm of gradients: 0.49667288476575405
l2 norm of weights: 3.7535085333785703

Iteration Number: 1880
Loss: 36.22975734393575
l2 norm of gradients: 0.49645003280762473
l2 norm of weights: 3.7524901561529806

Iteration Number: 1881
Loss: 36.21125282619761
l2 norm of gradients: 0.4962272756323024
l2 norm of weights: 3.751472247951337

Iteration Number: 1882
Loss: 36.19276113092731
l2 norm of gradients: 0.49600461323157863
l2 norm of weights: 3.750454808641733

Iteration Number: 1883
Loss: 36.174282249331135
l2 norm of gradients: 0.49578204559708355
l2 norm of weights: 3.7494378380923683

Iteration Number: 1884
Loss: 36.15581617262171
l2 norm of gradients: 0.4955595727202868
l2 norm of weights: 3.748421336171544

Iteration Number: 1885
Loss: 36.13736289201944
l2 norm of gradients: 0.49533719459249725
l2 norm of weights: 3.7474053027476675

Iteration Number: 1886
Loss: 36.11892239875007
l2 norm of gradients: 0.49511491120486456
l2 norm of weights: 3.746389737689251

Iteration Number: 1887
Loss: 36.10049468404899
l2 norm of gradients: 0.49489272254837813
l2 norm of weights: 3.745374640864912

Iteration Number: 1888
Loss: 36.082079739155944
l2 norm of gradients: 0.49467062861386846
l2 norm of weights: 3.7443600121433693

Iteration Number: 1889
Loss: 36.06367755531807
l2 norm of gradients: 0.4944486293920072
l2 norm of weights: 3.743345851393451

Iteration Number: 1890
Loss: 36.04528812379146
l2 norm of gradients: 0.4942267248733074
l2 norm of weights: 3.742332158484087

Iteration Number: 1891
Loss: 36.02691143583681
l2 norm of gradients: 0.49400491504812394
l2 norm of weights: 3.741318933284314

Iteration Number: 1892
Loss: 36.00854748272353
l2 norm of gradients: 0.4937831999066542
l2 norm of weights: 3.7403061756632723

Iteration Number: 1893
Loss: 35.99019625572739
l2 norm of gradients: 0.49356157943893797
l2 norm of weights: 3.7392938854902087

Iteration Number: 1894
Loss: 35.971857746131185
l2 norm of gradients: 0.49334005363485817
l2 norm of weights: 3.7382820626344744

Iteration Number: 1895
Loss: 35.95353194522465
l2 norm of gradients: 0.49311862248414084
l2 norm of weights: 3.7372707069655253

Iteration Number: 1896
Loss: 35.935218844305666
l2 norm of gradients: 0.49289728597635607
l2 norm of weights: 3.7362598183529245

Iteration Number: 1897
Loss: 35.91691843467642
l2 norm of gradients: 0.4926760441009176
l2 norm of weights: 3.735249396666339

Iteration Number: 1898
Loss: 35.89863070764859
l2 norm of gradients: 0.4924548968470841

```

12 norm of weights: 3.734239441775541
-----
Iteration Number: 1899
Loss: 35.88035565454055
12 norm of gradients: 0.4922338442039588
12 norm of weights: 3.7332299535504103
-----
Iteration Number: 1900
Loss: 35.862093266677036
12 norm of gradients: 0.49201288616049055
12 norm of weights: 3.7322209318609296
-----
Iteration Number: 1901
Loss: 35.843843535389944
12 norm of gradients: 0.49179202270547306
12 norm of weights: 3.7312123765771887
-----
Iteration Number: 1902
Loss: 35.82560645201787
12 norm of gradients: 0.4915712538275468
12 norm of weights: 3.730204287569383
-----
Iteration Number: 1903
Loss: 35.807382007907364
12 norm of gradients: 0.4913505795151979
12 norm of weights: 3.729196664707814
-----
Iteration Number: 1904
Loss: 35.789170194412
12 norm of gradients: 0.49112999975675975
12 norm of weights: 3.728189507862888
-----
Iteration Number: 1905
Loss: 35.77097100289105
12 norm of gradients: 0.4909095145404126
12 norm of weights: 3.7271828169051178
-----
Iteration Number: 1906
Loss: 35.75278442471143
12 norm of gradients: 0.49068912385418406
12 norm of weights: 3.7261765917051215
-----
Iteration Number: 1907
Loss: 35.73461045124791
12 norm of gradients: 0.4904688276859498
12 norm of weights: 3.725170832133624
-----
Iteration Number: 1908

```

```

Loss: 35.71644907388142
l2 norm of gradients: 0.49024862602343355
l2 norm of weights: 3.724165538061456
-----
Iteration Number: 1909
Loss: 35.69830028400044
l2 norm of gradients: 0.49002851885420795
l2 norm of weights: 3.723160709359554
-----
Iteration Number: 1910
Loss: 35.6801640730001
l2 norm of gradients: 0.4898085061656942
l2 norm of weights: 3.7221563458989593
-----
Iteration Number: 1911
Loss: 35.662040432282524
l2 norm of gradients: 0.48958858794516324
l2 norm of weights: 3.7211524475508218
-----
Iteration Number: 1912
Loss: 35.643929353256816
l2 norm of gradients: 0.4893687641797355
l2 norm of weights: 3.7201490141863953
-----
Iteration Number: 1913
Loss: 35.625830827340145
l2 norm of gradients: 0.4891490348563817
l2 norm of weights: 3.7191460456770415
-----
Iteration Number: 1914
Loss: 35.60774484595569
l2 norm of gradients: 0.48892939996192303
l2 norm of weights: 3.7181435418942272
-----
Iteration Number: 1915
Loss: 35.58967140053272
l2 norm of gradients: 0.4887098594830316
l2 norm of weights: 3.7171415027095245
-----
Iteration Number: 1916
Loss: 35.57161048250952
l2 norm of gradients: 0.48849041340623056
l2 norm of weights: 3.7161399279946146
-----
Iteration Number: 1917
Loss: 35.55356208333161
l2 norm of gradients: 0.48827106171789514
l2 norm of weights: 3.715138817621282

```

```

-----
Iteration Number: 1918
Loss: 35.53552619444846
l2 norm of gradients: 0.48805180440425244
l2 norm of weights: 3.7141381714614186
-----
Iteration Number: 1919
Loss: 35.5175028073202
l2 norm of gradients: 0.48783264145138194
l2 norm of weights: 3.713137989387023
-----
Iteration Number: 1920
Loss: 35.49949191341162
l2 norm of gradients: 0.487613572845216
l2 norm of weights: 3.7121382712702
-----
Iteration Number: 1921
Loss: 35.48149350419582
l2 norm of gradients: 0.48739459857154
l2 norm of weights: 3.7111390169831604
-----
Iteration Number: 1922
Loss: 35.46350757115232
l2 norm of gradients: 0.48717571861599335
l2 norm of weights: 3.7101402263982206
-----
Iteration Number: 1923
Loss: 35.44553410576753
l2 norm of gradients: 0.48695693296406906
l2 norm of weights: 3.709141899387806
-----
Iteration Number: 1924
Loss: 35.427573099535785
l2 norm of gradients: 0.4867382416011146
l2 norm of weights: 3.708144035824444
-----
Iteration Number: 1925
Loss: 35.4096245439568
l2 norm of gradients: 0.4865196445123322
l2 norm of weights: 3.7071466355807736
-----
Iteration Number: 1926
Loss: 35.39168843053982
l2 norm of gradients: 0.48630114168277944
l2 norm of weights: 3.706149698529535
-----
Iteration Number: 1927
Loss: 35.373764750798145

```



```

12 norm of gradients: 0.48608273309736905
12 norm of weights: 3.705153224543579
-----
Iteration Number: 1928
Loss: 35.355853496255264
12 norm of gradients: 0.4858644187408703
12 norm of weights: 3.7041572134958605
-----
Iteration Number: 1929
Loss: 35.33795465843906
12 norm of gradients: 0.4856461985979079
12 norm of weights: 3.703161665259441
-----
Iteration Number: 1930
Loss: 35.3200682288858
12 norm of gradients: 0.4854280726529642
12 norm of weights: 3.702166579707489
-----
Iteration Number: 1931
Loss: 35.30219419913818
12 norm of gradients: 0.485210040890378
12 norm of weights: 3.701171956713279
-----
Iteration Number: 1932
Loss: 35.28433256074684
12 norm of gradients: 0.48499210329434583
12 norm of weights: 3.7001777961501916
-----
Iteration Number: 1933
Loss: 35.26648330526871
12 norm of gradients: 0.48477425984892225
12 norm of weights: 3.6991840978917145
-----
Iteration Number: 1934
Loss: 35.24864642426675
12 norm of gradients: 0.48455651053801996
12 norm of weights: 3.698190861811441
-----
Iteration Number: 1935
Loss: 35.230821909313256
12 norm of gradients: 0.4843388553454102
12 norm of weights: 3.697198087783071
-----
Iteration Number: 1936
Loss: 35.213009751986355
12 norm of gradients: 0.4841212942547236
12 norm of weights: 3.696205775680411
-----

```

Iteration Number: 1937
Loss: 35.1952099438709
l2 norm of gradients: 0.48390382724945014
l2 norm of weights: 3.6952139253773737

Iteration Number: 1938
Loss: 35.17742247655914
l2 norm of gradients: 0.48368645431293955
l2 norm of weights: 3.6942225367479775

Iteration Number: 1939
Loss: 35.159647341650356
l2 norm of gradients: 0.48346917542840184
l2 norm of weights: 3.693231609666348

Iteration Number: 1940
Loss: 35.14188453075039
l2 norm of gradients: 0.48325199057890783
l2 norm of weights: 3.6922411440067155

Iteration Number: 1941
Loss: 35.12413403547367
l2 norm of gradients: 0.48303489974738945
l2 norm of weights: 3.691251139643419

Iteration Number: 1942
Loss: 35.10639584743955
l2 norm of gradients: 0.4828179029166398
l2 norm of weights: 3.6902615964509007

Iteration Number: 1943
Loss: 35.08866995827614
l2 norm of gradients: 0.48260100006931395
l2 norm of weights: 3.689272514303712

Iteration Number: 1944
Loss: 35.07095635961682
l2 norm of gradients: 0.4823841911879295
l2 norm of weights: 3.688283893076508

Iteration Number: 1945
Loss: 35.053255043103626
l2 norm of gradients: 0.48216747625486645
l2 norm of weights: 3.687295732644052

Iteration Number: 1946
Loss: 35.035566000385344
l2 norm of gradients: 0.48195085525236775

```

12 norm of weights: 3.6863080328812106
-----
Iteration Number: 1947
Loss: 35.01788922311669
12 norm of gradients: 0.4817343281625399
12 norm of weights: 3.685320793662959
-----
Iteration Number: 1948
Loss: 35.00022470296065
12 norm of gradients: 0.4815178949673536
12 norm of weights: 3.6843340148643775
-----
Iteration Number: 1949
Loss: 34.98257243158641
12 norm of gradients: 0.48130155564864324
12 norm of weights: 3.6833476963606517
-----
Iteration Number: 1950
Loss: 34.964932400670556
12 norm of gradients: 0.48108531018810824
12 norm of weights: 3.6823618380270746
-----
Iteration Number: 1951
Loss: 34.94730460189705
12 norm of gradients: 0.480869158567313
12 norm of weights: 3.681376439739043
-----
Iteration Number: 1952
Loss: 34.92968902695621
12 norm of gradients: 0.48065310076768736
12 norm of weights: 3.6803915013720627
-----
Iteration Number: 1953
Loss: 34.912085667545455
12 norm of gradients: 0.48043713677052713
12 norm of weights: 3.6794070228017413
-----
Iteration Number: 1954
Loss: 34.89449451536947
12 norm of gradients: 0.4802212665569943
12 norm of weights: 3.678423003903795
-----
Iteration Number: 1955
Loss: 34.87691556213955
12 norm of gradients: 0.48000549010811755
12 norm of weights: 3.677439444554045
-----
Iteration Number: 1956

```

```

Loss: 34.85934879957512
12 norm of gradients: 0.4797898074047927
12 norm of weights: 3.6764563446284173
-----
Iteration Number: 1957
Loss: 34.841794219401166
12 norm of gradients: 0.479574218427783
12 norm of weights: 3.675473704002945
-----
Iteration Number: 1958
Loss: 34.82425181334998
12 norm of gradients: 0.4793587231577196
12 norm of weights: 3.6744915225537653
-----
Iteration Number: 1959
Loss: 34.80672157316244
12 norm of gradients: 0.479143321575102
12 norm of weights: 3.6735098001571216
-----
Iteration Number: 1960
Loss: 34.7892034905839
12 norm of gradients: 0.4789280136602983
12 norm of weights: 3.6725285366893616
-----
Iteration Number: 1961
Loss: 34.77169755736889
12 norm of gradients: 0.47871279939354605
12 norm of weights: 3.671547732026941
-----
Iteration Number: 1962
Loss: 34.75420376527702
12 norm of gradients: 0.4784976787549518
12 norm of weights: 3.6705673860464176
-----
Iteration Number: 1963
Loss: 34.736722106077266
12 norm of gradients: 0.4782826517244924
12 norm of weights: 3.6695874986244563
-----
Iteration Number: 1964
Loss: 34.719252571544104
12 norm of gradients: 0.4780677182820152
12 norm of weights: 3.668608069637827
-----
Iteration Number: 1965
Loss: 34.701795153457994
12 norm of gradients: 0.47785287840723767
12 norm of weights: 3.667629098963404

```

```

-----
Iteration Number: 1966
Loss: 34.6843498436083
l2 norm of gradients: 0.47763813207974914
l2 norm of weights: 3.666650586478166
-----
Iteration Number: 1967
Loss: 34.66691663379091
l2 norm of gradients: 0.4774234792790101
l2 norm of weights: 3.6656725320591996
-----
Iteration Number: 1968
Loss: 34.6494955158081
l2 norm of gradients: 0.4772089199843531
l2 norm of weights: 3.6646949355836926
-----
Iteration Number: 1969
Loss: 34.632086481469436
l2 norm of gradients: 0.4769944541749832
l2 norm of weights: 3.66371779692894
-----
Iteration Number: 1970
Loss: 34.614689522591455
l2 norm of gradients: 0.47678008182997805
l2 norm of weights: 3.6627411159723406
-----
Iteration Number: 1971
Loss: 34.59730463099832
l2 norm of gradients: 0.47656580292828876
l2 norm of weights: 3.6617648925913975
-----
Iteration Number: 1972
Loss: 34.579931798520086
l2 norm of gradients: 0.47635161744873994
l2 norm of weights: 3.660789126663719
-----
Iteration Number: 1973
Loss: 34.56257101699445
l2 norm of gradients: 0.47613752537003035
l2 norm of weights: 3.659813818067017
-----
Iteration Number: 1974
Loss: 34.54522227826588
l2 norm of gradients: 0.47592352667073273
l2 norm of weights: 3.6588389666791095
-----
Iteration Number: 1975
Loss: 34.52788557418607

```

```

12 norm of gradients: 0.4757096213292955
12 norm of weights: 3.6578645723779175
-----
Iteration Number: 1976
Loss: 34.51056089661248
12 norm of gradients: 0.4754958093240417
12 norm of weights: 3.656890635041466
-----
Iteration Number: 1977
Loss: 34.49324823741198
12 norm of gradients: 0.47528209063317034
12 norm of weights: 3.6559171545478835
-----
Iteration Number: 1978
Loss: 34.475947588456464
12 norm of gradients: 0.47506846523475654
12 norm of weights: 3.6549441307754047
-----
Iteration Number: 1979
Loss: 34.45865894162517
12 norm of gradients: 0.4748549331067518
12 norm of weights: 3.6539715636023664
-----
Iteration Number: 1980
Loss: 34.44138228880455
12 norm of gradients: 0.47464149422698476
12 norm of weights: 3.6529994529072103
-----
Iteration Number: 1981
Loss: 34.42411762188721
12 norm of gradients: 0.4744281485731611
12 norm of weights: 3.65202779856848
-----
Iteration Number: 1982
Loss: 34.40686493277425
12 norm of gradients: 0.4742148961228647
12 norm of weights: 3.651056600464826
-----
Iteration Number: 1983
Loss: 34.38962421337258
12 norm of gradients: 0.47400173685355734
12 norm of weights: 3.6500858584749984
-----
Iteration Number: 1984
Loss: 34.37239545559642
12 norm of gradients: 0.47378867074257935
12 norm of weights: 3.6491155724778537
-----

```

Iteration Number: 1985
Loss: 34.355178651367225
l2 norm of gradients: 0.47357569776715036
l2 norm of weights: 3.64814574235235

Iteration Number: 1986
Loss: 34.33797379261226
l2 norm of gradients: 0.4733628179043692
l2 norm of weights: 3.6471763679775493

Iteration Number: 1987
Loss: 34.32078087126693
l2 norm of gradients: 0.47315003113121457
l2 norm of weights: 3.6462074492326173

Iteration Number: 1988
Loss: 34.303599879273506
l2 norm of gradients: 0.47293733742454547
l2 norm of weights: 3.6452389859968206

Iteration Number: 1989
Loss: 34.28643080858011
l2 norm of gradients: 0.47272473676110177
l2 norm of weights: 3.6442709781495313

Iteration Number: 1990
Loss: 34.26927365114299
l2 norm of gradients: 0.472512229117504
l2 norm of weights: 3.6433034255702226

Iteration Number: 1991
Loss: 34.25212839892492
l2 norm of gradients: 0.47229981447025493
l2 norm of weights: 3.6423363281384704

Iteration Number: 1992
Loss: 34.23499504389603
l2 norm of gradients: 0.47208749279573864
l2 norm of weights: 3.6413696857339537

Iteration Number: 1993
Loss: 34.21787357803246
l2 norm of gradients: 0.4718752640702218
l2 norm of weights: 3.6404034982364535

Iteration Number: 1994
Loss: 34.200763993317125
l2 norm of gradients: 0.4716631282698539

```

12 norm of weights: 3.6394377655258534
-----
Iteration Number: 1995
Loss: 34.18366628174113
12 norm of gradients: 0.47145108537066754
12 norm of weights: 3.6384724874821384
-----
Iteration Number: 1996
Loss: 34.16658043530169
12 norm of gradients: 0.47123913534857914
12 norm of weights: 3.637507663985397
-----
Iteration Number: 1997
Loss: 34.14950644600378
12 norm of gradients: 0.47102727817938894
12 norm of weights: 3.6365432949158185
-----
Iteration Number: 1998
Loss: 34.13244430585739
12 norm of gradients: 0.47081551383878173
12 norm of weights: 3.6355793801536933
-----
Iteration Number: 1999
Loss: 34.11539400688122
12 norm of gradients: 0.47060384230232716
12 norm of weights: 3.6346159195794154
-----
Iteration Number: 2000
Loss: 34.098355541100446
12 norm of gradients: 0.4703922635454806
12 norm of weights: 3.633652913073478
-----
Iteration Number: 2001
Loss: 34.08132890054629
12 norm of gradients: 0.47018077754358234
12 norm of weights: 3.6326903605164778
-----
Iteration Number: 2002
Loss: 34.06431407725821
12 norm of gradients: 0.46996938427185947
12 norm of weights: 3.6317282617891116
-----
Iteration Number: 2003
Loss: 34.047311063282294
12 norm of gradients: 0.4697580837054256
12 norm of weights: 3.630766616772178
-----
Iteration Number: 2004

```



```

Loss: 34.03031985066965
l2 norm of gradients: 0.469546875819281
l2 norm of weights: 3.6298054253465755
-----
Iteration Number: 2005
Loss: 34.01334043148115
l2 norm of gradients: 0.4693357605883137
l2 norm of weights: 3.6288446873933045
-----
Iteration Number: 2006
Loss: 33.996372797782634
l2 norm of gradients: 0.4691247379872995
l2 norm of weights: 3.627884402793465
-----
Iteration Number: 2007
Loss: 33.9794169416469
l2 norm of gradients: 0.46891380799090243
l2 norm of weights: 3.626924571428259
-----
Iteration Number: 2008
Loss: 33.96247285515484
l2 norm of gradients: 0.46870297057367505
l2 norm of weights: 3.6259651931789882
-----
Iteration Number: 2009
Loss: 33.94554053039295
l2 norm of gradients: 0.46849222571005933
l2 norm of weights: 3.625006267927054
-----
Iteration Number: 2010
Loss: 33.92861995945502
l2 norm of gradients: 0.46828157337438664
l2 norm of weights: 3.624047795553958
-----
Iteration Number: 2011
Loss: 33.9117111344423
l2 norm of gradients: 0.4680710135408784
l2 norm of weights: 3.6230897759413025
-----
Iteration Number: 2012
Loss: 33.894814047461324
l2 norm of gradients: 0.46786054618364625
l2 norm of weights: 3.6221322089707897
-----
Iteration Number: 2013
Loss: 33.87792869062759
l2 norm of gradients: 0.46765017127669284
l2 norm of weights: 3.621175094524221

```

```

-----
Iteration Number: 2014
Loss: 33.86105505606184
l2 norm of gradients: 0.46743988879391196
l2 norm of weights: 3.6202184324834965
-----
Iteration Number: 2015
Loss: 33.844193135892475
l2 norm of gradients: 0.46722969870908915
l2 norm of weights: 3.6192622227306175
-----
Iteration Number: 2016
Loss: 33.82734292225426
l2 norm of gradients: 0.46701960099590184
l2 norm of weights: 3.6183064651476826
-----
Iteration Number: 2017
Loss: 33.81050440728907
l2 norm of gradients: 0.4668095956279203
l2 norm of weights: 3.617351159616891
-----
Iteration Number: 2018
Loss: 33.793677583145644
l2 norm of gradients: 0.4665996825786075
l2 norm of weights: 3.6163963060205386
-----
Iteration Number: 2019
Loss: 33.776862441979205
l2 norm of gradients: 0.4663898618213199
l2 norm of weights: 3.615441904241024
-----
Iteration Number: 2020
Loss: 33.760058975952646
l2 norm of gradients: 0.4661801333293076
l2 norm of weights: 3.61448795416084
-----
Iteration Number: 2021
Loss: 33.743267177234664
l2 norm of gradients: 0.46597049707571514
l2 norm of weights: 3.6135344556625797
-----
Iteration Number: 2022
Loss: 33.72648703800107
l2 norm of gradients: 0.4657609530335816
l2 norm of weights: 3.612581408628935
-----
Iteration Number: 2023
Loss: 33.70971855043451

```

```

12 norm of gradients: 0.4655515011758411
12 norm of weights: 3.611628812942695
-----
Iteration Number: 2024
Loss: 33.692961706725605
12 norm of gradients: 0.4653421414753234
12 norm of weights: 3.6106766684867466
-----
Iteration Number: 2025
Loss: 33.67621649906965
12 norm of gradients: 0.465132873904754
12 norm of weights: 3.6097249751440748
-----
Iteration Number: 2026
Loss: 33.65948291967036
12 norm of gradients: 0.464923698436755
12 norm of weights: 3.608773732797762
-----
Iteration Number: 2027
Loss: 33.64276096073789
12 norm of gradients: 0.464714615043845
12 norm of weights: 3.607822941330988
-----
Iteration Number: 2028
Loss: 33.62605061448881
12 norm of gradients: 0.46450562369844023
12 norm of weights: 3.60687260062703
-----
Iteration Number: 2029
Loss: 33.6093518731464
12 norm of gradients: 0.46429672437285413
12 norm of weights: 3.605922710569261
-----
Iteration Number: 2030
Loss: 33.5926647289417
12 norm of gradients: 0.4640879170392983
12 norm of weights: 3.6049732710411524
-----
Iteration Number: 2031
Loss: 33.5759891741114
12 norm of gradients: 0.4638792016698831
12 norm of weights: 3.6040242819262716
-----
Iteration Number: 2032
Loss: 33.559325200899906
12 norm of gradients: 0.46367057823661767
12 norm of weights: 3.603075743108283
-----

```

Iteration Number: 2033
Loss: 33.54267280155685
l2 norm of gradients: 0.46346204671141045
l2 norm of weights: 3.6021276544709457

Iteration Number: 2034
Loss: 33.526031968340966
l2 norm of gradients: 0.46325360706606955
l2 norm of weights: 3.601180015898117

Iteration Number: 2035
Loss: 33.50940269351609
l2 norm of gradients: 0.4630452592723038
l2 norm of weights: 3.6002328272737483

Iteration Number: 2036
Loss: 33.49278496935283
l2 norm of gradients: 0.462837003301722
l2 norm of weights: 3.599286088481888

Iteration Number: 2037
Loss: 33.476178788129424
l2 norm of gradients: 0.4626288391258346
l2 norm of weights: 3.5983397994066793

Iteration Number: 2038
Loss: 33.45958414212994
l2 norm of gradients: 0.46242076671605303
l2 norm of weights: 3.597393959932361

Iteration Number: 2039
Loss: 33.44300102364631
l2 norm of gradients: 0.46221278604369104
l2 norm of weights: 3.596448569943267

Iteration Number: 2040
Loss: 33.426429424975694
l2 norm of gradients: 0.46200489707996484
l2 norm of weights: 3.5955036293238263

Iteration Number: 2041
Loss: 33.40986933842339
l2 norm of gradients: 0.46179709979599287
l2 norm of weights: 3.5945591379585626

Iteration Number: 2042
Loss: 33.39332075630087
l2 norm of gradients: 0.4615893941627972

```

12 norm of weights: 3.5936150957320945
-----
Iteration Number: 2043
Loss: 33.3767836709258
12 norm of gradients: 0.46138178015130366
12 norm of weights: 3.592671502529133
-----
Iteration Number: 2044
Loss: 33.360258074623864
12 norm of gradients: 0.4611742577323419
12 norm of weights: 3.5917283582344868
-----
Iteration Number: 2045
Loss: 33.34374395972583
12 norm of gradients: 0.4609668268766461
12 norm of weights: 3.590785662733056
-----
Iteration Number: 2046
Loss: 33.32724131857083
12 norm of gradients: 0.4607594875548554
12 norm of weights: 3.5898434159098347
-----
Iteration Number: 2047
Loss: 33.31075014350306
12 norm of gradients: 0.4605522397375144
12 norm of weights: 3.5889016176499116
-----
Iteration Number: 2048
Loss: 33.29427042687476
12 norm of gradients: 0.46034508339507346
12 norm of weights: 3.5879602678384677
-----
Iteration Number: 2049
Loss: 33.277802161044555
12 norm of gradients: 0.4601380184978891
12 norm of weights: 3.587019366360778
-----
Iteration Number: 2050
Loss: 33.261345338377154
12 norm of gradients: 0.4599310450162245
12 norm of weights: 3.5860789131022095
-----
Iteration Number: 2051
Loss: 33.244899951244676
12 norm of gradients: 0.45972416292025
12 norm of weights: 3.585138907948223
-----
Iteration Number: 2052

```

Loss: 33.228465992025185
l2 norm of gradients: 0.45951737218004346
l2 norm of weights: 3.584199350784371

Iteration Number: 2053
Loss: 33.21204345310446
l2 norm of gradients: 0.45931067276559046
l2 norm of weights: 3.5832602414962986

Iteration Number: 2054
Loss: 33.19563232687389
l2 norm of gradients: 0.4591040646467855
l2 norm of weights: 3.5823215799697437

Iteration Number: 2055
Loss: 33.17923260573215
l2 norm of gradients: 0.4588975477934311
l2 norm of weights: 3.581383366090535

Iteration Number: 2056
Loss: 33.1628442820841
l2 norm of gradients: 0.4586911221752399
l2 norm of weights: 3.5804455997445928

Iteration Number: 2057
Loss: 33.14646734834174
l2 norm of gradients: 0.4584847877618334
l2 norm of weights: 3.57950828081793

Iteration Number: 2058
Loss: 33.1301017969238
l2 norm of gradients: 0.4582785445227436
l2 norm of weights: 3.5785714091966505

Iteration Number: 2059
Loss: 33.11374762025488
l2 norm of gradients: 0.4580723924274131
l2 norm of weights: 3.577634984766948

Iteration Number: 2060
Loss: 33.097404810766896
l2 norm of gradients: 0.45786633144519523
l2 norm of weights: 3.5766990074151086

Iteration Number: 2061
Loss: 33.08107336089877
l2 norm of gradients: 0.4576603615453549
l2 norm of weights: 3.575763477027508

```

-----
Iteration Number: 2062
Loss: 33.06475326309495
l2 norm of gradients: 0.4574544826970686
l2 norm of weights: 3.574828393490612
-----
Iteration Number: 2063
Loss: 33.04844450980656
l2 norm of gradients: 0.4572486948694254
l2 norm of weights: 3.5738937566909783
-----
Iteration Number: 2064
Loss: 33.032147093493066
l2 norm of gradients: 0.4570429980314267
l2 norm of weights: 3.572959566515252
-----
Iteration Number: 2065
Loss: 33.01586100661835
l2 norm of gradients: 0.45683739215198726
l2 norm of weights: 3.57202582285017
-----
Iteration Number: 2066
Loss: 32.99958624165404
l2 norm of gradients: 0.4566318771999352
l2 norm of weights: 3.571092525582557
-----
Iteration Number: 2067
Loss: 32.983322791078365
l2 norm of gradients: 0.4564264531440128
l2 norm of weights: 3.5701596745993283
-----
Iteration Number: 2068
Loss: 32.96707064737591
l2 norm of gradients: 0.4562211199528766
l2 norm of weights: 3.569227269787488
-----
Iteration Number: 2069
Loss: 32.95082980303795
l2 norm of gradients: 0.45601587759509793
l2 norm of weights: 3.568295311034128
-----
Iteration Number: 2070
Loss: 32.934600250562205
l2 norm of gradients: 0.45581072603916334
l2 norm of weights: 3.5673637982264292
-----
Iteration Number: 2071
Loss: 32.91838198245326

```

```

12 norm of gradients: 0.4556056652534754
12 norm of weights: 3.5664327312516613
-----
Iteration Number: 2072
Loss: 32.90217499122136
12 norm of gradients: 0.45540069520635246
12 norm of weights: 3.565502109997182
-----
Iteration Number: 2073
Loss: 32.88597926938503
12 norm of gradients: 0.45519581586602953
12 norm of weights: 3.5645719343504356
-----
Iteration Number: 2074
Loss: 32.86979480946753
12 norm of gradients: 0.4549910272006585
12 norm of weights: 3.5636422041989557
-----
Iteration Number: 2075
Loss: 32.85362160400007
12 norm of gradients: 0.4547863291783088
12 norm of weights: 3.5627129194303624
-----
Iteration Number: 2076
Loss: 32.83745964551956
12 norm of gradients: 0.45458172176696765
12 norm of weights: 3.561784079932363
-----
Iteration Number: 2077
Loss: 32.82130892656957
12 norm of gradients: 0.4543772049345404
12 norm of weights: 3.560855685592751
-----
Iteration Number: 2078
Loss: 32.805169439700435
12 norm of gradients: 0.45417277864885136
12 norm of weights: 3.559927736299408
-----
Iteration Number: 2079
Loss: 32.78904117746916
12 norm of gradients: 0.45396844287764365
12 norm of weights: 3.559000231940302
-----
Iteration Number: 2080
Loss: 32.77292413243865
12 norm of gradients: 0.4537641975885803
12 norm of weights: 3.5580731724034838
-----

```


Iteration Number: 2081
Loss: 32.7568182971793
l2 norm of gradients: 0.45356004274924405
l2 norm of weights: 3.5571465575770946

Iteration Number: 2082
Loss: 32.740723664267215
l2 norm of gradients: 0.45335597832713787
l2 norm of weights: 3.5562203873493585

Iteration Number: 2083
Loss: 32.72464022628493
l2 norm of gradients: 0.45315200428968594
l2 norm of weights: 3.555294661608586

Iteration Number: 2084
Loss: 32.70856797582151
l2 norm of gradients: 0.4529481206042336
l2 norm of weights: 3.5543693802431733

Iteration Number: 2085
Loss: 32.6925069054738
l2 norm of gradients: 0.45274432723804786
l2 norm of weights: 3.5534445431415995

Iteration Number: 2086
Loss: 32.676457007843226
l2 norm of gradients: 0.45254062415831775
l2 norm of weights: 3.5525201501924304

Iteration Number: 2087
Loss: 32.66041827553899
l2 norm of gradients: 0.45233701133215504
l2 norm of weights: 3.5515962012843154

Iteration Number: 2088
Loss: 32.644390701175865
l2 norm of gradients: 0.45213348872659426
l2 norm of weights: 3.5506726963059876

Iteration Number: 2089
Loss: 32.62837427737663
l2 norm of gradients: 0.45193005630859356
l2 norm of weights: 3.5497496351462643

Iteration Number: 2090
Loss: 32.61236899676839
l2 norm of gradients: 0.4517267140450345

```
12 norm of weights: 3.5488270176940477
-----
Iteration Number: 2091
Loss: 32.596374851986305
12 norm of gradients: 0.4515234619027235
12 norm of weights: 3.5479048438383214
-----
Iteration Number: 2092
Loss: 32.58039183567116
12 norm of gradients: 0.45132029984839134
12 norm of weights: 3.546983113468153
-----
Iteration Number: 2093
Loss: 32.56441994047063
12 norm of gradients: 0.4511172278486938
12 norm of weights: 3.5460618264726933
-----
Iteration Number: 2094
Loss: 32.54845915903902
12 norm of gradients: 0.4509142458702122
12 norm of weights: 3.545140982741175
-----
Iteration Number: 2095
Loss: 32.53250948403635
12 norm of gradients: 0.45071135387945427
12 norm of weights: 3.544220582162914
-----
Iteration Number: 2096
Loss: 32.51657090812974
12 norm of gradients: 0.45050855184285354
12 norm of weights: 3.543300624627308
-----
Iteration Number: 2097
Loss: 32.500643423992
12 norm of gradients: 0.4503058397267708
12 norm of weights: 3.5423811100238356
-----
Iteration Number: 2098
Loss: 32.484727024302934
12 norm of gradients: 0.45010321749749377
12 norm of weights: 3.5414620382420585
-----
Iteration Number: 2099
Loss: 32.46882170174881
12 norm of gradients: 0.44990068512123804
12 norm of weights: 3.540543409171619
-----
Iteration Number: 2100
```

Loss: 32.45292744902217
l2 norm of gradients: 0.44969824256414725
l2 norm of weights: 3.5396252227022393

Iteration Number: 2101
Loss: 32.43704425882128
l2 norm of gradients: 0.4494958897922934
l2 norm of weights: 3.5387074787237247

Iteration Number: 2102
Loss: 32.42117212385178
l2 norm of gradients: 0.44929362677167767
l2 norm of weights: 3.5377901771259586

Iteration Number: 2103
Loss: 32.405311036825054
l2 norm of gradients: 0.4490914534682305
l2 norm of weights: 3.536873317798907

Iteration Number: 2104
Loss: 32.38946099045955
l2 norm of gradients: 0.448889369847812
l2 norm of weights: 3.5359569006326135

Iteration Number: 2105
Loss: 32.3736219774793
l2 norm of gradients: 0.4486873758762127
l2 norm of weights: 3.5350409255172037

Iteration Number: 2106
Loss: 32.35779399061456
l2 norm of gradients: 0.44848547151915363
l2 norm of weights: 3.5341253923428804

Iteration Number: 2107
Loss: 32.34197702260289
l2 norm of gradients: 0.4482836567422871
l2 norm of weights: 3.5332103009999276

Iteration Number: 2108
Loss: 32.32617106618771
l2 norm of gradients: 0.4480819315111966
l2 norm of weights: 3.5322956513787065

Iteration Number: 2109
Loss: 32.31037611411846
l2 norm of gradients: 0.4478802957913976
l2 norm of weights: 3.531381443369658

```

-----
Iteration Number: 2110
Loss: 32.29459215915136
l2 norm of gradients: 0.4476787495483381
l2 norm of weights: 3.5304676768633008
-----
Iteration Number: 2111
Loss: 32.27881919404879
l2 norm of gradients: 0.44747729274739867
l2 norm of weights: 3.5295543517502317
-----
Iteration Number: 2112
Loss: 32.26305721157956
l2 norm of gradients: 0.4472759253538931
l2 norm of weights: 3.528641467921126
-----
Iteration Number: 2113
Loss: 32.247306204518345
l2 norm of gradients: 0.4470746473330688
l2 norm of weights: 3.5277290252667353
-----
Iteration Number: 2114
Loss: 32.231566165646754
l2 norm of gradients: 0.44687345865010736
l2 norm of weights: 3.5268170236778893
-----
Iteration Number: 2115
Loss: 32.21583708775233
l2 norm of gradients: 0.44667235927012444
l2 norm of weights: 3.5259054630454942
-----
Iteration Number: 2116
Loss: 32.20011896362841
l2 norm of gradients: 0.44647134915817077
l2 norm of weights: 3.5249943432605337
-----
Iteration Number: 2117
Loss: 32.1844117860762
l2 norm of gradients: 0.4462704282792326
l2 norm of weights: 3.524083664214067
-----
Iteration Number: 2118
Loss: 32.16871554790134
l2 norm of gradients: 0.4460695965982314
l2 norm of weights: 3.5231734257972294
-----
Iteration Number: 2119
Loss: 32.15303024191691

```

```

12 norm of gradients: 0.4458688540800252
12 norm of weights: 3.5222636279012325
-----
Iteration Number: 2120
Loss: 32.13735586094191
12 norm of gradients: 0.4456682006894082
12 norm of weights: 3.5213542704173637
-----
Iteration Number: 2121
Loss: 32.121692397801375
12 norm of gradients: 0.4454676363911119
12 norm of weights: 3.520445353236985
-----
Iteration Number: 2122
Loss: 32.10603984532685
12 norm of gradients: 0.445267161149805
12 norm of weights: 3.519536876251533
-----
Iteration Number: 2123
Loss: 32.090398196356176
12 norm of gradients: 0.44506677493009406
12 norm of weights: 3.5186288393525205
-----
Iteration Number: 2124
Loss: 32.07476744373304
12 norm of gradients: 0.4448664776965239
12 norm of weights: 3.517721242431534
-----
Iteration Number: 2125
Loss: 32.05914758030811
12 norm of gradients: 0.44466626941357784
12 norm of weights: 3.5168140853802328
-----
Iteration Number: 2126
Loss: 32.04353859893736
12 norm of gradients: 0.4444661500456783
12 norm of weights: 3.515907368090352
-----
Iteration Number: 2127
Loss: 32.02794049248347
12 norm of gradients: 0.44426611955718753
12 norm of weights: 3.5150010904536995
-----
Iteration Number: 2128
Loss: 32.01235325381527
12 norm of gradients: 0.44406617791240693
12 norm of weights: 3.5140952523621567
-----

```

Iteration Number: 2129
Loss: 31.99677687580799
l2 norm of gradients: 0.44386632507557894
l2 norm of weights: 3.513189853707677

Iteration Number: 2130
Loss: 31.98121135134254
l2 norm of gradients: 0.4436665610108862
l2 norm of weights: 3.5122848943822875

Iteration Number: 2131
Loss: 31.965656673306608
l2 norm of gradients: 0.4434668856824529
l2 norm of weights: 3.511380374278088

Iteration Number: 2132
Loss: 31.95011283459361
l2 norm of gradients: 0.44326729905434453
l2 norm of weights: 3.510476293287249

Iteration Number: 2133
Loss: 31.934579828103256
l2 norm of gradients: 0.44306780109056876
l2 norm of weights: 3.509572651302013

Iteration Number: 2134
Loss: 31.919057646741532
l2 norm of gradients: 0.4428683917550753
l2 norm of weights: 3.5086694482146967

Iteration Number: 2135
Loss: 31.903546283420404
l2 norm of gradients: 0.44266907101175695
l2 norm of weights: 3.5077666839176844

Iteration Number: 2136
Loss: 31.888045731057876
l2 norm of gradients: 0.4424698388244497
l2 norm of weights: 3.5068643583034333

Iteration Number: 2137
Loss: 31.872555982578806
l2 norm of gradients: 0.442270695156933
l2 norm of weights: 3.505962471264471

Iteration Number: 2138
Loss: 31.85707703091355
l2 norm of gradients: 0.44207163997293053

```

12 norm of weights: 3.5050610226933947
-----
Iteration Number: 2139
Loss: 31.841608868998264
12 norm of gradients: 0.4418726732361104
12 norm of weights: 3.5041600124828727
-----
Iteration Number: 2140
Loss: 31.82615148977618
12 norm of gradients: 0.4416737949100854
12 norm of weights: 3.5032594405256425
-----
Iteration Number: 2141
Loss: 31.810704886195733
12 norm of gradients: 0.44147500495841385
12 norm of weights: 3.502359306714511
-----
Iteration Number: 2142
Loss: 31.79526905121198
12 norm of gradients: 0.44127630334459966
12 norm of weights: 3.501459610942355
-----
Iteration Number: 2143
Loss: 31.77984397778614
12 norm of gradients: 0.4410776900320927
12 norm of weights: 3.5005603531021188
-----
Iteration Number: 2144
Loss: 31.764429658885405
12 norm of gradients: 0.4408791649842896
12 norm of weights: 3.4996615330868157
-----
Iteration Number: 2145
Loss: 31.749026087482328
12 norm of gradients: 0.4406807281645336
12 norm of weights: 3.498763150789529
-----
Iteration Number: 2146
Loss: 31.733633256556782
12 norm of gradients: 0.4404823795361157
12 norm of weights: 3.4978652061034072
-----
Iteration Number: 2147
Loss: 31.718251159094123
12 norm of gradients: 0.44028411906227427
12 norm of weights: 3.4969676989216683
-----
Iteration Number: 2148

```

Loss: 31.702879788085433
l2 norm of gradients: 0.4400859467061959
l2 norm of weights: 3.4960706291375976

Iteration Number: 2149
Loss: 31.687519136528188
l2 norm of gradients: 0.4398878624310158
l2 norm of weights: 3.4951739966445463

Iteration Number: 2150
Loss: 31.67216919742617
l2 norm of gradients: 0.43968986619981826
l2 norm of weights: 3.494277801335933

Iteration Number: 2151
Loss: 31.656829963788788
l2 norm of gradients: 0.43949195797563667
l2 norm of weights: 3.4933820431052434

Iteration Number: 2152
Loss: 31.641501428631504
l2 norm of gradients: 0.43929413772145454
l2 norm of weights: 3.4924867218460283

Iteration Number: 2153
Loss: 31.626183584976026
l2 norm of gradients: 0.43909640540020506
l2 norm of weights: 3.4915918374519057

Iteration Number: 2154
Loss: 31.610876425849575
l2 norm of gradients: 0.43889876097477254
l2 norm of weights: 3.4906973898165568

Iteration Number: 2155
Loss: 31.595579944286413
l2 norm of gradients: 0.43870120440799204
l2 norm of weights: 3.489803378833731

Iteration Number: 2156
Loss: 31.580294133325612
l2 norm of gradients: 0.43850373566265005
l2 norm of weights: 3.48890980439724

Iteration Number: 2157
Loss: 31.56501898601278
l2 norm of gradients: 0.4383063547014847
l2 norm of weights: 3.488016666400962


```

-----
Iteration Number: 2158
Loss: 31.549754495399966
l2 norm of gradients: 0.4381090614871867
l2 norm of weights: 3.4871239647388386
-----
Iteration Number: 2159
Loss: 31.534500654544242
l2 norm of gradients: 0.4379118559823989
l2 norm of weights: 3.4862316993048754
-----
Iteration Number: 2160
Loss: 31.51925745650929
l2 norm of gradients: 0.4377147381497178
l2 norm of weights: 3.4853398699931417
-----
Iteration Number: 2161
Loss: 31.504024894364424
l2 norm of gradients: 0.43751770795169254
l2 norm of weights: 3.484448476697771
-----
Iteration Number: 2162
Loss: 31.488802961185293
l2 norm of gradients: 0.43732076535082665
l2 norm of weights: 3.4835575193129595
-----
Iteration Number: 2163
Loss: 31.473591650053283
l2 norm of gradients: 0.43712391030957765
l2 norm of weights: 3.4826669977329647
-----
Iteration Number: 2164
Loss: 31.45839095405568
l2 norm of gradients: 0.4369271427903577
l2 norm of weights: 3.48177691185211
-----
Iteration Number: 2165
Loss: 31.44320086628566
l2 norm of gradients: 0.43673046275553423
l2 norm of weights: 3.4808872615647766
-----
Iteration Number: 2166
Loss: 31.428021379842136
l2 norm of gradients: 0.43653387016742956
l2 norm of weights: 3.4799980467654112
-----
Iteration Number: 2167
Loss: 31.41285248783074

```

```

12 norm of gradients: 0.43633736498832215
12 norm of weights: 3.4791092673485204
-----
Iteration Number: 2168
Loss: 31.397694183362045
12 norm of gradients: 0.4361409471804469
12 norm of weights: 3.4782209232086716
-----
Iteration Number: 2169
Loss: 31.382546459553012
12 norm of gradients: 0.4359446167059949
12 norm of weights: 3.4773330142404943
-----
Iteration Number: 2170
Loss: 31.36740930952674
12 norm of gradients: 0.4357483735271144
12 norm of weights: 3.4764455403386787
-----
Iteration Number: 2171
Loss: 31.352282726411527
12 norm of gradients: 0.4355522176059112
12 norm of weights: 3.4755585013979737
-----
Iteration Number: 2172
Loss: 31.337166703341737
12 norm of gradients: 0.4353561489044486
12 norm of weights: 3.47467189731319
-----
Iteration Number: 2173
Loss: 31.322061233457966
12 norm of gradients: 0.43516016738474844
12 norm of weights: 3.473785727979197
-----
Iteration Number: 2174
Loss: 31.3069663099068
12 norm of gradients: 0.4349642730087909
12 norm of weights: 3.4728999932909232
-----
Iteration Number: 2175
Loss: 31.29188192583995
12 norm of gradients: 0.4347684657385154
12 norm of weights: 3.472014693143357
-----
Iteration Number: 2176
Loss: 31.27680807441564
12 norm of gradients: 0.43457274553582026
12 norm of weights: 3.4711298274315454
-----

```

Iteration Number: 2177
Loss: 31.26174474879748
l2 norm of gradients: 0.4343771123625642
l2 norm of weights: 3.470245396050594

Iteration Number: 2178
Loss: 31.246691942154925
l2 norm of gradients: 0.4341815661805657
l2 norm of weights: 3.4693613988956646

Iteration Number: 2179
Loss: 31.231649647663833
l2 norm of gradients: 0.4339861069516039
l2 norm of weights: 3.46847783586198

Iteration Number: 2180
Loss: 31.21661785850501
l2 norm of gradients: 0.4337907346374187
l2 norm of weights: 3.467594706844818

Iteration Number: 2181
Loss: 31.20159656786602
l2 norm of gradients: 0.4335954491997116
l2 norm of weights: 3.466712011739514

Iteration Number: 2182
Loss: 31.18658576893936
l2 norm of gradients: 0.433400250600146
l2 norm of weights: 3.4658297504414617

Iteration Number: 2183
Loss: 31.17158545492349
l2 norm of gradients: 0.4332051388003468
l2 norm of weights: 3.4649479228461093

Iteration Number: 2184
Loss: 31.15659561902321
l2 norm of gradients: 0.433010113761902
l2 norm of weights: 3.464066528848963

Iteration Number: 2185
Loss: 31.141616254448774
l2 norm of gradients: 0.4328151754463622
l2 norm of weights: 3.4631855683455837

Iteration Number: 2186
Loss: 31.1266473544159
l2 norm of gradients: 0.43262032381524135

```

12 norm of weights: 3.462305041231588
-----
Iteration Number: 2187
Loss: 31.1116889121463
12 norm of gradients: 0.432425558830017
12 norm of weights: 3.461424947402649
-----
Iteration Number: 2188
Loss: 31.096740920867592
12 norm of gradients: 0.4322308804521308
12 norm of weights: 3.460545286754493
-----
Iteration Number: 2189
Loss: 31.081803373812942
12 norm of gradients: 0.43203628864298893
12 norm of weights: 3.4596660591829025
-----
Iteration Number: 2190
Loss: 31.066876264221044
12 norm of gradients: 0.43184178336396234
12 norm of weights: 3.458787264583713
-----
Iteration Number: 2191
Loss: 31.051959585336974
12 norm of gradients: 0.43164736457638714
12 norm of weights: 3.4579089028528154
-----
Iteration Number: 2192
Loss: 31.037053330410583
12 norm of gradients: 0.43145303224156495
12 norm of weights: 3.4570309738861527
-----
Iteration Number: 2193
Loss: 31.02215749269869
12 norm of gradients: 0.4312587863207635
12 norm of weights: 3.4561534775797234
-----
Iteration Number: 2194
Loss: 31.00727206546243
12 norm of gradients: 0.4310646267752169
12 norm of weights: 3.4552764138295764
-----
Iteration Number: 2195
Loss: 30.992397041969703
12 norm of gradients: 0.43087055356612586
12 norm of weights: 3.454399782531816
-----
Iteration Number: 2196

```

```
Loss: 30.977532415493346
12 norm of gradients: 0.43067656665465837
12 norm of weights: 3.4535235835825966
-----
Iteration Number: 2197
Loss: 30.962678179312437
12 norm of gradients: 0.4304826660019498
12 norm of weights: 3.452647816878127
-----
Iteration Number: 2198
Loss: 30.94783432671153
12 norm of gradients: 0.4302888515691034
12 norm of weights: 3.4517724823146656
-----
Iteration Number: 2199
Loss: 30.933000850980644
12 norm of gradients: 0.43009512331719074
12 norm of weights: 3.450897579788523
-----
Iteration Number: 2200
Loss: 30.918177745415612
12 norm of gradients: 0.4299014812072519
12 norm of weights: 3.4500231091960623
-----
Iteration Number: 2201
Loss: 30.903365003317976
12 norm of gradients: 0.42970792520029616
12 norm of weights: 3.449149070433696
-----
Iteration Number: 2202
Loss: 30.888562617994914
12 norm of gradients: 0.4295144552573021
12 norm of weights: 3.448275463397887
-----
Iteration Number: 2203
Loss: 30.87377058275896
12 norm of gradients: 0.4293210713392179
12 norm of weights: 3.4474022879851494
-----
Iteration Number: 2204
Loss: 30.858988890928824
12 norm of gradients: 0.429127773406962
12 norm of weights: 3.4465295440920465
-----
Iteration Number: 2205
Loss: 30.84421753582807
12 norm of gradients: 0.42893456142142355
12 norm of weights: 3.445657231615191
```

```

-----
Iteration Number: 2206
Loss: 30.829456510786464
l2 norm of gradients: 0.4287414353434623
l2 norm of weights: 3.444785350451246
-----
Iteration Number: 2207
Loss: 30.814705809139166
l2 norm of gradients: 0.4285483951339095
l2 norm of weights: 3.4439139004969213
-----
Iteration Number: 2208
Loss: 30.79996542422695
l2 norm of gradients: 0.42835544075356763
l2 norm of weights: 3.4430428816489784
-----
Iteration Number: 2209
Loss: 30.785235349396093
l2 norm of gradients: 0.4281625721632117
l2 norm of weights: 3.4421722938042243
-----
Iteration Number: 2210
Loss: 30.77051557799876
l2 norm of gradients: 0.4279697893235886
l2 norm of weights: 3.441302136859515
-----
Iteration Number: 2211
Loss: 30.755806103392022
l2 norm of gradients: 0.4277770921954182
l2 norm of weights: 3.4404324107117548
-----
Iteration Number: 2212
Loss: 30.741106918939124
l2 norm of gradients: 0.4275844807393936
l2 norm of weights: 3.439563115257894
-----
Iteration Number: 2213
Loss: 30.72641801800865
l2 norm of gradients: 0.4273919549161812
l2 norm of weights: 3.4386942503949314
-----
Iteration Number: 2214
Loss: 30.711739393974682
l2 norm of gradients: 0.42719951468642114
l2 norm of weights: 3.4378258160199113
-----
Iteration Number: 2215
Loss: 30.697071040217047

```

```

12 norm of gradients: 0.4270071600107281
12 norm of weights: 3.4369578120299247
-----
Iteration Number: 2216
Loss: 30.682412950120558
12 norm of gradients: 0.426814890849691
12 norm of weights: 3.4360902383221092
-----
Iteration Number: 2217
Loss: 30.66776511707623
12 norm of gradients: 0.426622707163874
12 norm of weights: 3.435223094793647
-----
Iteration Number: 2218
Loss: 30.65312753447997
12 norm of gradients: 0.4264306089138163
12 norm of weights: 3.4343563813417672
-----
Iteration Number: 2219
Loss: 30.638500195733727
12 norm of gradients: 0.426238596060033
12 norm of weights: 3.4334900978637424
-----
Iteration Number: 2220
Loss: 30.623883094244505
12 norm of gradients: 0.4260466685630151
12 norm of weights: 3.4326242442568913
-----
Iteration Number: 2221
Loss: 30.6092762234251
12 norm of gradients: 0.42585482638323
12 norm of weights: 3.4317588204185765
-----
Iteration Number: 2222
Loss: 30.594679576693704
12 norm of gradients: 0.42566306948112187
12 norm of weights: 3.4308938262462045
-----
Iteration Number: 2223
Loss: 30.580093147473708
12 norm of gradients: 0.42547139781711213
12 norm of weights: 3.430029261637226
-----
Iteration Number: 2224
Loss: 30.56551692919414
12 norm of gradients: 0.42527981135159953
12 norm of weights: 3.429165126489135
-----

```

Iteration Number: 2225
Loss: 30.550950915289786
l2 norm of gradients: 0.42508831004496045
l2 norm of weights: 3.428301420699469

Iteration Number: 2226
Loss: 30.53639509920021
l2 norm of gradients: 0.4248968938575501
l2 norm of weights: 3.4274381441658077

Iteration Number: 2227
Loss: 30.52184947437134
l2 norm of gradients: 0.42470556274970156
l2 norm of weights: 3.426575296785774

Iteration Number: 2228
Loss: 30.507314034253447
l2 norm of gradients: 0.4245143166817273
l2 norm of weights: 3.4257128784570328

Iteration Number: 2229
Loss: 30.492788772303136
l2 norm of gradients: 0.42432315561391876
l2 norm of weights: 3.4248508890772906

Iteration Number: 2230
Loss: 30.47827368198177
l2 norm of gradients: 0.4241320795065471
l2 norm of weights: 3.423989328544296

Iteration Number: 2231
Loss: 30.4637687567567
l2 norm of gradients: 0.42394108831986355
l2 norm of weights: 3.4231281967558393

Iteration Number: 2232
Loss: 30.449273990100007
l2 norm of gradients: 0.4237501820140994
l2 norm of weights: 3.4222674936097497

Iteration Number: 2233
Loss: 30.434789375489544
l2 norm of gradients: 0.42355936054946697
l2 norm of weights: 3.421407219003899

Iteration Number: 2234
Loss: 30.420314906408805
l2 norm of gradients: 0.42336862388615937


```

12 norm of weights: 3.4205473728361993
-----
Iteration Number: 2235
Loss: 30.405850576346097
12 norm of gradients: 0.42317797198435125
12 norm of weights: 3.419687955004601
-----
Iteration Number: 2236
Loss: 30.391396378795168
12 norm of gradients: 0.4229874048041988
12 norm of weights: 3.418828965407096
-----
Iteration Number: 2237
Loss: 30.376952307255728
12 norm of gradients: 0.4227969223058405
12 norm of weights: 3.4179704039417147
-----
Iteration Number: 2238
Loss: 30.362518355232098
12 norm of gradients: 0.422606524449397
12 norm of weights: 3.4171122705065264
-----
Iteration Number: 2239
Loss: 30.348094516234173
12 norm of gradients: 0.42241621119497214
12 norm of weights: 3.4162545649996403
-----
Iteration Number: 2240
Loss: 30.33368078377718
12 norm of gradients: 0.4222259825026525
12 norm of weights: 3.4153972873192022
-----
Iteration Number: 2241
Loss: 30.319277151381982
12 norm of gradients: 0.42203583833250813
12 norm of weights: 3.4145404373633967
-----
Iteration Number: 2242
Loss: 30.30488361257429
12 norm of gradients: 0.4218457786445933
12 norm of weights: 3.413684015030447
-----
Iteration Number: 2243
Loss: 30.290500160885447
12 norm of gradients: 0.4216558033989459
12 norm of weights: 3.412828020218613
-----
Iteration Number: 2244

```

```

Loss: 30.276126789851567
12 norm of gradients: 0.42146591255558885
12 norm of weights: 3.4119724528261917
-----
Iteration Number: 2245
Loss: 30.261763493014683
12 norm of gradients: 0.42127610607452953
12 norm of weights: 3.411117312751517
-----
Iteration Number: 2246
Loss: 30.247410263921957
12 norm of gradients: 0.4210863839157606
12 norm of weights: 3.410262599892959
-----
Iteration Number: 2247
Loss: 30.233067096125627
12 norm of gradients: 0.4208967460392604
12 norm of weights: 3.409408314148926
-----
Iteration Number: 2248
Loss: 30.218733983183018
12 norm of gradients: 0.42070719240499305
12 norm of weights: 3.408554455417858
-----
Iteration Number: 2249
Loss: 30.204410918657192
12 norm of gradients: 0.4205177229729087
12 norm of weights: 3.4077010235982352
-----
Iteration Number: 2250
Loss: 30.19009789611634
12 norm of gradients: 0.42032833770294425
12 norm of weights: 3.4068480185885694
-----
Iteration Number: 2251
Loss: 30.17579490913357
12 norm of gradients: 0.4201390365550234
12 norm of weights: 3.4059954402874104
-----
Iteration Number: 2252
Loss: 30.16150195128741
12 norm of gradients: 0.4199498194890572
12 norm of weights: 3.4051432885933393
-----
Iteration Number: 2253
Loss: 30.14721901616168
12 norm of gradients: 0.4197606864649441
12 norm of weights: 3.404291563404975

```

```

-----
Iteration Number: 2254
Loss: 30.13294609734549
l2 norm of gradients: 0.4195716374425704
l2 norm of weights: 3.403440264620967
-----
Iteration Number: 2255
Loss: 30.118683188432676
l2 norm of gradients: 0.4193826723818109
l2 norm of weights: 3.4025893921400012
-----
Iteration Number: 2256
Loss: 30.104430283022804
l2 norm of gradients: 0.4191937912425286
l2 norm of weights: 3.4017389458607954
-----
Iteration Number: 2257
Loss: 30.090187374720536
l2 norm of gradients: 0.41900499398457564
l2 norm of weights: 3.4008889256821004
-----
Iteration Number: 2258
Loss: 30.075954457135467
l2 norm of gradients: 0.41881628056779324
l2 norm of weights: 3.4000393315027004
-----
Iteration Number: 2259
Loss: 30.061731523882575
l2 norm of gradients: 0.4186276509520123
l2 norm of weights: 3.3991901632214123
-----
Iteration Number: 2260
Loss: 30.047518568581676
l2 norm of gradients: 0.4184391050970535
l2 norm of weights: 3.3983414207370837
-----
Iteration Number: 2261
Loss: 30.03331558485819
l2 norm of gradients: 0.4182506429627276
l2 norm of weights: 3.397493103948595
-----
Iteration Number: 2262
Loss: 30.019122566342507
l2 norm of gradients: 0.4180622645088364
l2 norm of weights: 3.3966452127548576
-----
Iteration Number: 2263
Loss: 30.00493950667007

```

```

12 norm of gradients: 0.41787396969517177
12 norm of weights: 3.395797747054815
-----
Iteration Number: 2264
Loss: 29.9907663994815
12 norm of gradients: 0.4176857584815174
12 norm of weights: 3.3949507067474407
-----
Iteration Number: 2265
Loss: 29.97660323842268
12 norm of gradients: 0.41749763082764835
12 norm of weights: 3.3941040917317387
-----
Iteration Number: 2266
Loss: 29.962450017144178
12 norm of gradients: 0.41730958669333135
12 norm of weights: 3.3932579019067433
-----
Iteration Number: 2267
Loss: 29.948306729301986
12 norm of gradients: 0.4171216260383255
12 norm of weights: 3.3924121371715197
-----
Iteration Number: 2268
Loss: 29.934173368557538
12 norm of gradients: 0.41693374882238204
12 norm of weights: 3.3915667974251607
-----
Iteration Number: 2269
Loss: 29.920049928576674
12 norm of gradients: 0.4167459550052456
12 norm of weights: 3.39072188256679
-----
Iteration Number: 2270
Loss: 29.90593640303065
12 norm of gradients: 0.4165582445466532
12 norm of weights: 3.38987739249556
-----
Iteration Number: 2271
Loss: 29.891832785595778
12 norm of gradients: 0.416370617406336
12 norm of weights: 3.389033327110652
-----
Iteration Number: 2272
Loss: 29.877739069953247
12 norm of gradients: 0.4161830735440184
12 norm of weights: 3.3881896863112746
-----

```

Iteration Number: 2273
Loss: 29.863655249789804
l2 norm of gradients: 0.4159956129194191
l2 norm of weights: 3.3873464699966656

Iteration Number: 2274
Loss: 29.8495813187969
l2 norm of gradients: 0.4158082354922512
l2 norm of weights: 3.3865036780660898

Iteration Number: 2275
Loss: 29.835517270670795
l2 norm of gradients: 0.4156209412222224
l2 norm of weights: 3.3856613104188398

Iteration Number: 2276
Loss: 29.821463099113124
l2 norm of gradients: 0.41543373006903556
l2 norm of weights: 3.3848193669542357

Iteration Number: 2277
Loss: 29.807418797830387
l2 norm of gradients: 0.41524660199238883
l2 norm of weights: 3.383977847571624

Iteration Number: 2278
Loss: 29.793384360534283
l2 norm of gradients: 0.415059556951976
l2 norm of weights: 3.3831367521703766

Iteration Number: 2279
Loss: 29.779359780941228
l2 norm of gradients: 0.41487259490748674
l2 norm of weights: 3.382296080649894

Iteration Number: 2280
Loss: 29.765345052772954
l2 norm of gradients: 0.4146857158186072
l2 norm of weights: 3.381455832909601

Iteration Number: 2281
Loss: 29.751340169755935
l2 norm of gradients: 0.41449891964501967
l2 norm of weights: 3.380616008848948

Iteration Number: 2282
Loss: 29.737345125621697
l2 norm of gradients: 0.4143122063464041

```

12 norm of weights: 3.3797766083674126
-----
Iteration Number: 2283
Loss: 29.723359914106787
12 norm of gradients: 0.41412557588243676
12 norm of weights: 3.3789376313644937
-----
Iteration Number: 2284
Loss: 29.70938452895255
12 norm of gradients: 0.4139390282127921
12 norm of weights: 3.378099077739718
-----
Iteration Number: 2285
Loss: 29.695418963905816
12 norm of gradients: 0.413752563297142
12 norm of weights: 3.377260947392636
-----
Iteration Number: 2286
Loss: 29.681463212717446
12 norm of gradients: 0.41356618109515675
12 norm of weights: 3.376423240222821
-----
Iteration Number: 2287
Loss: 29.66751726914434
12 norm of gradients: 0.4133798815665048
12 norm of weights: 3.375585956129872
-----
Iteration Number: 2288
Loss: 29.653581126947206
12 norm of gradients: 0.4131936646708534
12 norm of weights: 3.3747490950134096
-----
Iteration Number: 2289
Loss: 29.639654779892453
12 norm of gradients: 0.4130075303678689
12 norm of weights: 3.373912656773079
-----
Iteration Number: 2290
Loss: 29.62573822175129
12 norm of gradients: 0.412821478617217
12 norm of weights: 3.3730766413085473
-----
Iteration Number: 2291
Loss: 29.611831446299675
12 norm of gradients: 0.41263550937856286
12 norm of weights: 3.372241048519505
-----
Iteration Number: 2292

```

```

Loss: 29.597934447318366
l2 norm of gradients: 0.41244962261157175
l2 norm of weights: 3.371405878305665
-----
Iteration Number: 2293
Loss: 29.584047218593327
l2 norm of gradients: 0.41226381827590913
l2 norm of weights: 3.3705711305667596
-----
Iteration Number: 2294
Loss: 29.570169753915163
l2 norm of gradients: 0.41207809633124104
l2 norm of weights: 3.369736805202547
-----
Iteration Number: 2295
Loss: 29.55630204707965
l2 norm of gradients: 0.411892456737234
l2 norm of weights: 3.3689029021128034
-----
Iteration Number: 2296
Loss: 29.54244409188687
l2 norm of gradients: 0.4117068994535562
l2 norm of weights: 3.3680694211973274
-----
Iteration Number: 2297
Loss: 29.528595882142298
l2 norm of gradients: 0.4115214244398768
l2 norm of weights: 3.3672363623559387
-----
Iteration Number: 2298
Loss: 29.514757411656184
l2 norm of gradients: 0.41133603165586696
l2 norm of weights: 3.3664037254884764
-----
Iteration Number: 2299
Loss: 29.500928674243
l2 norm of gradients: 0.41115072106119954
l2 norm of weights: 3.3655715104948003
-----
Iteration Number: 2300
Loss: 29.487109663723402
l2 norm of gradients: 0.41096549261554993
l2 norm of weights: 3.3647397172747904
-----
Iteration Number: 2301
Loss: 29.47330037392135
l2 norm of gradients: 0.41078034627859605
l2 norm of weights: 3.3639083457283463

```

```
-----
Iteration Number: 2302
Loss: 29.45950079866657
l2 norm of gradients: 0.41059528201001877
l2 norm of weights: 3.363077395755386
-----
Iteration Number: 2303
Loss: 29.4457109317934
l2 norm of gradients: 0.4104102997695019
l2 norm of weights: 3.362246867255848
-----
Iteration Number: 2304
Loss: 29.431930767140752
l2 norm of gradients: 0.4102253995167329
l2 norm of weights: 3.361416760129688
-----
Iteration Number: 2305
Loss: 29.418160298552483
l2 norm of gradients: 0.41004058121140297
l2 norm of weights: 3.3605870742768804
-----
Iteration Number: 2306
Loss: 29.40439951987753
l2 norm of gradients: 0.40985584481320725
l2 norm of weights: 3.3597578095974185
-----
Iteration Number: 2307
Loss: 29.39064842496913
l2 norm of gradients: 0.4096711902818451
l2 norm of weights: 3.358928965991313
-----
Iteration Number: 2308
Loss: 29.37690700768542
l2 norm of gradients: 0.4094866175770207
l2 norm of weights: 3.3581005433585918
-----
Iteration Number: 2309
Loss: 29.363175261889584
l2 norm of gradients: 0.4093021266584429
l2 norm of weights: 3.3572725415993
-----
Iteration Number: 2310
Loss: 29.349453181449167
l2 norm of gradients: 0.4091177174858259
l2 norm of weights: 3.3564449606135005
-----
Iteration Number: 2311
Loss: 29.335740760236874
```



```

12 norm of gradients: 0.40893339001888923
12 norm of weights: 3.355617800301272
-----
Iteration Number: 2312
Loss: 29.322037992129612
12 norm of gradients: 0.40874914421735814
12 norm of weights: 3.3547910605627105
-----
Iteration Number: 2313
Loss: 29.30834487100971
12 norm of gradients: 0.40856498004096387
12 norm of weights: 3.3539647412979265
-----
Iteration Number: 2314
Loss: 29.294661390763586
12 norm of gradients: 0.40838089744944395
12 norm of weights: 3.353138842407047
-----
Iteration Number: 2315
Loss: 29.280987545282553
12 norm of gradients: 0.4081968964025426
12 norm of weights: 3.3523133637902163
-----
Iteration Number: 2316
Loss: 29.267323328462957
12 norm of gradients: 0.40801297686001053
12 norm of weights: 3.3514883053475906
-----
Iteration Number: 2317
Loss: 29.2536687342053
12 norm of gradients: 0.407829138781606
12 norm of weights: 3.350663666979343
-----
Iteration Number: 2318
Loss: 29.24002375641536
12 norm of gradients: 0.4076453821270944
12 norm of weights: 3.3498394485856613
-----
Iteration Number: 2319
Loss: 29.22638838900319
12 norm of gradients: 0.40746170685624866
12 norm of weights: 3.349015650066747
-----
Iteration Number: 2320
Loss: 29.212762625883634
12 norm of gradients: 0.40727811292885
12 norm of weights: 3.3481922713228163
-----

```

Iteration Number: 2321
Loss: 29.199146460976046
l2 norm of gradients: 0.40709460030468736
l2 norm of weights: 3.3473693122540986

Iteration Number: 2322
Loss: 29.1855398882049
l2 norm of gradients: 0.4069111689435586
l2 norm of weights: 3.3465467727608367

Iteration Number: 2323
Loss: 29.17194290149875
l2 norm of gradients: 0.40672781880527026
l2 norm of weights: 3.345724652743287

Iteration Number: 2324
Loss: 29.158355494791326
l2 norm of gradients: 0.40654454984963734
l2 norm of weights: 3.3449029521017186

Iteration Number: 2325
Loss: 29.14477766202058
l2 norm of gradients: 0.4063613620364849
l2 norm of weights: 3.3440816707364136

Iteration Number: 2326
Loss: 29.131209397129204
l2 norm of gradients: 0.40617825532564705
l2 norm of weights: 3.3432608085476656

Iteration Number: 2327
Loss: 29.117650694064626
l2 norm of gradients: 0.4059952296769679
l2 norm of weights: 3.3424403654357815

Iteration Number: 2328
Loss: 29.10410154677867
l2 norm of gradients: 0.40581228505030154
l2 norm of weights: 3.3416203413010788

Iteration Number: 2329
Loss: 29.09056194922819
l2 norm of gradients: 0.4056294214055123
l2 norm of weights: 3.340800736043887

Iteration Number: 2330
Loss: 29.07703189537405
l2 norm of gradients: 0.4054466387024754

```

12 norm of weights: 3.339981549564547
-----
Iteration Number: 2331
Loss: 29.06351137918218
12 norm of gradients: 0.40526393690107665
12 norm of weights: 3.339162781763411
-----
Iteration Number: 2332
Loss: 29.050000394622934
12 norm of gradients: 0.4050813159612131
12 norm of weights: 3.3383444325408402
-----
Iteration Number: 2333
Loss: 29.036498935671123
12 norm of gradients: 0.4048987758427932
12 norm of weights: 3.3375265017972082
-----
Iteration Number: 2334
Loss: 29.023006996306393
12 norm of gradients: 0.4047163165057369
12 norm of weights: 3.336708989432898
-----
Iteration Number: 2335
Loss: 29.009524570512493
12 norm of gradients: 0.4045339379099763
12 norm of weights: 3.3358918953483014
-----
Iteration Number: 2336
Loss: 28.996051652278158
12 norm of gradients: 0.40435164001545537
12 norm of weights: 3.335075219443821
-----
Iteration Number: 2337
Loss: 28.982588235596715
12 norm of gradients: 0.40416942278213064
12 norm of weights: 3.334258961619869
-----
Iteration Number: 2338
Loss: 28.969134314465318
12 norm of gradients: 0.4039872861699713
12 norm of weights: 3.3334431217768654
-----
Iteration Number: 2339
Loss: 28.955689882886755
12 norm of gradients: 0.40380523013895936
12 norm of weights: 3.332627699815239
-----
Iteration Number: 2340

```

```
Loss: 28.942254934867236
12 norm of gradients: 0.4036232546490901
12 norm of weights: 3.3318126956354286
-----
Iteration Number: 2341
Loss: 28.92882946441823
12 norm of gradients: 0.40344135966037226
12 norm of weights: 3.3309981091378793
-----
Iteration Number: 2342
Loss: 28.915413465555503
12 norm of gradients: 0.40325954513282797
12 norm of weights: 3.3301839402230446
-----
Iteration Number: 2343
Loss: 28.902006932299162
12 norm of gradients: 0.4030778110264938
12 norm of weights: 3.329370188791387
-----
Iteration Number: 2344
Loss: 28.888609858673867
12 norm of gradients: 0.40289615730141987
12 norm of weights: 3.328556854743375
-----
Iteration Number: 2345
Loss: 28.875222238708886
12 norm of gradients: 0.4027145839176712
12 norm of weights: 3.327743937979484
-----
Iteration Number: 2346
Loss: 28.861844066437726
12 norm of gradients: 0.4025330908353273
12 norm of weights: 3.3269314384001976
-----
Iteration Number: 2347
Loss: 28.84847533589872
12 norm of gradients: 0.4023516780144825
12 norm of weights: 3.326119355906005
-----
Iteration Number: 2348
Loss: 28.835116041134448
12 norm of gradients: 0.40217034541524643
12 norm of weights: 3.325307690397401
-----
Iteration Number: 2349
Loss: 28.8217661761917
12 norm of gradients: 0.40198909299774405
12 norm of weights: 3.324496441774888
```

```

-----
Iteration Number: 2350
Loss: 28.80842573512217
l2 norm of gradients: 0.40180792072211585
l2 norm of weights: 3.3236856099389733
-----
Iteration Number: 2351
Loss: 28.795094711981637
l2 norm of gradients: 0.4016268285485185
l2 norm of weights: 3.32287519479017
-----
Iteration Number: 2352
Loss: 28.78177310083058
l2 norm of gradients: 0.40144581643712446
l2 norm of weights: 3.322065196228996
-----
Iteration Number: 2353
Loss: 28.768460895733618
l2 norm of gradients: 0.4012648843481229
l2 norm of weights: 3.321255614155975
-----
Iteration Number: 2354
Loss: 28.755158090759828
l2 norm of gradients: 0.4010840322417191
l2 norm of weights: 3.320446448471633
-----
Iteration Number: 2355
Loss: 28.741864679982996
l2 norm of gradients: 0.40090326007813576
l2 norm of weights: 3.319637699076505
-----
Iteration Number: 2356
Loss: 28.728580657480926
l2 norm of gradients: 0.4007225678176122
l2 norm of weights: 3.3188293658711254
-----
Iteration Number: 2357
Loss: 28.715306017335983
l2 norm of gradients: 0.4005419554204052
l2 norm of weights: 3.318021448756035
-----
Iteration Number: 2358
Loss: 28.702040753635004
l2 norm of gradients: 0.4003614228467892
l2 norm of weights: 3.3172139476317777
-----
Iteration Number: 2359
Loss: 28.688784860468985

```

```

12 norm of gradients: 0.4001809700570562
12 norm of weights: 3.3164068623989014
-----
Iteration Number: 2360
Loss: 28.67553833193339
12 norm of gradients: 0.4000005970115164
12 norm of weights: 3.315600192957956
-----
Iteration Number: 2361
Loss: 28.662301162128077
12 norm of gradients: 0.3998203036704982
12 norm of weights: 3.3147939392094954
-----
Iteration Number: 2362
Loss: 28.649073345157298
12 norm of gradients: 0.39964008999434847
12 norm of weights: 3.313988101054076
-----
Iteration Number: 2363
Loss: 28.635854875129425
12 norm of gradients: 0.39945995594343264
12 norm of weights: 3.3131826783922547
-----
Iteration Number: 2364
Loss: 28.622645746157517
12 norm of gradients: 0.39927990147813525
12 norm of weights: 3.312377671124593
-----
Iteration Number: 2365
Loss: 28.609445952358612
12 norm of gradients: 0.39909992655886006
12 norm of weights: 3.3115730791516533
-----
Iteration Number: 2366
Loss: 28.59625548785416
12 norm of gradients: 0.39892003114602986
12 norm of weights: 3.310768902373999
-----
Iteration Number: 2367
Loss: 28.583074346770182
12 norm of gradients: 0.39874021520008757
12 norm of weights: 3.3099651406921957
-----
Iteration Number: 2368
Loss: 28.569902523236717
12 norm of gradients: 0.3985604786814954
12 norm of weights: 3.309161794006809
-----

```

Iteration Number: 2369
Loss: 28.556740011388122
l2 norm of gradients: 0.3983808215507362
l2 norm of weights: 3.308358862218407

Iteration Number: 2370
Loss: 28.543586805363155
l2 norm of gradients: 0.39820124376831245
l2 norm of weights: 3.3075563452275567

Iteration Number: 2371
Loss: 28.530442899304823
l2 norm of gradients: 0.3980217452947477
l2 norm of weights: 3.3067542429348262

Iteration Number: 2372
Loss: 28.5173082873604
l2 norm of gradients: 0.39784232609058606
l2 norm of weights: 3.3059525552407836

Iteration Number: 2373
Loss: 28.504182963681416
l2 norm of gradients: 0.3976629861163924
l2 norm of weights: 3.305151282045997

Iteration Number: 2374
Loss: 28.491066922423762
l2 norm of gradients: 0.3974837253327529
l2 norm of weights: 3.3043504232510337

Iteration Number: 2375
Loss: 28.477960157747447
l2 norm of gradients: 0.39730454370027524
l2 norm of weights: 3.3035499787564606

Iteration Number: 2376
Loss: 28.464862663816735
l2 norm of gradients: 0.3971254411795884
l2 norm of weights: 3.3027499484628433

Iteration Number: 2377
Loss: 28.45177443480029
l2 norm of gradients: 0.3969464177313434
l2 norm of weights: 3.301950332270746

Iteration Number: 2378
Loss: 28.43869546487074
l2 norm of gradients: 0.39676747331621337

```

12 norm of weights: 3.3011511300807337
-----
Iteration Number: 2379
Loss: 28.425625748204997
12 norm of gradients: 0.39658860789489353
12 norm of weights: 3.3003523417933667
-----
Iteration Number: 2380
Loss: 28.41256527898467
12 norm of gradients: 0.39640982142810144
12 norm of weights: 3.2995539673092042
-----
Iteration Number: 2381
Loss: 28.399514051394895
12 norm of gradients: 0.3962311138765776
12 norm of weights: 3.2987560065288046
-----
Iteration Number: 2382
Loss: 28.38647205962532
12 norm of gradients: 0.3960524852010852
12 norm of weights: 3.297958459352723
-----
Iteration Number: 2383
Loss: 28.373439297869876
12 norm of gradients: 0.39587393536241067
12 norm of weights: 3.2971613256815115
-----
Iteration Number: 2384
Loss: 28.360415760326532
12 norm of gradients: 0.3956954643213636
12 norm of weights: 3.2963646054157203
-----
Iteration Number: 2385
Loss: 28.347401441197484
12 norm of gradients: 0.39551707203877695
12 norm of weights: 3.295568298455895
-----
Iteration Number: 2386
Loss: 28.33439633468919
12 norm of gradients: 0.395338758475508
12 norm of weights: 3.294772404702579
-----
Iteration Number: 2387
Loss: 28.321400435012
12 norm of gradients: 0.39516052359243725
12 norm of weights: 3.2939769240563117
-----
Iteration Number: 2388

```


Loss: 28.308413736380846
l2 norm of gradients: 0.39498236735046965
l2 norm of weights: 3.2931818564176294

Iteration Number: 2389
Loss: 28.29543623301434
l2 norm of gradients: 0.3948042897105345
l2 norm of weights: 3.292387201687063

Iteration Number: 2390
Loss: 28.2824679191357
l2 norm of gradients: 0.3946262906335855
l2 norm of weights: 3.29159295976514

Iteration Number: 2391
Loss: 28.269508788971898
l2 norm of gradients: 0.39444837008060124
l2 norm of weights: 3.2907991305523825

Iteration Number: 2392
Loss: 28.2565588367543
l2 norm of gradients: 0.39427052801258505
l2 norm of weights: 3.2900057139493093

Iteration Number: 2393
Loss: 28.243618056718148
l2 norm of gradients: 0.3940927643905658
l2 norm of weights: 3.2892127098564328

Iteration Number: 2394
Loss: 28.23068644310317
l2 norm of gradients: 0.39391507917559704
l2 norm of weights: 3.2884201181742605

Iteration Number: 2395
Loss: 28.217763990152797
l2 norm of gradients: 0.39373747232875844
l2 norm of weights: 3.2876279388032943

Iteration Number: 2396
Loss: 28.204850692114846
l2 norm of gradients: 0.3935599438111551
l2 norm of weights: 3.2868361716440315

Iteration Number: 2397
Loss: 28.1919465432411
l2 norm of gradients: 0.39338249358391814
l2 norm of weights: 3.2860448165969616

```

-----
Iteration Number: 2398
Loss: 28.17905153778738
l2 norm of gradients: 0.3932051216082049
l2 norm of weights: 3.285253873562569
-----
Iteration Number: 2399
Loss: 28.16616567001376
l2 norm of gradients: 0.3930278278451988
l2 norm of weights: 3.284463342441332
-----
Iteration Number: 2400
Loss: 28.15328893418423
l2 norm of gradients: 0.3928506122561099
l2 norm of weights: 3.2836732231337207
-----
Iteration Number: 2401
Loss: 28.140421324567015
l2 norm of gradients: 0.3926734748021749
l2 norm of weights: 3.2828835155402003
-----
Iteration Number: 2402
Loss: 28.127562835434233
l2 norm of gradients: 0.3924964154446578
l2 norm of weights: 3.2820942195612273
-----
Iteration Number: 2403
Loss: 28.114713461061996
l2 norm of gradients: 0.3923194341448488
l2 norm of weights: 3.281305335097252
-----
Iteration Number: 2404
Loss: 28.101873195730747
l2 norm of gradients: 0.3921425308640664
l2 norm of weights: 3.2805168620487164
-----
Iteration Number: 2405
Loss: 28.089042033724766
l2 norm of gradients: 0.39196570556365573
l2 norm of weights: 3.279728800316055
-----
Iteration Number: 2406
Loss: 28.076219969332353
l2 norm of gradients: 0.39178895820499027
l2 norm of weights: 3.278941149799694
-----
Iteration Number: 2407
Loss: 28.063406996845913

```

```

12 norm of gradients: 0.3916122887494706
12 norm of weights: 3.2781539104000523
-----
Iteration Number: 2408
Loss: 28.050603110561756
12 norm of gradients: 0.39143569715852594
12 norm of weights: 3.277367082017539
-----
Iteration Number: 2409
Loss: 28.037808304780434
12 norm of gradients: 0.39125918339361343
12 norm of weights: 3.2765806645525553
-----
Iteration Number: 2410
Loss: 28.025022573806254
12 norm of gradients: 0.39108274741621857
12 norm of weights: 3.2757946579054935
-----
Iteration Number: 2411
Loss: 28.012245911947623
12 norm of gradients: 0.39090638918785564
12 norm of weights: 3.2750090619767365
-----
Iteration Number: 2412
Loss: 27.999478313517006
12 norm of gradients: 0.3907301086700674
12 norm of weights: 3.274223876666658
-----
Iteration Number: 2413
Loss: 27.986719772830646
12 norm of gradients: 0.39055390582442584
12 norm of weights: 3.273439101875622
-----
Iteration Number: 2414
Loss: 27.973970284208928
12 norm of gradients: 0.39037778061253187
12 norm of weights: 3.272654737503984
-----
Iteration Number: 2415
Loss: 27.961229841976248
12 norm of gradients: 0.3902017329960156
12 norm of weights: 3.271870783452086
-----
Iteration Number: 2416
Loss: 27.948498440460806
12 norm of gradients: 0.3900257629365371
12 norm of weights: 3.271087239620264
-----

```

Iteration Number: 2417
Loss: 27.93577607399487
l2 norm of gradients: 0.38984987039578556
l2 norm of weights: 3.270304105908842

Iteration Number: 2418
Loss: 27.92306273691464
l2 norm of gradients: 0.38967405533547994
l2 norm of weights: 3.269521382218131

Iteration Number: 2419
Loss: 27.9103584235602
l2 norm of gradients: 0.38949831771736976
l2 norm of weights: 3.268739068448435

Iteration Number: 2420
Loss: 27.897663128275695
l2 norm of gradients: 0.3893226575032343
l2 norm of weights: 3.2679571645000447

Iteration Number: 2421
Loss: 27.884976845409074
l2 norm of gradients: 0.38914707465488324
l2 norm of weights: 3.2671756702732395

Iteration Number: 2422
Loss: 27.872299569312162
l2 norm of gradients: 0.3889715691341569
l2 norm of weights: 3.2663945856682877

Iteration Number: 2423
Loss: 27.85963129434092
l2 norm of gradients: 0.38879614090292625
l2 norm of weights: 3.2656139105854467

Iteration Number: 2424
Loss: 27.846972014855062
l2 norm of gradients: 0.3886207899230929
l2 norm of weights: 3.264833644924961

Iteration Number: 2425
Loss: 27.834321725218068
l2 norm of gradients: 0.3884455161565899
l2 norm of weights: 3.2640537885870624

Iteration Number: 2426
Loss: 27.821680419797698
l2 norm of gradients: 0.3882703195653812

```

12 norm of weights: 3.2632743414719716
-----
Iteration Number: 2427
Loss: 27.809048092965234
12 norm of gradients: 0.3880952001114621
12 norm of weights: 3.2624953034798962
-----
Iteration Number: 2428
Loss: 27.796424739096015
12 norm of gradients: 0.38792015775685973
12 norm of weights: 3.2617166745110313
-----
Iteration Number: 2429
Loss: 27.783810352569198
12 norm of gradients: 0.38774519246363254
12 norm of weights: 3.260938454465559
-----
Iteration Number: 2430
Loss: 27.77120492776777
12 norm of gradients: 0.3875703041938712
12 norm of weights: 3.2601606432436476
-----
Iteration Number: 2431
Loss: 27.758608459078733
12 norm of gradients: 0.38739549290969805
12 norm of weights: 3.259383240745453
-----
Iteration Number: 2432
Loss: 27.746020940892862
12 norm of gradients: 0.38722075857326793
12 norm of weights: 3.2586062468711168
-----
Iteration Number: 2433
Loss: 27.733442367604614
12 norm of gradients: 0.38704610114676785
12 norm of weights: 3.2578296615207676
-----
Iteration Number: 2434
Loss: 27.7208727336126
12 norm of gradients: 0.3868715205924175
12 norm of weights: 3.2570534845945183
-----
Iteration Number: 2435
Loss: 27.708312033319054
12 norm of gradients: 0.38669701687246893
12 norm of weights: 3.25627771599247
-----
Iteration Number: 2436

```

```

Loss: 27.695760261130104
12 norm of gradients: 0.3865225899492075
12 norm of weights: 3.2555023556147082
-----
Iteration Number: 2437
Loss: 27.68321741145553
12 norm of gradients: 0.3863482397849512
12 norm of weights: 3.254727403361303
-----
Iteration Number: 2438
Loss: 27.670683478709297
12 norm of gradients: 0.38617396634205114
12 norm of weights: 3.253952859132312
-----
Iteration Number: 2439
Loss: 27.658158457308847
12 norm of gradients: 0.385999769582892
12 norm of weights: 3.2531787228277746
-----
Iteration Number: 2440
Loss: 27.64564234167547
12 norm of gradients: 0.38582564946989173
12 norm of weights: 3.2524049943477182
-----
Iteration Number: 2441
Loss: 27.633135126234528
12 norm of gradients: 0.38565160596550196
12 norm of weights: 3.251631673592153
-----
Iteration Number: 2442
Loss: 27.62063680541479
12 norm of gradients: 0.38547763903220816
12 norm of weights: 3.2508587604610746
-----
Iteration Number: 2443
Loss: 27.608147373648958
12 norm of gradients: 0.3853037486325294
12 norm of weights: 3.2500862548544616
-----
Iteration Number: 2444
Loss: 27.59566682537359
12 norm of gradients: 0.3851299347290193
12 norm of weights: 3.249314156672278
-----
Iteration Number: 2445
Loss: 27.583195155029042
12 norm of gradients: 0.38495619728426544
12 norm of weights: 3.2485424658144706

```

```

-----
Iteration Number: 2446
Loss: 27.570732357059246
l2 norm of gradients: 0.38478253626088976
l2 norm of weights: 3.247771182180971
-----
Iteration Number: 2447
Loss: 27.558278425912025
l2 norm of gradients: 0.3846089516215487
l2 norm of weights: 3.247000305671693
-----
Iteration Number: 2448
Loss: 27.545833356038912
l2 norm of gradients: 0.38443544332893353
l2 norm of weights: 3.2462298361865343
-----
Iteration Number: 2449
Loss: 27.533397141895104
l2 norm of gradients: 0.38426201134577026
l2 norm of weights: 3.2454597736253765
-----
Iteration Number: 2450
Loss: 27.520969777939776
l2 norm of gradients: 0.38408865563481975
l2 norm of weights: 3.2446901178880823
-----
Iteration Number: 2451
Loss: 27.508551258635674
l2 norm of gradients: 0.38391537615887816
l2 norm of weights: 3.2439208688744987
-----
Iteration Number: 2452
Loss: 27.496141578449084
l2 norm of gradients: 0.3837421728807769
l2 norm of weights: 3.243152026484454
-----
Iteration Number: 2453
Loss: 27.483740731850407
l2 norm of gradients: 0.38356904576338247
l2 norm of weights: 3.24238359061776
-----
Iteration Number: 2454
Loss: 27.471348713313443
l2 norm of gradients: 0.38339599476959746
l2 norm of weights: 3.24161556117421
-----
Iteration Number: 2455
Loss: 27.458965517315907

```

```

12 norm of gradients: 0.3832230198623595
12 norm of weights: 3.240847938053579
-----
Iteration Number: 2456
Loss: 27.446591138338942
12 norm of gradients: 0.38305012100464253
12 norm of weights: 3.2400807211556244
-----
Iteration Number: 2457
Loss: 27.43422557086768
12 norm of gradients: 0.3828772981594563
12 norm of weights: 3.2393139103800856
-----
Iteration Number: 2458
Loss: 27.42186880939086
12 norm of gradients: 0.38270455128984665
12 norm of weights: 3.2385475056266824
-----
Iteration Number: 2459
Loss: 27.40952084840076
12 norm of gradients: 0.3825318803588957
12 norm of weights: 3.2377815067951157
-----
Iteration Number: 2460
Loss: 27.397181682393388
12 norm of gradients: 0.3823592853297218
12 norm of weights: 3.237015913785069
-----
Iteration Number: 2461
Loss: 27.38485130586856
12 norm of gradients: 0.38218676616548014
12 norm of weights: 3.236250726496204
-----
Iteration Number: 2462
Loss: 27.372529713329644
12 norm of gradients: 0.38201432282936215
12 norm of weights: 3.235485944828167
-----
Iteration Number: 2463
Loss: 27.36021689928363
12 norm of gradients: 0.3818419552845964
12 norm of weights: 3.2347215686805817
-----
Iteration Number: 2464
Loss: 27.34791285824122
12 norm of gradients: 0.38166966349444814
12 norm of weights: 3.233957597953052
-----

```


Iteration Number: 2465
Loss: 27.335617584716758
l2 norm of gradients: 0.38149744742221986
l2 norm of weights: 3.233194032545165

Iteration Number: 2466
Loss: 27.323331073228083
l2 norm of gradients: 0.3813253070312511
l2 norm of weights: 3.2324308723564843

Iteration Number: 2467
Loss: 27.311053318297024
l2 norm of gradients: 0.3811532422849188
l2 norm of weights: 3.231668117286555

Iteration Number: 2468
Loss: 27.29878431444867
l2 norm of gradients: 0.3809812531466373
l2 norm of weights: 3.2309057672349026

Iteration Number: 2469
Loss: 27.28652405621178
l2 norm of gradients: 0.38080933957985835
l2 norm of weights: 3.2301438221010303

Iteration Number: 2470
Loss: 27.274272538118915
l2 norm of gradients: 0.38063750154807185
l2 norm of weights: 3.2293822817844218

Iteration Number: 2471
Loss: 27.262029754706163
l2 norm of gradients: 0.3804657390148051
l2 norm of weights: 3.2286211461845395

Iteration Number: 2472
Loss: 27.249795700513047
l2 norm of gradients: 0.3802940519436236
l2 norm of weights: 3.227860415200825

Iteration Number: 2473
Loss: 27.237570370083017
l2 norm of gradients: 0.38012244029813086
l2 norm of weights: 3.227100088732699

Iteration Number: 2474
Loss: 27.22535375796273
l2 norm of gradients: 0.37995090404196846

```

12 norm of weights: 3.2263401666795595
-----
Iteration Number: 2475
Loss: 27.213145858702664
12 norm of gradients: 0.37977944313881656
12 norm of weights: 3.225580648940784
-----
Iteration Number: 2476
Loss: 27.20094666856836
12 norm of gradients: 0.37960805755239374
12 norm of weights: 3.2248215354157286
-----
Iteration Number: 2477
Loss: 27.188756176982913
12 norm of gradients: 0.379436747246457
12 norm of weights: 3.2240628260037267
-----
Iteration Number: 2478
Loss: 27.176574383641896
12 norm of gradients: 0.37926551218480203
12 norm of weights: 3.2233045206040893
-----
Iteration Number: 2479
Loss: 27.16440128139847
12 norm of gradients: 0.37909435233126365
12 norm of weights: 3.2225466191161076
-----
Iteration Number: 2480
Loss: 27.152236864820978
12 norm of gradients: 0.37892326764971535
12 norm of weights: 3.2217891214390466
-----
Iteration Number: 2481
Loss: 27.140081128481174
12 norm of gradients: 0.37875225810406965
12 norm of weights: 3.2210320274721527
-----
Iteration Number: 2482
Loss: 27.127934066954374
12 norm of gradients: 0.37858132365827835
12 norm of weights: 3.2202753371146464
-----
Iteration Number: 2483
Loss: 27.115795674819417
12 norm of gradients: 0.3784104642763328
12 norm of weights: 3.2195190502657276
-----
Iteration Number: 2484

```

```

Loss: 27.103665946658808
l2 norm of gradients: 0.3782396799222632
l2 norm of weights: 3.218763166824572
-----
Iteration Number: 2485
Loss: 27.09154487705843
l2 norm of gradients: 0.37806897056013966
l2 norm of weights: 3.2180076866903318
-----
Iteration Number: 2486
Loss: 27.07943246060764
l2 norm of gradients: 0.37789833615407203
l2 norm of weights: 3.217252609762137
-----
Iteration Number: 2487
Loss: 27.06732869189943
l2 norm of gradients: 0.3777277766682096
l2 norm of weights: 3.2164979359390933
-----
Iteration Number: 2488
Loss: 27.0552335655303
l2 norm of gradients: 0.37755729206674177
l2 norm of weights: 3.2157436651202835
-----
Iteration Number: 2489
Loss: 27.043147076100226
l2 norm of gradients: 0.37738688231389783
l2 norm of weights: 3.2149897972047654
-----
Iteration Number: 2490
Loss: 27.031069218212505
l2 norm of gradients: 0.377216547373947
l2 norm of weights: 3.2142363320915734
-----
Iteration Number: 2491
Loss: 27.018999986474256
l2 norm of gradients: 0.37704628721119904
l2 norm of weights: 3.2134832696797186
-----
Iteration Number: 2492
Loss: 27.006939375495755
l2 norm of gradients: 0.3768761017900038
l2 norm of weights: 3.2127306098681867
-----
Iteration Number: 2493
Loss: 26.994887379890944
l2 norm of gradients: 0.3767059910747518
l2 norm of weights: 3.2119783525559393

```

```

-----
Iteration Number: 2494
Loss: 26.982843994277253
l2 norm of gradients: 0.3765359550298737
l2 norm of weights: 3.2112264976419134
-----
Iteration Number: 2495
Loss: 26.970809213275466
l2 norm of gradients: 0.376365993619841
l2 norm of weights: 3.2104750450250217
-----
Iteration Number: 2496
Loss: 26.95878303150983
l2 norm of gradients: 0.3761961068091661
l2 norm of weights: 3.2097239946041505
-----
Iteration Number: 2497
Loss: 26.94676544360822
l2 norm of gradients: 0.3760262945624021
l2 norm of weights: 3.2089733462781633
-----
Iteration Number: 2498
Loss: 26.93475644420168
l2 norm of gradients: 0.37585655684414293
l2 norm of weights: 3.208223099945897
-----
Iteration Number: 2499
Loss: 26.922756027924798
l2 norm of gradients: 0.37568689361902385
l2 norm of weights: 3.2074732555061636
-----
Iteration Number: 2500
Loss: 26.91076418941591
l2 norm of gradients: 0.3755173048517211
l2 norm of weights: 3.206723812857749
-----
Iteration Number: 2501
Loss: 26.89878092331622
l2 norm of gradients: 0.3753477905069523
l2 norm of weights: 3.2059747718994145
-----
Iteration Number: 2502
Loss: 26.88680622427082
l2 norm of gradients: 0.3751783505494762
l2 norm of weights: 3.2052261325298947
-----
Iteration Number: 2503
Loss: 26.87484008692804

```

```

12 norm of gradients: 0.3750089849440933
12 norm of weights: 3.2044778946478987
-----
Iteration Number: 2504
Loss: 26.86288250593952
12 norm of gradients: 0.3748396936556458
12 norm of weights: 3.2037300581521095
-----
Iteration Number: 2505
Loss: 26.850933475960638
12 norm of gradients: 0.3746704766490168
12 norm of weights: 3.202982622941184
-----
Iteration Number: 2506
Loss: 26.838992991649782
12 norm of gradients: 0.3745013338891322
12 norm of weights: 3.2022355889137524
-----
Iteration Number: 2507
Loss: 26.82706104766902
12 norm of gradients: 0.374332265340959
12 norm of weights: 3.2014889559684176
-----
Iteration Number: 2508
Loss: 26.815137638683712
12 norm of gradients: 0.3741632709695065
12 norm of weights: 3.200742724003759
-----
Iteration Number: 2509
Loss: 26.803222759362473
12 norm of gradients: 0.373994350739826
12 norm of weights: 3.199996892918326
-----
Iteration Number: 2510
Loss: 26.791316404377586
12 norm of gradients: 0.373825504617011
12 norm of weights: 3.1992514626106416
-----
Iteration Number: 2511
Loss: 26.779418568404516
12 norm of gradients: 0.3736567325661969
12 norm of weights: 3.1985064329792032
-----
Iteration Number: 2512
Loss: 26.76752924612207
12 norm of gradients: 0.37348803455256196
12 norm of weights: 3.1977618039224796
-----

```

Iteration Number: 2513
Loss: 26.75564843221247
l2 norm of gradients: 0.37331941054132656
l2 norm of weights: 3.1970175753389123

Iteration Number: 2514
Loss: 26.743776121361442
l2 norm of gradients: 0.37315086049775364
l2 norm of weights: 3.196273747126917

Iteration Number: 2515
Loss: 26.731912308257808
l2 norm of gradients: 0.3729823843871488
l2 norm of weights: 3.1955303191848796

Iteration Number: 2516
Loss: 26.720056987593875
l2 norm of gradients: 0.37281398217486034
l2 norm of weights: 3.19478729141116

Iteration Number: 2517
Loss: 26.70821015406534
l2 norm of gradients: 0.3726456538262794
l2 norm of weights: 3.194044663704089

Iteration Number: 2518
Loss: 26.69637180237112
l2 norm of gradients: 0.3724773993068399
l2 norm of weights: 3.1933024359619697

Iteration Number: 2519
Loss: 26.684541927213548
l2 norm of gradients: 0.37230921858201865
l2 norm of weights: 3.192560608083077

Iteration Number: 2520
Loss: 26.67272052329823
l2 norm of gradients: 0.37214111161733576
l2 norm of weights: 3.1918191799656594

Iteration Number: 2521
Loss: 26.660907585334154
l2 norm of gradients: 0.3719730783783546
l2 norm of weights: 3.1910781515079334

Iteration Number: 2522
Loss: 26.649103108033522
l2 norm of gradients: 0.3718051188306812

```

12 norm of weights: 3.1903375226080897
-----
Iteration Number: 2523
Loss: 26.637307086111978
12 norm of gradients: 0.37163723293996537
12 norm of weights: 3.1895972931642893
-----
Iteration Number: 2524
Loss: 26.62551951428837
12 norm of gradients: 0.3714694206719001
12 norm of weights: 3.18857463074665
-----
Iteration Number: 2525
Loss: 26.613740387284928
12 norm of gradients: 0.37130168199222213
12 norm of weights: 3.18811803223732
-----
Iteration Number: 2526
Loss: 26.60196969982706
12 norm of gradients: 0.37113401686671155
12 norm of weights: 3.187379000550329
-----
Iteration Number: 2527
Loss: 26.590207446643472
12 norm of gradients: 0.3709664252611921
12 norm of weights: 3.186640367911737
-----
Iteration Number: 2528
Loss: 26.57845362246641
12 norm of gradients: 0.37079890714153135
12 norm of weights: 3.1859021342195604
-----
Iteration Number: 2529
Loss: 26.566708222030975
12 norm of gradients: 0.37063146247364065
12 norm of weights: 3.1851642993717855
-----
Iteration Number: 2530
Loss: 26.55497124007591
12 norm of gradients: 0.37046409122347507
12 norm of weights: 3.1844268632663693
-----
Iteration Number: 2531
Loss: 26.543242671342995
12 norm of gradients: 0.37029679335703386
12 norm of weights: 3.183689825801239
-----
Iteration Number: 2532

```

```

Loss: 26.53152251057747
l2 norm of gradients: 0.37012956884036025
l2 norm of weights: 3.1829531868742924
-----
Iteration Number: 2533
Loss: 26.51981075252758
l2 norm of gradients: 0.3699624176395416
l2 norm of weights: 3.1822169463833974
-----
Iteration Number: 2534
Loss: 26.508107391945092
l2 norm of gradients: 0.36979533972070944
l2 norm of weights: 3.181481104226391
-----
Iteration Number: 2535
Loss: 26.49641242358483
l2 norm of gradients: 0.36962833505003956
l2 norm of weights: 3.1807456603010813
-----
Iteration Number: 2536
Loss: 26.48472584220484
l2 norm of gradients: 0.3694614035937522
l2 norm of weights: 3.1800106145052447
-----
Iteration Number: 2537
Loss: 26.47304764256663
l2 norm of gradients: 0.36929454531811184
l2 norm of weights: 3.179275966736628
-----
Iteration Number: 2538
Loss: 26.461377819434738
l2 norm of gradients: 0.36912776018942767
l2 norm of weights: 3.178541716892948
-----
Iteration Number: 2539
Loss: 26.44971636757697
l2 norm of gradients: 0.3689610481740533
l2 norm of weights: 3.17780786487189
-----
Iteration Number: 2540
Loss: 26.43806328176432
l2 norm of gradients: 0.3687944092383869
l2 norm of weights: 3.1770744105711093
-----
Iteration Number: 2541
Loss: 26.426418556771175
l2 norm of gradients: 0.3686278433488717
l2 norm of weights: 3.1763413538882292

```



```
-----  
Iteration Number: 2542  
Loss: 26.41478218737489  
l2 norm of gradients: 0.36846135047199535  
l2 norm of weights: 3.1756086947208426  
-----  
Iteration Number: 2543  
Loss: 26.403154168356174  
l2 norm of gradients: 0.36829493057429064  
l2 norm of weights: 3.174876432966512  
-----  
Iteration Number: 2544  
Loss: 26.391534494498952  
l2 norm of gradients: 0.36812858362233497  
l2 norm of weights: 3.174144568522768  
-----  
Iteration Number: 2545  
Loss: 26.379923160590234  
l2 norm of gradients: 0.36796230958275095  
l2 norm of weights: 3.17341310128711  
-----  
Iteration Number: 2546  
Loss: 26.3683201614203  
l2 norm of gradients: 0.3677961084222063  
l2 norm of weights: 3.172682031157006  
-----  
Iteration Number: 2547  
Loss: 26.35672549178257  
l2 norm of gradients: 0.3676299801074136  
l2 norm of weights: 3.1719513580298915  
-----  
Iteration Number: 2548  
Loss: 26.34513914647374  
l2 norm of gradients: 0.36746392460513094  
l2 norm of weights: 3.1712210818031723  
-----  
Iteration Number: 2549  
Loss: 26.333561120293634  
l2 norm of gradients: 0.36729794188216147  
l2 norm of weights: 3.1704912023742216  
-----  
Iteration Number: 2550  
Loss: 26.321991408045122  
l2 norm of gradients: 0.36713203190535354  
l2 norm of weights: 3.1697617196403796  
-----  
Iteration Number: 2551  
Loss: 26.310430004534403
```

```

12 norm of gradients: 0.3669661946416013
12 norm of weights: 3.169032633498956
-----
Iteration Number: 2552
Loss: 26.298876904570815
12 norm of gradients: 0.3668004300578439
12 norm of weights: 3.1683039438472274
-----
Iteration Number: 2553
Loss: 26.287332102966786
12 norm of gradients: 0.3666347381210663
12 norm of weights: 3.1675756505824393
-----
Iteration Number: 2554
Loss: 26.27579559453794
12 norm of gradients: 0.3664691187982989
12 norm of weights: 3.1668477536018034
-----
Iteration Number: 2555
Loss: 26.264267374102992
12 norm of gradients: 0.3663035720566178
12 norm of weights: 3.1661202528025005
-----
Iteration Number: 2556
Loss: 26.25274743648392
12 norm of gradients: 0.36613809786314466
12 norm of weights: 3.1653931480816784
-----
Iteration Number: 2557
Loss: 26.24123577650572
12 norm of gradients: 0.365972696185047
12 norm of weights: 3.1646664393364516
-----
Iteration Number: 2558
Loss: 26.229732388996496
12 norm of gradients: 0.3658073669895385
12 norm of weights: 3.1639401264639027
-----
Iteration Number: 2559
Loss: 26.218237268787693
12 norm of gradients: 0.3656421102438778
12 norm of weights: 3.1632142093610813
-----
Iteration Number: 2560
Loss: 26.206750410713596
12 norm of gradients: 0.3654769259153704
12 norm of weights: 3.1624886879250034
-----

```

Iteration Number: 2561
Loss: 26.195271809611782
l2 norm of gradients: 0.36531181397136725
l2 norm of weights: 3.1617635620526525

Iteration Number: 2562
Loss: 26.183801460322854
l2 norm of gradients: 0.36514677437926557
l2 norm of weights: 3.16103883164098

Iteration Number: 2563
Loss: 26.172339357690714
l2 norm of gradients: 0.3649818071065087
l2 norm of weights: 3.160314496586902

Iteration Number: 2564
Loss: 26.160885496562106
l2 norm of gradients: 0.36481691212058587
l2 norm of weights: 3.159590556787303

Iteration Number: 2565
Loss: 26.14943987178693
l2 norm of gradients: 0.3646520893890329
l2 norm of weights: 3.158867012139033

Iteration Number: 2566
Loss: 26.138002478218343
l2 norm of gradients: 0.36448733887943136
l2 norm of weights: 3.15814386253891

Iteration Number: 2567
Loss: 26.12657331071246
l2 norm of gradients: 0.3643226605594097
l2 norm of weights: 3.157421107883716

Iteration Number: 2568
Loss: 26.115152364128516
l2 norm of gradients: 0.3641580543966424
l2 norm of weights: 3.1566987480702013

Iteration Number: 2569
Loss: 26.10373963332876
l2 norm of gradients: 0.3639935203588504
l2 norm of weights: 3.1559767829950816

Iteration Number: 2570
Loss: 26.09233511317863
l2 norm of gradients: 0.36382905841380114

```

12 norm of weights: 3.155255212555039
-----
Iteration Number: 2571
Loss: 26.080938798546477
12 norm of gradients: 0.36366466852930857
12 norm of weights: 3.1545340366467207
-----
Iteration Number: 2572
Loss: 26.069550684303934
12 norm of gradients: 0.3635003506732333
12 norm of weights: 3.153813255166741
-----
Iteration Number: 2573
Loss: 26.058170765325507
12 norm of gradients: 0.3633361048134824
12 norm of weights: 3.15309286801168
-----
Iteration Number: 2574
Loss: 26.04679903648882
12 norm of gradients: 0.36317193091800953
12 norm of weights: 3.1523728750780826
-----
Iteration Number: 2575
Loss: 26.035435492674512
12 norm of gradients: 0.3630078289548155
12 norm of weights: 3.1516532762624596
-----
Iteration Number: 2576
Loss: 26.024080128766396
12 norm of gradients: 0.3628437988919472
12 norm of weights: 3.150934071461288
-----
Iteration Number: 2577
Loss: 26.01273293965123
12 norm of gradients: 0.362679840697499
12 norm of weights: 3.1502152605710094
-----
Iteration Number: 2578
Loss: 26.001393920218717
12 norm of gradients: 0.3625159543396117
12 norm of weights: 3.1494968434880315
-----
Iteration Number: 2579
Loss: 25.99006306536181
12 norm of gradients: 0.36235213978647307
12 norm of weights: 3.1487788201087272
-----
Iteration Number: 2580

```

Loss: 25.97874036997633
l2 norm of gradients: 0.362188397006318
l2 norm of weights: 3.1480611903294347

Iteration Number: 2581
Loss: 25.96742582896118
l2 norm of gradients: 0.36202472596742813
l2 norm of weights: 3.147343954046456

Iteration Number: 2582
Loss: 25.956119437218202
l2 norm of gradients: 0.36186112663813214
l2 norm of weights: 3.146627111156059

Iteration Number: 2583
Loss: 25.94482118965244
l2 norm of gradients: 0.3616975989868059
l2 norm of weights: 3.145910661554478

Iteration Number: 2584
Loss: 25.933531081171747
l2 norm of gradients: 0.3615341429818724
l2 norm of weights: 3.1451946051379096

Iteration Number: 2585
Loss: 25.922249106687048
l2 norm of gradients: 0.3613707585918017
l2 norm of weights: 3.1444789418025167

Iteration Number: 2586
Loss: 25.9109752611123
l2 norm of gradients: 0.361207445785111
l2 norm of weights: 3.143763671444427

Iteration Number: 2587
Loss: 25.899709539364522
l2 norm of gradients: 0.3610442045303648
l2 norm of weights: 3.143048793959732

Iteration Number: 2588
Loss: 25.88845193636351
l2 norm of gradients: 0.3608810347961749
l2 norm of weights: 3.1423343092444878

Iteration Number: 2589
Loss: 25.877202447032285
l2 norm of gradients: 0.3607179365512003
l2 norm of weights: 3.141620217194716

```

-----
Iteration Number: 2590
Loss: 25.865961066296702
l2 norm of gradients: 0.36055490976414734
l2 norm of weights: 3.1409065177064006
-----
Iteration Number: 2591
Loss: 25.854727789085622
l2 norm of gradients: 0.3603919544037701
l2 norm of weights: 3.140193210675492
-----
Iteration Number: 2592
Loss: 25.843502610330876
l2 norm of gradients: 0.36022907043886937
l2 norm of weights: 3.139480295997904
-----
Iteration Number: 2593
Loss: 25.8322855249674
l2 norm of gradients: 0.36006625783829393
l2 norm of weights: 3.138767773569514
-----
Iteration Number: 2594
Loss: 25.82107652793291
l2 norm of gradients: 0.35990351657094016
l2 norm of weights: 3.138055643286163
-----
Iteration Number: 2595
Loss: 25.80987561416814
l2 norm of gradients: 0.3597408466057514
l2 norm of weights: 3.1373439050436582
-----
Iteration Number: 2596
Loss: 25.79868277861685
l2 norm of gradients: 0.35957824791171894
l2 norm of weights: 3.136632558737769
-----
Iteration Number: 2597
Loss: 25.787498016225594
l2 norm of gradients: 0.35941572045788156
l2 norm of weights: 3.135921604264228
-----
Iteration Number: 2598
Loss: 25.776321321944124
l2 norm of gradients: 0.3592532642133258
l2 norm of weights: 3.1352110415187338
-----
Iteration Number: 2599
Loss: 25.76515269072491

```

```

12 norm of gradients: 0.3590908791471856
12 norm of weights: 3.134500870396946
-----
Iteration Number: 2600
Loss: 25.753992117523456
12 norm of gradients: 0.35892856522864275
12 norm of weights: 3.1337910907944897
-----
Iteration Number: 2601
Loss: 25.74283959729819
12 norm of gradients: 0.35876632242692663
12 norm of weights: 3.133081702606953
-----
Iteration Number: 2602
Loss: 25.731695125010503
12 norm of gradients: 0.3586041507113146
12 norm of weights: 3.132372705729887
-----
Iteration Number: 2603
Loss: 25.720558695624593
12 norm of gradients: 0.35844205005113144
12 norm of weights: 3.131664100058807
-----
Iteration Number: 2604
Loss: 25.709430304107766
12 norm of gradients: 0.35828002041574997
12 norm of weights: 3.1309558854891906
-----
Iteration Number: 2605
Loss: 25.698309945430115
12 norm of gradients: 0.35811806177459093
12 norm of weights: 3.1302480619164794
-----
Iteration Number: 2606
Loss: 25.68719761456463
12 norm of gradients: 0.35795617409712266
12 norm of weights: 3.129540629236078
-----
Iteration Number: 2607
Loss: 25.676093306487328
12 norm of gradients: 0.3577943573528615
12 norm of weights: 3.1288335873433533
-----
Iteration Number: 2608
Loss: 25.664997016177004
12 norm of gradients: 0.35763261151137177
12 norm of weights: 3.128126936133637
-----

```

Iteration Number: 2609
Loss: 25.653908738615456
l2 norm of gradients: 0.35747093654226564
l2 norm of weights: 3.127420675502222

Iteration Number: 2610
Loss: 25.642828468787307
l2 norm of gradients: 0.35730933241520324
l2 norm of weights: 3.1267148053443647

Iteration Number: 2611
Loss: 25.631756201680208
l2 norm of gradients: 0.35714779909989275
l2 norm of weights: 3.1260093255552848

Iteration Number: 2612
Loss: 25.620691932284423
l2 norm of gradients: 0.35698633656609025
l2 norm of weights: 3.125304236030164

Iteration Number: 2613
Loss: 25.609635655593394
l2 norm of gradients: 0.35682494478360005
l2 norm of weights: 3.1245995366641472

Iteration Number: 2614
Loss: 25.59858736660326
l2 norm of gradients: 0.3566636237222742
l2 norm of weights: 3.1238952273523424

Iteration Number: 2615
Loss: 25.587547060313177
l2 norm of gradients: 0.35650237335201335
l2 norm of weights: 3.123191307989818

Iteration Number: 2616
Loss: 25.576514731724973
l2 norm of gradients: 0.3563411936427656
l2 norm of weights: 3.122487778471607

Iteration Number: 2617
Loss: 25.565490375843616
l2 norm of gradients: 0.3561800845645276
l2 norm of weights: 3.121784638692705

Iteration Number: 2618
Loss: 25.55447398767667
l2 norm of gradients: 0.3560190460873441


```
12 norm of weights: 3.1210818885480682
-----
Iteration Number: 2619
Loss: 25.54346556223469
12 norm of gradients: 0.3558580781813078
12 norm of weights: 3.120379527932617
-----
Iteration Number: 2620
Loss: 25.53246509453111
12 norm of gradients: 0.35569718081655993
12 norm of weights: 3.119677556741232
-----
Iteration Number: 2621
Loss: 25.521472579582184
12 norm of gradients: 0.3555363539632895
12 norm of weights: 3.118975974868758
-----
Iteration Number: 2622
Loss: 25.51048801240694
12 norm of gradients: 0.35537559759173415
12 norm of weights: 3.1182747822100003
-----
Iteration Number: 2623
Loss: 25.499511388027383
12 norm of gradients: 0.3552149116721795
12 norm of weights: 3.1175739786597285
-----
Iteration Number: 2624
Loss: 25.48854270146832
12 norm of gradients: 0.35505429617495926
12 norm of weights: 3.1168735641126712
-----
Iteration Number: 2625
Loss: 25.47758194775729
12 norm of gradients: 0.354893751070456
12 norm of weights: 3.1161735384635207
-----
Iteration Number: 2626
Loss: 25.466629121924775
12 norm of gradients: 0.3547332763291001
12 norm of weights: 3.1154739016069315
-----
Iteration Number: 2627
Loss: 25.455684219004088
12 norm of gradients: 0.3545728719213702
12 norm of weights: 3.114774653437519
-----
Iteration Number: 2628
```

Loss: 25.444747234031198
l2 norm of gradients: 0.3544125378177935
l2 norm of weights: 3.1140757938498607

Iteration Number: 2629
Loss: 25.433818162045213
l2 norm of gradients: 0.35425227398894554
l2 norm of weights: 3.1133773227384967

Iteration Number: 2630
Loss: 25.42289699808773
l2 norm of gradients: 0.3540920804054499
l2 norm of weights: 3.112679239997927

Iteration Number: 2631
Loss: 25.411983737203357
l2 norm of gradients: 0.3539319570379789
l2 norm of weights: 3.1119815455226143

Iteration Number: 2632
Loss: 25.40107837443941
l2 norm of gradients: 0.3537719038572529
l2 norm of weights: 3.1112842392069835

Iteration Number: 2633
Loss: 25.390180904846094
l2 norm of gradients: 0.35361192083404086
l2 norm of weights: 3.1105873209454193

Iteration Number: 2634
Loss: 25.379291323476306
l2 norm of gradients: 0.35345200793916
l2 norm of weights: 3.109890790632269

Iteration Number: 2635
Loss: 25.368409625385883
l2 norm of gradients: 0.35329216514347617
l2 norm of weights: 3.1091946481618415

Iteration Number: 2636
Loss: 25.357535805633322
l2 norm of gradients: 0.35313239241790323
l2 norm of weights: 3.1084988934284064

Iteration Number: 2637
Loss: 25.34666985927994
l2 norm of gradients: 0.3529726897334038
l2 norm of weights: 3.1078035263261943

```

-----
Iteration Number: 2638
Loss: 25.335811781389886
l2 norm of gradients: 0.35281305706098887
l2 norm of weights: 3.1071085467493984
-----
Iteration Number: 2639
Loss: 25.32496156703006
l2 norm of gradients: 0.3526534943717177
l2 norm of weights: 3.106413954592171
-----
Iteration Number: 2640
Loss: 25.31411921127013
l2 norm of gradients: 0.3524940016366984
l2 norm of weights: 3.1057197497486286
-----
Iteration Number: 2641
Loss: 25.303284709182567
l2 norm of gradients: 0.352334578827087
l2 norm of weights: 3.1050259321128455
-----
Iteration Number: 2642
Loss: 25.292458055842616
l2 norm of gradients: 0.3521752259140884
l2 norm of weights: 3.1043325015788596
-----
Iteration Number: 2643
Loss: 25.281639246328165
l2 norm of gradients: 0.3520159428689558
l2 norm of weights: 3.1036394580406674
-----
Iteration Number: 2644
Loss: 25.27082827572004
l2 norm of gradients: 0.3518567296629909
l2 norm of weights: 3.1029468013922292
-----
Iteration Number: 2645
Loss: 25.260025139101735
l2 norm of gradients: 0.3516975862675439
l2 norm of weights: 3.102254531527464
-----
Iteration Number: 2646
Loss: 25.249229831559575
l2 norm of gradients: 0.35153851265401354
l2 norm of weights: 3.1015626483402525
-----
Iteration Number: 2647
Loss: 25.23844234818243

```

```

12 norm of gradients: 0.3513795087938469
12 norm of weights: 3.1008711517244354
-----
Iteration Number: 2648
Loss: 25.227662684062146
12 norm of gradients: 0.3512205746585396
12 norm of weights: 3.100180041573816
-----
Iteration Number: 2649
Loss: 25.216890834293192
12 norm of gradients: 0.35106171021963606
12 norm of weights: 3.099489317782156
-----
Iteration Number: 2650
Loss: 25.206126793972857
12 norm of gradients: 0.3509029154487287
12 norm of weights: 3.0987989802431803
-----
Iteration Number: 2651
Loss: 25.195370558201045
12 norm of gradients: 0.35074419031745885
12 norm of weights: 3.0981090288505717
-----
Iteration Number: 2652
Loss: 25.18462212208053
12 norm of gradients: 0.3505855347975161
12 norm of weights: 3.0974194634979764
-----
Iteration Number: 2653
Loss: 25.17388148071668
12 norm of gradients: 0.35042694886063874
12 norm of weights: 3.096730284078998
-----
Iteration Number: 2654
Loss: 25.16314862921769
12 norm of gradients: 0.3502684324786134
12 norm of weights: 3.0960414904872042
-----
Iteration Number: 2655
Loss: 25.152423562694455
12 norm of gradients: 0.3501099856232752
12 norm of weights: 3.0953530826161204
-----
Iteration Number: 2656
Loss: 25.14170627626056
12 norm of gradients: 0.349951608266508
12 norm of weights: 3.094665060359233
-----

```

Iteration Number: 2657
Loss: 25.130996765032258
l2 norm of gradients: 0.349793300380244
l2 norm of weights: 3.09397742360999

Iteration Number: 2658
Loss: 25.120295024128648
l2 norm of gradients: 0.34963506193646393
l2 norm of weights: 3.0932901722617987

Iteration Number: 2659
Loss: 25.109601048671443
l2 norm of gradients: 0.34947689290719713
l2 norm of weights: 3.0926033062080265

Iteration Number: 2660
Loss: 25.098914833784995
l2 norm of gradients: 0.3493187932645211
l2 norm of weights: 3.0919168253420017

Iteration Number: 2661
Loss: 25.08823637459653
l2 norm of gradients: 0.34916076298056237
l2 norm of weights: 3.091230729557013

Iteration Number: 2662
Loss: 25.077565666235806
l2 norm of gradients: 0.34900280202749545
l2 norm of weights: 3.090545018746308

Iteration Number: 2663
Loss: 25.066902703835357
l2 norm of gradients: 0.3488449103775438
l2 norm of weights: 3.0898596928030964

Iteration Number: 2664
Loss: 25.056247482530374
l2 norm of gradients: 0.3486870880029789
l2 norm of weights: 3.0891747516205466

Iteration Number: 2665
Loss: 25.0455999974588
l2 norm of gradients: 0.3485293348761212
l2 norm of weights: 3.088490195091788

Iteration Number: 2666
Loss: 25.034960243761063
l2 norm of gradients: 0.3483716509693392

```

12 norm of weights: 3.087806023109908
-----
Iteration Number: 2667
Loss: 25.024328216580518
12 norm of gradients: 0.3482140362550502
12 norm of weights: 3.0871222355679575
-----
Iteration Number: 2668
Loss: 25.013703911063097
12 norm of gradients: 0.3480564907057198
12 norm of weights: 3.0864388323589442
-----
Iteration Number: 2669
Loss: 25.003087322357363
12 norm of gradients: 0.3478990142938621
12 norm of weights: 3.0857558133758376
-----
Iteration Number: 2670
Loss: 24.992478445614523
12 norm of gradients: 0.3477416069920396
12 norm of weights: 3.085073178511567
-----
Iteration Number: 2671
Loss: 24.981877275988538
12 norm of gradients: 0.34758426877286347
12 norm of weights: 3.08439092765902
-----
Iteration Number: 2672
Loss: 24.971283808635953
12 norm of gradients: 0.34742699960899304
12 norm of weights: 3.0837090607110453
-----
Iteration Number: 2673
Loss: 24.96069803871605
12 norm of gradients: 0.34726979947313624
12 norm of weights: 3.0830275775604528
-----
Iteration Number: 2674
Loss: 24.950119961390683
12 norm of gradients: 0.3471126683380493
12 norm of weights: 3.0823464781000087
-----
Iteration Number: 2675
Loss: 24.93954957182438
12 norm of gradients: 0.34695560617653703
12 norm of weights: 3.0816657622224426
-----
Iteration Number: 2676

```

Loss: 24.92898686518435
l2 norm of gradients: 0.3467986129614524
l2 norm of weights: 3.080985429820441

Iteration Number: 2677
Loss: 24.91843183664042
l2 norm of gradients: 0.3466416886656971
l2 norm of weights: 3.080305480786652

Iteration Number: 2678
Loss: 24.907884481365024
l2 norm of gradients: 0.3464848332622209
l2 norm of weights: 3.079625915013682

Iteration Number: 2679
Loss: 24.89734479453325
l2 norm of gradients: 0.34632804672402207
l2 norm of weights: 3.078946732394099

Iteration Number: 2680
Loss: 24.88681277132288
l2 norm of gradients: 0.34617132902414727
l2 norm of weights: 3.078267932820428

Iteration Number: 2681
Loss: 24.876288406914227
l2 norm of gradients: 0.3460146801356914
l2 norm of weights: 3.0775895161851547

Iteration Number: 2682
Loss: 24.865771696490324
l2 norm of gradients: 0.34585810003179795
l2 norm of weights: 3.076911482380726

Iteration Number: 2683
Loss: 24.8552626352367
l2 norm of gradients: 0.3457015886856583
l2 norm of weights: 3.076233831299546

Iteration Number: 2684
Loss: 24.84476121834165
l2 norm of gradients: 0.34554514607051257
l2 norm of weights: 3.075556562833979

Iteration Number: 2685
Loss: 24.834267440995983
l2 norm of gradients: 0.34538877215964886
l2 norm of weights: 3.0748796768763498

```

-----
Iteration Number: 2686
Loss: 24.823781298393126
l2 norm of gradients: 0.34523246692640375
l2 norm of weights: 3.0742031733189408
-----
Iteration Number: 2687
Loss: 24.813302785729217
l2 norm of gradients: 0.345076230344162
l2 norm of weights: 3.073527052053996
-----
Iteration Number: 2688
Loss: 24.80283189820286
l2 norm of gradients: 0.34492006238635675
l2 norm of weights: 3.0728513129737167
-----
Iteration Number: 2689
Loss: 24.79236863101534
l2 norm of gradients: 0.34476396302646906
l2 norm of weights: 3.0721759559702653
-----
Iteration Number: 2690
Loss: 24.781912979370517
l2 norm of gradients: 0.34460793223802844
l2 norm of weights: 3.0715009809357623
-----
Iteration Number: 2691
Loss: 24.771464938474825
l2 norm of gradients: 0.3444519699946126
l2 norm of weights: 3.0708263877622883
-----
Iteration Number: 2692
Loss: 24.76102450353735
l2 norm of gradients: 0.34429607626984743
l2 norm of weights: 3.070152176341883
-----
Iteration Number: 2693
Loss: 24.750591669769758
l2 norm of gradients: 0.3441402510374068
l2 norm of weights: 3.069478346566545
-----
Iteration Number: 2694
Loss: 24.740166432386218
l2 norm of gradients: 0.34398449427101324
l2 norm of weights: 3.0688048983282337
-----
Iteration Number: 2695
Loss: 24.72974878660352

```



```

12 norm of gradients: 0.34382880594443666
12 norm of weights: 3.068131831518865
-----
Iteration Number: 2696
Loss: 24.71933872764113
12 norm of gradients: 0.3436731860314956
12 norm of weights: 3.0674591460303167
-----
Iteration Number: 2697
Loss: 24.70893625072096
12 norm of gradients: 0.34351763450605666
12 norm of weights: 3.0667868417544235
-----
Iteration Number: 2698
Loss: 24.698541351067583
12 norm of gradients: 0.3433621513420343
12 norm of weights: 3.066114918582982
-----
Iteration Number: 2699
Loss: 24.688154023908012
12 norm of gradients: 0.3432067365133913
12 norm of weights: 3.0654433764077456
-----
Iteration Number: 2700
Loss: 24.677774264472024
12 norm of gradients: 0.34305138999413815
12 norm of weights: 3.064772215120428
-----
Iteration Number: 2701
Loss: 24.6674020679918
12 norm of gradients: 0.34289611175833357
12 norm of weights: 3.0641014346127022
-----
Iteration Number: 2702
Loss: 24.657037429702118
12 norm of gradients: 0.3427409017800842
12 norm of weights: 3.063431034776199
-----
Iteration Number: 2703
Loss: 24.646680344840348
12 norm of gradients: 0.3425857600335448
12 norm of weights: 3.062761015502509
-----
Iteration Number: 2704
Loss: 24.636330808646342
12 norm of gradients: 0.3424306864929178
12 norm of weights: 3.0620913766831834
-----

```

Iteration Number: 2705
Loss: 24.625988816362593
l2 norm of gradients: 0.3422756811324538
l2 norm of weights: 3.0614221182097303

Iteration Number: 2706
Loss: 24.615654363234132
l2 norm of gradients: 0.3421207439264511
l2 norm of weights: 3.060753239973617

Iteration Number: 2707
Loss: 24.60532744450836
l2 norm of gradients: 0.341965874849256
l2 norm of weights: 3.0600847418662718

Iteration Number: 2708
Loss: 24.595008055435535
l2 norm of gradients: 0.34181107387526266
l2 norm of weights: 3.059416623779079

Iteration Number: 2709
Loss: 24.584696191268158
l2 norm of gradients: 0.341656340978913
l2 norm of weights: 3.0587488856033858

Iteration Number: 2710
Loss: 24.57439184726141
l2 norm of gradients: 0.34150167613469684
l2 norm of weights: 3.0580815272304944

Iteration Number: 2711
Loss: 24.564095018672912
l2 norm of gradients: 0.34134707931715164
l2 norm of weights: 3.057414548551668

Iteration Number: 2712
Loss: 24.553805700762986
l2 norm of gradients: 0.34119255050086267
l2 norm of weights: 3.0567479494581296

Iteration Number: 2713
Loss: 24.54352388879425
l2 norm of gradients: 0.341038089660463
l2 norm of weights: 3.056081729841059

Iteration Number: 2714
Loss: 24.53324957803205
l2 norm of gradients: 0.3408836967706333

```

12 norm of weights: 3.0554158895915964
-----
Iteration Number: 2715
Loss: 24.522982763744093
12 norm of gradients: 0.3407293718061021
12 norm of weights: 3.054750428600841
-----
Iteration Number: 2716
Loss: 24.512723441200695
12 norm of gradients: 0.34057511474164526
12 norm of weights: 3.054085346759849
-----
Iteration Number: 2717
Loss: 24.50247160567464
12 norm of gradients: 0.34042092555208653
12 norm of weights: 3.053420643959638
-----
Iteration Number: 2718
Loss: 24.492227252441268
12 norm of gradients: 0.3402668042122971
12 norm of weights: 3.0527563200911834
-----
Iteration Number: 2719
Loss: 24.481990376778338
12 norm of gradients: 0.34011275069719593
12 norm of weights: 3.0520923750454187
-----
Iteration Number: 2720
Loss: 24.4717609739662
12 norm of gradients: 0.3399587649817492
12 norm of weights: 3.0514288087132377
-----
Iteration Number: 2721
Loss: 24.46153903928768
12 norm of gradients: 0.3398048470409708
12 norm of weights: 3.050765620985492
-----
Iteration Number: 2722
Loss: 24.451324568028063
12 norm of gradients: 0.3396509968499222
12 norm of weights: 3.0501028117529927
-----
Iteration Number: 2723
Loss: 24.44111755547519
12 norm of gradients: 0.33949721438371194
12 norm of weights: 3.049440380906509
-----
Iteration Number: 2724

```

Loss: 24.430917996919337
l2 norm of gradients: 0.33934349961749644
l2 norm of weights: 3.04877832833677

Iteration Number: 2725
Loss: 24.42072588765327
l2 norm of gradients: 0.3391898525264791
l2 norm of weights: 3.0481166539344633

Iteration Number: 2726
Loss: 24.410541222972277
l2 norm of gradients: 0.33903627308591106
l2 norm of weights: 3.0474553575902337

Iteration Number: 2727
Loss: 24.4003639981741
l2 norm of gradients: 0.33888276127109046
l2 norm of weights: 3.046794439194687

Iteration Number: 2728
Loss: 24.390194208559027
l2 norm of gradients: 0.3387293170573628
l2 norm of weights: 3.046133898638387

Iteration Number: 2729
Loss: 24.380031849429685
l2 norm of gradients: 0.3385759404201211
l2 norm of weights: 3.0454737358118553

Iteration Number: 2730
Loss: 24.36987691609133
l2 norm of gradients: 0.33842263133480516
l2 norm of weights: 3.044813950605574

Iteration Number: 2731
Loss: 24.359729403851496
l2 norm of gradients: 0.33826938977690235
l2 norm of weights: 3.044154542909983

Iteration Number: 2732
Loss: 24.349589308020438
l2 norm of gradients: 0.33811621572194717
l2 norm of weights: 3.0434955126154812

Iteration Number: 2733
Loss: 24.33945662391061
l2 norm of gradients: 0.33796310914552097
l2 norm of weights: 3.0428368596124256

```

-----
Iteration Number: 2734
Loss: 24.329331346837144
l2 norm of gradients: 0.3378100700232525
l2 norm of weights: 3.042178583791133
-----
Iteration Number: 2735
Loss: 24.31921347211749
l2 norm of gradients: 0.3376570983308173
l2 norm of weights: 3.0415206850418786
-----
Iteration Number: 2736
Loss: 24.309102995071623
l2 norm of gradients: 0.3375041940439383
l2 norm of weights: 3.0408631632548957
-----
Iteration Number: 2737
Loss: 24.298999911021962
l2 norm of gradients: 0.337351357138385
l2 norm of weights: 3.040206018320377
-----
Iteration Number: 2738
Loss: 24.28890421529329
l2 norm of gradients: 0.3371985875899742
l2 norm of weights: 3.0395492501284744
-----
Iteration Number: 2739
Loss: 24.278815903212976
l2 norm of gradients: 0.3370458853745694
l2 norm of weights: 3.038892858569297
-----
Iteration Number: 2740
Loss: 24.26873497011077
l2 norm of gradients: 0.33689325046808116
l2 norm of weights: 3.038236843532915
-----
Iteration Number: 2741
Loss: 24.258661411318815
l2 norm of gradients: 0.3367406828464667
l2 norm of weights: 3.0375812049093542
-----
Iteration Number: 2742
Loss: 24.24859522217174
l2 norm of gradients: 0.33658818248573025
l2 norm of weights: 3.0369259425886024
-----
Iteration Number: 2743
Loss: 24.238536398006655

```

```

12 norm of gradients: 0.3364357493619225
12 norm of weights: 3.0362710564606035
-----
Iteration Number: 2744
Loss: 24.228484934162992
12 norm of gradients: 0.33628338345114134
12 norm of weights: 3.0356165464152616
-----
Iteration Number: 2745
Loss: 24.218440825982672
12 norm of gradients: 0.336131084729531
12 norm of weights: 3.034962412342439
-----
Iteration Number: 2746
Loss: 24.208404068810083
12 norm of gradients: 0.3359788531732825
12 norm of weights: 3.034308654131957
-----
Iteration Number: 2747
Loss: 24.198374657991945
12 norm of gradients: 0.3358266887586334
12 norm of weights: 3.0336552716735956
-----
Iteration Number: 2748
Loss: 24.188352588877496
12 norm of gradients: 0.335674591461868
12 norm of weights: 3.0330022648570933
-----
Iteration Number: 2749
Loss: 24.17833785681829
12 norm of gradients: 0.335522561259317
12 norm of weights: 3.032349633572147
-----
Iteration Number: 2750
Loss: 24.168330457168377
12 norm of gradients: 0.3353705981273576
12 norm of weights: 3.031697377708414
-----
Iteration Number: 2751
Loss: 24.15833038528423
12 norm of gradients: 0.3352187020424136
12 norm of weights: 3.031045497155508
-----
Iteration Number: 2752
Loss: 24.14833763652462
12 norm of gradients: 0.33506687298095517
12 norm of weights: 3.0303939918030034
-----

```

Iteration Number: 2753
Loss: 24.13835220625081
l2 norm of gradients: 0.33491511091949877
l2 norm of weights: 3.029742861540431

Iteration Number: 2754
Loss: 24.12837408982649
l2 norm of gradients: 0.3347634158346072
l2 norm of weights: 3.0290921062572833

Iteration Number: 2755
Loss: 24.118403282617678
l2 norm of gradients: 0.33461178770288985
l2 norm of weights: 3.028441725843009

Iteration Number: 2756
Loss: 24.108439779992857
l2 norm of gradients: 0.334460226501002
l2 norm of weights: 3.0277917201870177

Iteration Number: 2757
Loss: 24.098483577322877
l2 norm of gradients: 0.33430873220564544
l2 norm of weights: 3.0271420891786756

Iteration Number: 2758
Loss: 24.088534669980948
l2 norm of gradients: 0.3341573047935678
l2 norm of weights: 3.0264928327073095

Iteration Number: 2759
Loss: 24.07859305334269
l2 norm of gradients: 0.33400594424156327
l2 norm of weights: 3.025843950662203

Iteration Number: 2760
Loss: 24.068658722786196
l2 norm of gradients: 0.3338546505264718
l2 norm of weights: 3.0251954429326005

Iteration Number: 2761
Loss: 24.058731673691756
l2 norm of gradients: 0.33370342362517946
l2 norm of weights: 3.0245473094077036

Iteration Number: 2762
Loss: 24.048811901442285
l2 norm of gradients: 0.3335522635146185

```

12 norm of weights: 3.023899549976674
-----
Iteration Number: 2763
Loss: 24.03889940142282
12 norm of gradients: 0.3334011701717669
12 norm of weights: 3.0232521645286305
-----
Iteration Number: 2764
Loss: 24.028994169020983
12 norm of gradients: 0.33325014357364874
12 norm of weights: 3.022605152952653
-----
Iteration Number: 2765
Loss: 24.019096199626638
12 norm of gradients: 0.3330991836973338
12 norm of weights: 3.021958515137778
-----
Iteration Number: 2766
Loss: 24.0092054886321
12 norm of gradients: 0.33294829051993796
12 norm of weights: 3.0213122509730015
-----
Iteration Number: 2767
Loss: 23.999322031432005
12 norm of gradients: 0.3327974640186224
12 norm of weights: 3.020666360347278
-----
Iteration Number: 2768
Loss: 23.9894458234234
12 norm of gradients: 0.3326467041705947
12 norm of weights: 3.0200208431495223
-----
Iteration Number: 2769
Loss: 23.979576860005583
12 norm of gradients: 0.33249601095310755
12 norm of weights: 3.0193756992686063
-----
Iteration Number: 2770
Loss: 23.969715136580383
12 norm of gradients: 0.3323453843434596
12 norm of weights: 3.018730928593361
-----
Iteration Number: 2771
Loss: 23.95986064855185
12 norm of gradients: 0.332194824318995
12 norm of weights: 3.0180865310125773
-----
Iteration Number: 2772

```


Loss: 23.950013391326447
l2 norm of gradients: 0.33204433085710344
l2 norm of weights: 3.0174425064150032

Iteration Number: 2773
Loss: 23.940173360312947
l2 norm of gradients: 0.33189390393522017
l2 norm of weights: 3.0167988546893474

Iteration Number: 2774
Loss: 23.930340550922597
l2 norm of gradients: 0.33174354353082586
l2 norm of weights: 3.0161555757242757

Iteration Number: 2775
Loss: 23.920514958568763
l2 norm of gradients: 0.3315932496214467
l2 norm of weights: 3.0155126694084142

Iteration Number: 2776
Loss: 23.910696578667398
l2 norm of gradients: 0.3314430221846541
l2 norm of weights: 3.014870135630347

Iteration Number: 2777
Loss: 23.90088540663664
l2 norm of gradients: 0.33129286119806484
l2 norm of weights: 3.0142279742786178

Iteration Number: 2778
Loss: 23.891081437897014
l2 norm of gradients: 0.33114276663934106
l2 norm of weights: 3.013586185241728

Iteration Number: 2779
Loss: 23.88128466787134
l2 norm of gradients: 0.33099273848618993
l2 norm of weights: 3.012944768408139

Iteration Number: 2780
Loss: 23.87149509198492
l2 norm of gradients: 0.33084277671636386
l2 norm of weights: 3.0123037236662706

Iteration Number: 2781
Loss: 23.861712705665212
l2 norm of gradients: 0.33069288130766056
l2 norm of weights: 3.011663050904502

```

-----
Iteration Number: 2782
Loss: 23.851937504342033
l2 norm of gradients: 0.3305430522379226
l2 norm of weights: 3.0110227500111697
-----
Iteration Number: 2783
Loss: 23.842169483447634
l2 norm of gradients: 0.3303932894850377
l2 norm of weights: 3.0103828208745718
-----
Iteration Number: 2784
Loss: 23.832408638416453
l2 norm of gradients: 0.33024359302693845
l2 norm of weights: 3.0097432633829633
-----
Iteration Number: 2785
Loss: 23.822654964685377
l2 norm of gradients: 0.3300939628416023
l2 norm of weights: 3.009104077424559
-----
Iteration Number: 2786
Loss: 23.81290845769349
l2 norm of gradients: 0.3299443989070519
l2 norm of weights: 3.008465262887532
-----
Iteration Number: 2787
Loss: 23.803169112882284
l2 norm of gradients: 0.32979490120135446
l2 norm of weights: 3.007826819660015
-----
Iteration Number: 2788
Loss: 23.793436925695502
l2 norm of gradients: 0.3296454697026219
l2 norm of weights: 3.0071887476301
-----
Iteration Number: 2789
Loss: 23.78371189157924
l2 norm of gradients: 0.329496104389011
l2 norm of weights: 3.006551046685837
-----
Iteration Number: 2790
Loss: 23.77399400598191
l2 norm of gradients: 0.3293468052387232
l2 norm of weights: 3.005913716715235
-----
Iteration Number: 2791
Loss: 23.764283264354145

```

```

12 norm of gradients: 0.32919757223000445
12 norm of weights: 3.0052767576062647
-----
Iteration Number: 2792
Loss: 23.754579662149006
12 norm of gradients: 0.32904840534114543
12 norm of weights: 3.0046401692468514
-----
Iteration Number: 2793
Loss: 23.744883194821732
12 norm of gradients: 0.32889930455048116
12 norm of weights: 3.004003951524883
-----
Iteration Number: 2794
Loss: 23.735193857829962
12 norm of gradients: 0.3287502698363913
12 norm of weights: 3.003368104328205
-----
Iteration Number: 2795
Loss: 23.725511646633578
12 norm of gradients: 0.3286013011772997
12 norm of weights: 3.0027326275446216
-----
Iteration Number: 2796
Loss: 23.715836556694747
12 norm of gradients: 0.3284523985516747
12 norm of weights: 3.002097521061898
-----
Iteration Number: 2797
Loss: 23.706168583477897
12 norm of gradients: 0.32830356193802884
12 norm of weights: 3.001462784767756
-----
Iteration Number: 2798
Loss: 23.69650772244987
12 norm of gradients: 0.328154791314919
12 norm of weights: 3.0008284185498795
-----
Iteration Number: 2799
Loss: 23.686853969079674
12 norm of gradients: 0.32800608666094627
12 norm of weights: 3.000194422295908
-----
Iteration Number: 2800
Loss: 23.677207318838647
12 norm of gradients: 0.32785744795475574
12 norm of weights: 2.9995607958934434
-----

```

Iteration Number: 2801
Loss: 23.667567767200364
l2 norm of gradients: 0.3277088751750366
l2 norm of weights: 2.9989275392300443

Iteration Number: 2802
Loss: 23.657935309640767
l2 norm of gradients: 0.3275603683005223
l2 norm of weights: 2.99829465219323

Iteration Number: 2803
Loss: 23.648309941637955
l2 norm of gradients: 0.3274119273099899
l2 norm of weights: 2.9976621346704797

Iteration Number: 2804
Loss: 23.6386916586724
l2 norm of gradients: 0.3272635521822606
l2 norm of weights: 2.9970299865492294

Iteration Number: 2805
Loss: 23.629080456226788
l2 norm of gradients: 0.3271152428961994
l2 norm of weights: 2.996398207716876

Iteration Number: 2806
Loss: 23.619476329786128
l2 norm of gradients: 0.32696699943071517
l2 norm of weights: 2.9957667980607754

Iteration Number: 2807
Loss: 23.609879274837617
l2 norm of gradients: 0.32681882176476024
l2 norm of weights: 2.995135757468243

Iteration Number: 2808
Loss: 23.600289286870737
l2 norm of gradients: 0.32667070987733104
l2 norm of weights: 2.994505085826553

Iteration Number: 2809
Loss: 23.59070636137732
l2 norm of gradients: 0.3265226637474674
l2 norm of weights: 2.9938747830229393

Iteration Number: 2810
Loss: 23.581130493851315
l2 norm of gradients: 0.32637468335425274

```

12 norm of weights: 2.993244848944595
-----
Iteration Number: 2811
Loss: 23.571561679789024
12 norm of gradients: 0.32622676867681394
12 norm of weights: 2.9926152834786723
-----
Iteration Number: 2812
Loss: 23.561999914688982
12 norm of gradients: 0.3260789196943215
12 norm of weights: 2.9919860865122834
-----
Iteration Number: 2813
Loss: 23.552445194051987
12 norm of gradients: 0.3259311363859892
12 norm of weights: 2.9913572579324987
-----
Iteration Number: 2814
Loss: 23.542897513381018
12 norm of gradients: 0.32578341873107414
12 norm of weights: 2.99072879762635
-----
Iteration Number: 2815
Loss: 23.533356868181368
12 norm of gradients: 0.32563576670887684
12 norm of weights: 2.990100705480827
-----
Iteration Number: 2816
Loss: 23.52382325396058
12 norm of gradients: 0.32548818029874094
12 norm of weights: 2.9894729813828786
-----
Iteration Number: 2817
Loss: 23.514296666228358
12 norm of gradients: 0.3253406594800533
12 norm of weights: 2.988845625219414
-----
Iteration Number: 2818
Loss: 23.504777100496767
12 norm of gradients: 0.32519320423224357
12 norm of weights: 2.988218636877302
-----
Iteration Number: 2819
Loss: 23.49526455228001
12 norm of gradients: 0.32504581453478504
12 norm of weights: 2.987592016243371
-----
Iteration Number: 2820

```

Loss: 23.485759017094527
l2 norm of gradients: 0.3248984903671936
l2 norm of weights: 2.986965763204407

Iteration Number: 2821
Loss: 23.47626049045905
l2 norm of gradients: 0.3247512317090281
l2 norm of weights: 2.986339877647158

Iteration Number: 2822
Loss: 23.466768967894495
l2 norm of gradients: 0.32460403853989034
l2 norm of weights: 2.9857143594583313

Iteration Number: 2823
Loss: 23.457284444924024
l2 norm of gradients: 0.32445691083942485
l2 norm of weights: 2.985089208524592

Iteration Number: 2824
Loss: 23.44780691707303
l2 norm of gradients: 0.324309848587319
l2 norm of weights: 2.9844644247325665

Iteration Number: 2825
Loss: 23.438336379869085
l2 norm of gradients: 0.3241628517633027
l2 norm of weights: 2.98384000796884

Iteration Number: 2826
Loss: 23.428872828842064
l2 norm of gradients: 0.3240159203471486
l2 norm of weights: 2.983215958119957

Iteration Number: 2827
Loss: 23.419416259523967
l2 norm of gradients: 0.3238690543186719
l2 norm of weights: 2.9825922750724243

Iteration Number: 2828
Loss: 23.40996666744904
l2 norm of gradients: 0.3237222536577303
l2 norm of weights: 2.9819689587127045

Iteration Number: 2829
Loss: 23.400524048153788
l2 norm of gradients: 0.3235755183442239
l2 norm of weights: 2.981346008927223

```

-----
Iteration Number: 2830
Loss: 23.391088397176862
l2 norm of gradients: 0.3234288483580952
l2 norm of weights: 2.980723425602363
-----
Iteration Number: 2831
Loss: 23.381659710059175
l2 norm of gradients: 0.323282243679329
l2 norm of weights: 2.9801012086244683
-----
Iteration Number: 2832
Loss: 23.372237982343815
l2 norm of gradients: 0.3231357042879525
l2 norm of weights: 2.979479357879843
-----
Iteration Number: 2833
Loss: 23.362823209576064
l2 norm of gradients: 0.3229892301640348
l2 norm of weights: 2.97885787325475
-----
Iteration Number: 2834
Loss: 23.353415387303443
l2 norm of gradients: 0.3228428212876875
l2 norm of weights: 2.9782367546354123
-----
Iteration Number: 2835
Loss: 23.344014511075624
l2 norm of gradients: 0.3226964776390636
l2 norm of weights: 2.977616001908014
-----
Iteration Number: 2836
Loss: 23.33462057644451
l2 norm of gradients: 0.322550199198359
l2 norm of weights: 2.976995614958697
-----
Iteration Number: 2837
Loss: 23.325233578964234
l2 norm of gradients: 0.32240398594581093
l2 norm of weights: 2.976375593673564
-----
Iteration Number: 2838
Loss: 23.315853514191033
l2 norm of gradients: 0.32225783786169865
l2 norm of weights: 2.9757559379386795
-----
Iteration Number: 2839
Loss: 23.306480377683407

```

```

12 norm of gradients: 0.3221117549263432
12 norm of weights: 2.9751366476400642
-----
Iteration Number: 2840
Loss: 23.29711416500197
12 norm of gradients: 0.3219657371201074
12 norm of weights: 2.974517722663702
-----
Iteration Number: 2841
Loss: 23.28775487170962
12 norm of gradients: 0.3218197844233957
12 norm of weights: 2.9738991628955356
-----
Iteration Number: 2842
Loss: 23.27840249337136
12 norm of gradients: 0.32167389681665426
12 norm of weights: 2.9732809682214674
-----
Iteration Number: 2843
Loss: 23.26905702555441
12 norm of gradients: 0.3215280742803708
12 norm of weights: 2.9726631385273605
-----
Iteration Number: 2844
Loss: 23.259718463828136
12 norm of gradients: 0.32138231679507423
12 norm of weights: 2.972045673699038
-----
Iteration Number: 2845
Loss: 23.250386803764144
12 norm of gradients: 0.32123662434133526
12 norm of weights: 2.9714285736222825
-----
Iteration Number: 2846
Loss: 23.24106204093615
12 norm of gradients: 0.32109099689976583
12 norm of weights: 2.9708118381828372
-----
Iteration Number: 2847
Loss: 23.23174417092006
12 norm of gradients: 0.32094543445101903
12 norm of weights: 2.9701954672664064
-----
Iteration Number: 2848
Loss: 23.222433189293966
12 norm of gradients: 0.3207999369757893
12 norm of weights: 2.9695794607586525
-----

```


Iteration Number: 2849
Loss: 23.2131290916381
l2 norm of gradients: 0.32065450445481214
l2 norm of weights: 2.9689638185452

Iteration Number: 2850
Loss: 23.203831873534895
l2 norm of gradients: 0.3205091368688644
l2 norm of weights: 2.9683485405116325

Iteration Number: 2851
Loss: 23.1945415305689
l2 norm of gradients: 0.32036383419876363
l2 norm of weights: 2.9677336265434944

Iteration Number: 2852
Loss: 23.185258058326912
l2 norm of gradients: 0.3202185964253684
l2 norm of weights: 2.96711907652629

Iteration Number: 2853
Loss: 23.175981452397775
l2 norm of gradients: 0.32007342352957824
l2 norm of weights: 2.966504890345485

Iteration Number: 2854
Loss: 23.166711708372546
l2 norm of gradients: 0.3199283154923336
l2 norm of weights: 2.9658910678865036

Iteration Number: 2855
Loss: 23.157448821844458
l2 norm of gradients: 0.3197832722946154
l2 norm of weights: 2.965277609034732

Iteration Number: 2856
Loss: 23.148192788408835
l2 norm of gradients: 0.3196382939174455
l2 norm of weights: 2.9646645136755168

Iteration Number: 2857
Loss: 23.138943603663222
l2 norm of gradients: 0.31949338034188624
l2 norm of weights: 2.964051781694163

Iteration Number: 2858
Loss: 23.129701263207256
l2 norm of gradients: 0.3193485315490406

```
l2 norm of weights: 2.963439412975939
-----
Iteration Number: 2859
Loss: 23.12046576264274
l2 norm of gradients: 0.31920374752005193
l2 norm of weights: 2.9628274074060714
-----
Iteration Number: 2860
Loss: 23.11123709757361
l2 norm of gradients: 0.319059028236104
l2 norm of weights: 2.9622157648697485
-----
Iteration Number: 2861
Loss: 23.102015263605985
l2 norm of gradients: 0.318914373678421
l2 norm of weights: 2.961604485252119
-----
Iteration Number: 2862
Loss: 23.092800256348053
l2 norm of gradients: 0.3187697838282675
l2 norm of weights: 2.960993568438291
-----
Iteration Number: 2863
Loss: 23.08359207141019
l2 norm of gradients: 0.31862525866694785
l2 norm of weights: 2.9603830143133347
-----
Iteration Number: 2864
Loss: 23.074390704404934
l2 norm of gradients: 0.318480798175807
l2 norm of weights: 2.9597728227622806
-----
Iteration Number: 2865
Loss: 23.06519615094685
l2 norm of gradients: 0.31833640233622973
l2 norm of weights: 2.95916299367012
-----
Iteration Number: 2866
Loss: 23.0560084066527
l2 norm of gradients: 0.3181920711296409
l2 norm of weights: 2.9585535269218037
-----
Iteration Number: 2867
Loss: 23.046827467141433
l2 norm of gradients: 0.31804780453750503
l2 norm of weights: 2.957944422402245
-----
Iteration Number: 2868
```

```

Loss: 23.037653328034004
l2 norm of gradients: 0.317903602541327
l2 norm of weights: 2.9573356799963157
-----
Iteration Number: 2869
Loss: 23.02848598495359
l2 norm of gradients: 0.3177594651226509
l2 norm of weights: 2.956727299588852
-----
Iteration Number: 2870
Loss: 23.019325433525417
l2 norm of gradients: 0.317615392263061
l2 norm of weights: 2.956119281064646
-----
Iteration Number: 2871
Loss: 23.01017166937691
l2 norm of gradients: 0.31747138394418095
l2 norm of weights: 2.955511624308455
-----
Iteration Number: 2872
Loss: 23.001024688137505
l2 norm of gradients: 0.31732744014767406
l2 norm of weights: 2.9549043292049952
-----
Iteration Number: 2873
Loss: 22.991884485438856
l2 norm of gradients: 0.317183560855243
l2 norm of weights: 2.9542973956389433
-----
Iteration Number: 2874
Loss: 22.98275105691468
l2 norm of gradients: 0.3170397460486301
l2 norm of weights: 2.953690823494939
-----
Iteration Number: 2875
Loss: 22.973624398200837
l2 norm of gradients: 0.3168959957096169
l2 norm of weights: 2.95308461265758
-----
Iteration Number: 2876
Loss: 22.964504504935224
l2 norm of gradients: 0.3167523098200242
l2 norm of weights: 2.9524787630114266
-----
Iteration Number: 2877
Loss: 22.955391372757916
l2 norm of gradients: 0.3166086883617121
l2 norm of weights: 2.951873274441002

```

```

-----
Iteration Number: 2878
Loss: 22.94628499731107
l2 norm of gradients: 0.31646513131657966
l2 norm of weights: 2.9512681468307864
-----
Iteration Number: 2879
Loss: 22.93718537423896
l2 norm of gradients: 0.3163216386665652
l2 norm of weights: 2.950663380065224
-----
Iteration Number: 2880
Loss: 22.92809249918795
l2 norm of gradients: 0.316178210393646
l2 norm of weights: 2.95005897402872
-----
Iteration Number: 2881
Loss: 22.91900636780645
l2 norm of gradients: 0.31603484647983815
l2 norm of weights: 2.9494549286056393
-----
Iteration Number: 2882
Loss: 22.909926975745087
l2 norm of gradients: 0.31589154690719673
l2 norm of weights: 2.9488512436803096
-----
Iteration Number: 2883
Loss: 22.900854318656457
l2 norm of gradients: 0.3157483116578154
l2 norm of weights: 2.948247919137018
-----
Iteration Number: 2884
Loss: 22.89178839219534
l2 norm of gradients: 0.31560514071382684
l2 norm of weights: 2.947644954860015
-----
Iteration Number: 2885
Loss: 22.882729192018527
l2 norm of gradients: 0.3154620340574018
l2 norm of weights: 2.94704235073351
-----
Iteration Number: 2886
Loss: 22.87367671378498
l2 norm of gradients: 0.3153189916707501
l2 norm of weights: 2.9464401066416763
-----
Iteration Number: 2887
Loss: 22.86463095315568

```

```

12 norm of gradients: 0.31517601353612007
12 norm of weights: 2.945838222468646
-----
Iteration Number: 2888
Loss: 22.855591905793734
12 norm of gradients: 0.3150330996357979
12 norm of weights: 2.945236698098515
-----
Iteration Number: 2889
Loss: 22.846559567364313
12 norm of gradients: 0.31489024995210857
12 norm of weights: 2.944635533415339
-----
Iteration Number: 2890
Loss: 22.83753393353467
12 norm of gradients: 0.3147474644674152
12 norm of weights: 2.944034728303135
-----
Iteration Number: 2891
Loss: 22.82851499997412
12 norm of gradients: 0.314604743164119
12 norm of weights: 2.9434342826458826
-----
Iteration Number: 2892
Loss: 22.819502762354116
12 norm of gradients: 0.31446208602465936
12 norm of weights: 2.9428341963275226
-----
Iteration Number: 2893
Loss: 22.810497216348082
12 norm of gradients: 0.3143194930315137
12 norm of weights: 2.9422344692319564
-----
Iteration Number: 2894
Loss: 22.801498357631623
12 norm of gradients: 0.3141769641671974
12 norm of weights: 2.941635101243049
-----
Iteration Number: 2895
Loss: 22.792506181882302
12 norm of gradients: 0.3140344994142637
12 norm of weights: 2.9410360922446244
-----
Iteration Number: 2896
Loss: 22.783520684779873
12 norm of gradients: 0.31389209875530355
12 norm of weights: 2.94043744212047
-----

```

Iteration Number: 2897
Loss: 22.77454186200606
l2 norm of gradients: 0.31374976217294576
l2 norm of weights: 2.939839150754335

Iteration Number: 2898
Loss: 22.76556970924467
l2 norm of gradients: 0.31360748964985674
l2 norm of weights: 2.939241218029929

Iteration Number: 2899
Loss: 22.756604222181622
l2 norm of gradients: 0.31346528116874034
l2 norm of weights: 2.9386436438309254

Iteration Number: 2900
Loss: 22.747645396504826
l2 norm of gradients: 0.31332313671233814
l2 norm of weights: 2.938046428040956

Iteration Number: 2901
Loss: 22.73869322790431
l2 norm of gradients: 0.31318105626342907
l2 norm of weights: 2.937449570543619

Iteration Number: 2902
Loss: 22.729747712072108
l2 norm of gradients: 0.31303903980482917
l2 norm of weights: 2.93685307122247

Iteration Number: 2903
Loss: 22.720808844702326
l2 norm of gradients: 0.31289708731939203
l2 norm of weights: 2.9362569299610293

Iteration Number: 2904
Loss: 22.71187662149113
l2 norm of gradients: 0.31275519879000835
l2 norm of weights: 2.935661146642779

Iteration Number: 2905
Loss: 22.702951038136742
l2 norm of gradients: 0.31261337419960583
l2 norm of weights: 2.9350657211511617

Iteration Number: 2906
Loss: 22.694032090339427
l2 norm of gradients: 0.3124716135311494

```

12 norm of weights: 2.934470653369583
-----
Iteration Number: 2907
Loss: 22.685119773801453
12 norm of gradients: 0.31232991676764066
12 norm of weights: 2.9338759431814108
-----
Iteration Number: 2908
Loss: 22.67621408422718
12 norm of gradients: 0.3121882838921183
12 norm of weights: 2.9332815904699734
-----
Iteration Number: 2909
Loss: 22.667315017323038
12 norm of gradients: 0.31204671488765773
12 norm of weights: 2.932687595118564
-----
Iteration Number: 2910
Loss: 22.65842256879739
12 norm of gradients: 0.31190520973737107
12 norm of weights: 2.932093957010435
-----
Iteration Number: 2911
Loss: 22.64953673436075
12 norm of gradients: 0.3117637684244072
12 norm of weights: 2.931500676028803
-----
Iteration Number: 2912
Loss: 22.64065750972562
12 norm of gradients: 0.3116223909319514
12 norm of weights: 2.9309077520568465
-----
Iteration Number: 2913
Loss: 22.631784890606514
12 norm of gradients: 0.31148107724322543
12 norm of weights: 2.9303151849777054
-----
Iteration Number: 2914
Loss: 22.622918872720007
12 norm of gradients: 0.31133982734148763
12 norm of weights: 2.9297229746744824
-----
Iteration Number: 2915
Loss: 22.614059451784737
12 norm of gradients: 0.31119864121003254
12 norm of weights: 2.9291311210302426
-----
Iteration Number: 2916

```

```

Loss: 22.60520662352128
l2 norm of gradients: 0.3110575188321908
l2 norm of weights: 2.928539623928013
-----
Iteration Number: 2917
Loss: 22.596360383652275
l2 norm of gradients: 0.31091646019132946
l2 norm of weights: 2.927948483250785
-----
Iteration Number: 2918
Loss: 22.587520727902486
l2 norm of gradients: 0.3107754652708516
l2 norm of weights: 2.927357698881509
-----
Iteration Number: 2919
Loss: 22.57868765199851
l2 norm of gradients: 0.31063453405419617
l2 norm of weights: 2.9267672707031003
-----
Iteration Number: 2920
Loss: 22.569861151669144
l2 norm of gradients: 0.3104936665248381
l2 norm of weights: 2.926177198598436
-----
Iteration Number: 2921
Loss: 22.56104122264509
l2 norm of gradients: 0.3103528626662883
l2 norm of weights: 2.925587482450356
-----
Iteration Number: 2922
Loss: 22.55222786065913
l2 norm of gradients: 0.31021212246209323
l2 norm of weights: 2.9249981221416634
-----
Iteration Number: 2923
Loss: 22.543421061446
l2 norm of gradients: 0.3100714458958352
l2 norm of weights: 2.924409117555122
-----
Iteration Number: 2924
Loss: 22.534620820742486
l2 norm of gradients: 0.309930832951132
l2 norm of weights: 2.9238204685734592
-----
Iteration Number: 2925
Loss: 22.52582713428742
l2 norm of gradients: 0.3097902836116369
l2 norm of weights: 2.923232175079366

```



```

-----
Iteration Number: 2926
Loss: 22.51703999782158
l2 norm of gradients: 0.3096497978610388
l2 norm of weights: 2.922644236955496
-----
Iteration Number: 2927
Loss: 22.508259407087746
l2 norm of gradients: 0.30950937568306186
l2 norm of weights: 2.9220566540844644
-----
Iteration Number: 2928
Loss: 22.499485357830775
l2 norm of gradients: 0.3093690170614654
l2 norm of weights: 2.92146942634885
-----
Iteration Number: 2929
Loss: 22.490717845797462
l2 norm of gradients: 0.3092287219800441
l2 norm of weights: 2.920882553631194
-----
Iteration Number: 2930
Loss: 22.481956866736645
l2 norm of gradients: 0.30908849042262765
l2 norm of weights: 2.920296035814001
-----
Iteration Number: 2931
Loss: 22.473202416399108
l2 norm of gradients: 0.3089483223730808
l2 norm of weights: 2.9197098727797384
-----
Iteration Number: 2932
Loss: 22.464454490537697
l2 norm of gradients: 0.30880821781530327
l2 norm of weights: 2.9191240644108363
-----
Iteration Number: 2933
Loss: 22.455713084907224
l2 norm of gradients: 0.30866817673322955
l2 norm of weights: 2.918538610589688
-----
Iteration Number: 2934
Loss: 22.44697819526447
l2 norm of gradients: 0.30852819911082896
l2 norm of weights: 2.9179535111986503
-----
Iteration Number: 2935
Loss: 22.438249817368224

```

```

12 norm of gradients: 0.30838828493210557
12 norm of weights: 2.9173687661200414
-----
Iteration Number: 2936
Loss: 22.42952794697932
12 norm of gradients: 0.308248434181098
12 norm of weights: 2.9167843752361455
-----
Iteration Number: 2937
Loss: 22.420812579860485
12 norm of gradients: 0.30810864684187944
12 norm of weights: 2.916200338429207
-----
Iteration Number: 2938
Loss: 22.412103711776503
12 norm of gradients: 0.3079689228985575
12 norm of weights: 2.915616655581435
-----
Iteration Number: 2939
Loss: 22.403401338494117
12 norm of gradients: 0.307829262335274
12 norm of weights: 2.9150333265750024
-----
Iteration Number: 2940
Loss: 22.39470545578204
12 norm of gradients: 0.3076896651362054
12 norm of weights: 2.914450351292044
-----
Iteration Number: 2941
Loss: 22.386016059411
12 norm of gradients: 0.307550131285562
12 norm of weights: 2.9138677296146582
-----
Iteration Number: 2942
Loss: 22.377333145153674
12 norm of gradients: 0.30741066076758844
12 norm of weights: 2.913285461424908
-----
Iteration Number: 2943
Loss: 22.368656708784716
12 norm of gradients: 0.30727125356656326
12 norm of weights: 2.912703546604819
-----
Iteration Number: 2944
Loss: 22.359986746080782
12 norm of gradients: 0.30713190966679893
12 norm of weights: 2.9121219850363786
-----

```

```

Iteration Number: 2945
Loss: 22.35132325282047
l2 norm of gradients: 0.30699262905264196
l2 norm of weights: 2.9115407766015404
-----
Iteration Number: 2946
Loss: 22.342666224784377
l2 norm of gradients: 0.3068534117084724
l2 norm of weights: 2.910959921182221
-----
Iteration Number: 2947
Loss: 22.33401565775503
l2 norm of gradients: 0.3067142576187041
l2 norm of weights: 2.9103794186602983
-----
Iteration Number: 2948
Loss: 22.325371547516983
l2 norm of gradients: 0.30657516676778446
l2 norm of weights: 2.909799268917616
-----
Iteration Number: 2949
Loss: 22.316733889856703
l2 norm of gradients: 0.3064361391401944
l2 norm of weights: 2.909219471835982
-----
Iteration Number: 2950
Loss: 22.308102680562644
l2 norm of gradients: 0.3062971747204485
l2 norm of weights: 2.908640027297165
-----
Iteration Number: 2951
Loss: 22.299477915425232
l2 norm of gradients: 0.3061582734930942
l2 norm of weights: 2.9080609351829003
-----
Iteration Number: 2952
Loss: 22.290859590236817
l2 norm of gradients: 0.3060194354427127
l2 norm of weights: 2.9074821953748855
-----
Iteration Number: 2953
Loss: 22.282247700791746
l2 norm of gradients: 0.305880660553918
l2 norm of weights: 2.9069038077547824
-----
Iteration Number: 2954
Loss: 22.273642242886307
l2 norm of gradients: 0.30574194881135747

```

```

12 norm of weights: 2.9063257722042164
-----
Iteration Number: 2955
Loss: 22.265043212318727
12 norm of gradients: 0.30560330019971116
12 norm of weights: 2.905748088604777
-----
Iteration Number: 2956
Loss: 22.256450604889217
12 norm of gradients: 0.3054647147036924
12 norm of weights: 2.905170756838017
-----
Iteration Number: 2957
Loss: 22.247864416399953
12 norm of gradients: 0.30532619230804714
12 norm of weights: 2.9045937767854553
-----
Iteration Number: 2958
Loss: 22.239284642654983
12 norm of gradients: 0.3051877329975541
12 norm of weights: 2.9040171483285717
-----
Iteration Number: 2959
Loss: 22.23071127946037
12 norm of gradients: 0.30504933675702445
12 norm of weights: 2.903440871348813
-----
Iteration Number: 2960
Loss: 22.22214432262412
12 norm of gradients: 0.3049110035713024
12 norm of weights: 2.902864945727588
-----
Iteration Number: 2961
Loss: 22.21358376795619
12 norm of gradients: 0.3047727334252641
12 norm of weights: 2.90228937134627
-----
Iteration Number: 2962
Loss: 22.20502961126839
12 norm of gradients: 0.30463452630381854
12 norm of weights: 2.9017141480861977
-----
Iteration Number: 2963
Loss: 22.196481848374596
12 norm of gradients: 0.30449638219190667
12 norm of weights: 2.9011392758286725
-----
Iteration Number: 2964

```

Loss: 22.187940475090553
l2 norm of gradients: 0.30435830107450196
l2 norm of weights: 2.9005647544549613

Iteration Number: 2965
Loss: 22.179405487233964
l2 norm of gradients: 0.30422028293660963
l2 norm of weights: 2.8999905838462943

Iteration Number: 2966
Loss: 22.17087688062444
l2 norm of gradients: 0.3040823277632673
l2 norm of weights: 2.899416763883867

Iteration Number: 2967
Loss: 22.16235465108356
l2 norm of gradients: 0.3039444355395443
l2 norm of weights: 2.8988432944448838

Iteration Number: 2968
Loss: 22.153838794434815
l2 norm of gradients: 0.30380660625054196
l2 norm of weights: 2.8982701754223315

Iteration Number: 2969
Loss: 22.14532930650364
l2 norm of gradients: 0.30366883988139326
l2 norm of weights: 2.8976974066854364

Iteration Number: 2970
Loss: 22.136826183117382
l2 norm of gradients: 0.30353113641726304
l2 norm of weights: 2.897124988119205

Iteration Number: 2971
Loss: 22.12832942010534
l2 norm of gradients: 0.30339349584334746
l2 norm of weights: 2.8965529196046544

Iteration Number: 2972
Loss: 22.119839013298694
l2 norm of gradients: 0.3032559181448745
l2 norm of weights: 2.8959812010227672

Iteration Number: 2973
Loss: 22.111354958530583
l2 norm of gradients: 0.3031184033071032
l2 norm of weights: 2.8954098322544897

```

-----
Iteration Number: 2974
Loss: 22.10287725163608
l2 norm of gradients: 0.30298095131532427
l2 norm of weights: 2.894838813180733
-----
Iteration Number: 2975
Loss: 22.094405888452137
l2 norm of gradients: 0.30284356215485947
l2 norm of weights: 2.894268143682374
-----
Iteration Number: 2976
Loss: 22.085940864817648
l2 norm of gradients: 0.3027062358110616
l2 norm of weights: 2.8936978236402524
-----
Iteration Number: 2977
Loss: 22.07748217657342
l2 norm of gradients: 0.30256897226931484
l2 norm of weights: 2.893127852935175
-----
Iteration Number: 2978
Loss: 22.06902981956215
l2 norm of gradients: 0.302431771515034
l2 norm of weights: 2.892558231447912
-----
Iteration Number: 2979
Loss: 22.06058378962851
l2 norm of gradients: 0.30229463353366487
l2 norm of weights: 2.8919889590591987
-----
Iteration Number: 2980
Loss: 22.052144082619037
l2 norm of gradients: 0.30215755831068414
l2 norm of weights: 2.8914200356497353
-----
Iteration Number: 2981
Loss: 22.043710694382145
l2 norm of gradients: 0.3020205458315991
l2 norm of weights: 2.890851461100188
-----
Iteration Number: 2982
Loss: 22.035283620768233
l2 norm of gradients: 0.30188359608194754
l2 norm of weights: 2.8902832352911862
-----
Iteration Number: 2983
Loss: 22.02686285762957

```

```

12 norm of gradients: 0.3017467090472979
12 norm of weights: 2.889715358103327
-----
Iteration Number: 2984
Loss: 22.01844840082032
12 norm of gradients: 0.30160988471324895
12 norm of weights: 2.8891478294171695
-----
Iteration Number: 2985
Loss: 22.010040246196528
12 norm of gradients: 0.3014731230654299
12 norm of weights: 2.8885806491132398
-----
Iteration Number: 2986
Loss: 22.0016383896162
12 norm of gradients: 0.30133642408950007
12 norm of weights: 2.88801381707203
-----
Iteration Number: 2987
Loss: 21.993242826939223
12 norm of gradients: 0.3011997877711489
12 norm of weights: 2.8874473331739954
-----
Iteration Number: 2988
Loss: 21.984853554027342
12 norm of gradients: 0.301063214096096
12 norm of weights: 2.8868811972995587
-----
Iteration Number: 2989
Loss: 21.97647056674422
12 norm of gradients: 0.3009267030500909
12 norm of weights: 2.886315409329106
-----
Iteration Number: 2990
Loss: 21.96809386095545
12 norm of gradients: 0.30079025461891284
12 norm of weights: 2.88574996914299
-----
Iteration Number: 2991
Loss: 21.959723432528452
12 norm of gradients: 0.3006538687883711
12 norm of weights: 2.8851848766215284
-----
Iteration Number: 2992
Loss: 21.951359277332596
12 norm of gradients: 0.3005175455443045
12 norm of weights: 2.8846201316450055
-----

```

Iteration Number: 2993
Loss: 21.943001391239108
l2 norm of gradients: 0.3003812848725812
l2 norm of weights: 2.884055734093669

Iteration Number: 2994
Loss: 21.9346497701211
l2 norm of gradients: 0.3002450867590994
l2 norm of weights: 2.883491683847734

Iteration Number: 2995
Loss: 21.926304409853568
l2 norm of gradients: 0.30010895118978625
l2 norm of weights: 2.88292798078738

Iteration Number: 2996
Loss: 21.917965306313455
l2 norm of gradients: 0.29997287815059837
l2 norm of weights: 2.8823646247927535

Iteration Number: 2997
Loss: 21.909632455379477
l2 norm of gradients: 0.2998368676275215
l2 norm of weights: 2.881801615743966

Iteration Number: 2998
Loss: 21.90130585293233
l2 norm of gradients: 0.29970091960657064
l2 norm of weights: 2.8812389535210943

Iteration Number: 2999
Loss: 21.89298549485454
l2 norm of gradients: 0.29956503407378965
l2 norm of weights: 2.8806766380041813

Iteration Number: 3000
Loss: 21.884671377030497
l2 norm of gradients: 0.29942921101525144
l2 norm of weights: 2.8801146690732367

Iteration Number: 3001
Loss: 21.876363495346514
l2 norm of gradients: 0.29929345041705774
l2 norm of weights: 2.879553046608235

Iteration Number: 3002
Loss: 21.86806184569073
l2 norm of gradients: 0.29915775226533897


```
12 norm of weights: 2.8789917704891166
-----
Iteration Number: 3003
Loss: 21.859766423953193
12 norm of gradients: 0.2990221165462542
12 norm of weights: 2.8784308405957884
-----
Iteration Number: 3004
Loss: 21.85147722602581
12 norm of gradients: 0.2988865432459911
12 norm of weights: 2.877870256808124
-----
Iteration Number: 3005
Loss: 21.84319424780236
12 norm of gradients: 0.2987510323507658
12 norm of weights: 2.877310019005961
-----
Iteration Number: 3006
Loss: 21.83491748517848
12 norm of gradients: 0.29861558384682263
12 norm of weights: 2.8767501270691054
-----
Iteration Number: 3007
Loss: 21.82664693405166
12 norm of gradients: 0.2984801977204344
12 norm of weights: 2.876190580877328
-----
Iteration Number: 3008
Loss: 21.81838259032131
12 norm of gradients: 0.29834487395790193
12 norm of weights: 2.875631380310366
-----
Iteration Number: 3009
Loss: 21.810124449888654
12 norm of gradients: 0.2982096125455542
12 norm of weights: 2.875072525247924
-----
Iteration Number: 3010
Loss: 21.801872508656782
12 norm of gradients: 0.29807441346974817
12 norm of weights: 2.8745140155696713
-----
Iteration Number: 3011
Loss: 21.793626762530646
12 norm of gradients: 0.29793927671686865
12 norm of weights: 2.873955851155244
-----
Iteration Number: 3012
```

Loss: 21.78538720741708
l2 norm of gradients: 0.29780420227332816
l2 norm of weights: 2.8733980318842463

Iteration Number: 3013
Loss: 21.777153839224738
l2 norm of gradients: 0.29766919012556714
l2 norm of weights: 2.872840557636246

Iteration Number: 3014
Loss: 21.768926653864163
l2 norm of gradients: 0.2975342402600533
l2 norm of weights: 2.8722834282907796

Iteration Number: 3015
Loss: 21.76070564724773
l2 norm of gradients: 0.29739935266328205
l2 norm of weights: 2.8717266437273494

Iteration Number: 3016
Loss: 21.752490815289665
l2 norm of gradients: 0.2972645273217762
l2 norm of weights: 2.8711702038254243

Iteration Number: 3017
Loss: 21.74428215390606
l2 norm of gradients: 0.2971297642220858
l2 norm of weights: 2.87061410846444

Iteration Number: 3018
Loss: 21.736079659014855
l2 norm of gradients: 0.29699506335078807
l2 norm of weights: 2.8700583575237983

Iteration Number: 3019
Loss: 21.727883326535817
l2 norm of gradients: 0.2968604246944875
l2 norm of weights: 2.8695029508828687

Iteration Number: 3020
Loss: 21.71969315239057
l2 norm of gradients: 0.2967258482398152
l2 norm of weights: 2.8689478884209874

Iteration Number: 3021
Loss: 21.71150913250258
l2 norm of gradients: 0.2965913339734296
l2 norm of weights: 2.8683931700174563

```

-----
Iteration Number: 3022
Loss: 21.70333126279715
l2 norm of gradients: 0.296456881882016
l2 norm of weights: 2.8678387955515454
-----
Iteration Number: 3023
Loss: 21.695159539201477
l2 norm of gradients: 0.296322491952286
l2 norm of weights: 2.8672847649024913
-----
Iteration Number: 3024
Loss: 21.686993957644493
l2 norm of gradients: 0.29618816417097804
l2 norm of weights: 2.866731077949497
-----
Iteration Number: 3025
Loss: 21.678834514057055
l2 norm of gradients: 0.2960538985248571
l2 norm of weights: 2.8661777345717336
-----
Iteration Number: 3026
Loss: 21.67068120437183
l2 norm of gradients: 0.2959196950007145
l2 norm of weights: 2.865624734648339
-----
Iteration Number: 3027
Loss: 21.662534024523303
l2 norm of gradients: 0.29578555358536796
l2 norm of weights: 2.865072078058417
-----
Iteration Number: 3028
Loss: 21.65439297044782
l2 norm of gradients: 0.2956514742656613
l2 norm of weights: 2.86451976468104
-----
Iteration Number: 3029
Loss: 21.64625803808356
l2 norm of gradients: 0.2955174570284646
l2 norm of weights: 2.8639677943952475
-----
Iteration Number: 3030
Loss: 21.638129223370477
l2 norm of gradients: 0.2953835018606739
l2 norm of weights: 2.863416167080046
-----
Iteration Number: 3031
Loss: 21.630006522250422

```

```

12 norm of gradients: 0.29524960874921113
12 norm of weights: 2.8628648826144087
-----
Iteration Number: 3032
Loss: 21.62188993066704
12 norm of gradients: 0.29511577768102415
12 norm of weights: 2.8623139408772773
-----
Iteration Number: 3033
Loss: 21.613779444565804
12 norm of gradients: 0.29498200864308627
12 norm of weights: 2.86176334174756
-----
Iteration Number: 3034
Loss: 21.605675059894015
12 norm of gradients: 0.2948483016223969
12 norm of weights: 2.8612130851041333
-----
Iteration Number: 3035
Loss: 21.5975767726008
12 norm of gradients: 0.2947146566059804
12 norm of weights: 2.8606631708258403
-----
Iteration Number: 3036
Loss: 21.58948457863709
12 norm of gradients: 0.2945810735808869
12 norm of weights: 2.8601135987914925
-----
Iteration Number: 3037
Loss: 21.581398473955645
12 norm of gradients: 0.29444755253419186
12 norm of weights: 2.859564368879869
-----
Iteration Number: 3038
Loss: 21.573318454511078
12 norm of gradients: 0.2943140934529959
12 norm of weights: 2.859015480969716
-----
Iteration Number: 3039
Loss: 21.56524451625975
12 norm of gradients: 0.2941806963244247
12 norm of weights: 2.8584669349397474
-----
Iteration Number: 3040
Loss: 21.557176655159882
12 norm of gradients: 0.29404736113562885
12 norm of weights: 2.8579187306686458
-----

```

Iteration Number: 3041
Loss: 21.549114867171504
l2 norm of gradients: 0.2939140878737842
l2 norm of weights: 2.8573708680350602

Iteration Number: 3042
Loss: 21.541059148256448
l2 norm of gradients: 0.2937808765260912
l2 norm of weights: 2.856823346917609

Iteration Number: 3043
Loss: 21.533009494378376
l2 norm of gradients: 0.293647727079775
l2 norm of weights: 2.8562761671948773

Iteration Number: 3044
Loss: 21.52496590150273
l2 norm of gradients: 0.29351463952208523
l2 norm of weights: 2.8557293287454195

Iteration Number: 3045
Loss: 21.51692836559677
l2 norm of gradients: 0.2933816138402965
l2 norm of weights: 2.8551828314477565

Iteration Number: 3046
Loss: 21.508896882629603
l2 norm of gradients: 0.2932486500217072
l2 norm of weights: 2.8546366751803784

Iteration Number: 3047
Loss: 21.500871448572063
l2 norm of gradients: 0.29311574805364055
l2 norm of weights: 2.8540908598217425

Iteration Number: 3048
Loss: 21.49285205939684
l2 norm of gradients: 0.29298290792344367
l2 norm of weights: 2.853545385250275

Iteration Number: 3049
Loss: 21.484838711078424
l2 norm of gradients: 0.2928501296184879
l2 norm of weights: 2.8530002513443704

Iteration Number: 3050
Loss: 21.476831399593088
l2 norm of gradients: 0.29271741312616845

```

12 norm of weights: 2.8524554579823906
-----
Iteration Number: 3051
Loss: 21.468830120918923
12 norm of gradients: 0.29258475843390463
12 norm of weights: 2.8519110050426675
-----
Iteration Number: 3052
Loss: 21.46083487103578
12 norm of gradients: 0.2924521655291394
12 norm of weights: 2.851366892403499
-----
Iteration Number: 3053
Loss: 21.45284564592538
12 norm of gradients: 0.2923196343993395
12 norm of weights: 2.850823119943154
-----
Iteration Number: 3054
Loss: 21.444862441571143
12 norm of gradients: 0.29218716503199504
12 norm of weights: 2.850279687539868
-----
Iteration Number: 3055
Loss: 21.43688525395835
12 norm of gradients: 0.29205475741461984
12 norm of weights: 2.849736595071846
-----
Iteration Number: 3056
Loss: 21.42891407907407
12 norm of gradients: 0.29192241153475096
12 norm of weights: 2.8491938424172605
-----
Iteration Number: 3057
Loss: 21.420948912907104
12 norm of gradients: 0.29179012737994886
12 norm of weights: 2.8486514294542546
-----
Iteration Number: 3058
Loss: 21.41298975144813
12 norm of gradients: 0.29165790493779703
12 norm of weights: 2.8481093560609376
-----
Iteration Number: 3059
Loss: 21.405036590689534
12 norm of gradients: 0.29152574419590216
12 norm of weights: 2.84756762211539
-----
Iteration Number: 3060

```

Loss: 21.397089426625538
l2 norm of gradients: 0.2913936451418937
l2 norm of weights: 2.8470262274956593

Iteration Number: 3061
Loss: 21.38914825525213
l2 norm of gradients: 0.29126160776342414
l2 norm of weights: 2.8464851720797624

Iteration Number: 3062
Loss: 21.381213072567075
l2 norm of gradients: 0.29112963204816866
l2 norm of weights: 2.8459444557456854

Iteration Number: 3063
Loss: 21.373283874569942
l2 norm of gradients: 0.2909977179838251
l2 norm of weights: 2.8454040783713834

Iteration Number: 3064
Loss: 21.365360657262055
l2 norm of gradients: 0.29086586555811367
l2 norm of weights: 2.8448640398347798

Iteration Number: 3065
Loss: 21.357443416646554
l2 norm of gradients: 0.2907340747587774
l2 norm of weights: 2.8443243400137677

Iteration Number: 3066
Loss: 21.349532148728297
l2 norm of gradients: 0.29060234557358106
l2 norm of weights: 2.8437849787862093

Iteration Number: 3067
Loss: 21.341626849513975
l2 norm of gradients: 0.29047067799031234
l2 norm of weights: 2.8432459560299352

Iteration Number: 3068
Loss: 21.333727515012022
l2 norm of gradients: 0.2903390719967805
l2 norm of weights: 2.8427072716227464

Iteration Number: 3069
Loss: 21.325834141232654
l2 norm of gradients: 0.2902075275808171
l2 norm of weights: 2.842168925442412

```

-----
Iteration Number: 3070
Loss: 21.317946724187873
l2 norm of gradients: 0.2900760447302754
l2 norm of weights: 2.8416309173666714
-----
Iteration Number: 3071
Loss: 21.310065259891438
l2 norm of gradients: 0.28994462343303057
l2 norm of weights: 2.841093247273233
-----
Iteration Number: 3072
Loss: 21.302189744358873
l2 norm of gradients: 0.28981326367697946
l2 norm of weights: 2.840555915039774
-----
Iteration Number: 3073
Loss: 21.294320173607474
l2 norm of gradients: 0.28968196545004055
l2 norm of weights: 2.840018920543942
-----
Iteration Number: 3074
Loss: 21.286456543656314
l2 norm of gradients: 0.28955072874015364
l2 norm of weights: 2.8394822636633537
-----
Iteration Number: 3075
Loss: 21.27859885052622
l2 norm of gradients: 0.2894195535352802
l2 norm of weights: 2.8389459442755958
-----
Iteration Number: 3076
Loss: 21.270747090239787
l2 norm of gradients: 0.2892884398234026
l2 norm of weights: 2.838409962258224
-----
Iteration Number: 3077
Loss: 21.262901258821383
l2 norm of gradients: 0.28915738759252463
l2 norm of weights: 2.837874317488764
-----
Iteration Number: 3078
Loss: 21.255061352297098
l2 norm of gradients: 0.28902639683067094
l2 norm of weights: 2.8373390098447118
-----
Iteration Number: 3079
Loss: 21.24722736669484

```



```

12 norm of gradients: 0.28889546752588724
12 norm of weights: 2.8368040392035314
-----
Iteration Number: 3080
Loss: 21.239399298044244
12 norm of gradients: 0.28876459966624013
12 norm of weights: 2.836269405442659
-----
Iteration Number: 3081
Loss: 21.231577142376704
12 norm of gradients: 0.2886337932398167
12 norm of weights: 2.8357351084394993
-----
Iteration Number: 3082
Loss: 21.223760895725356
12 norm of gradients: 0.2885030482347249
12 norm of weights: 2.835201148071427
-----
Iteration Number: 3083
Loss: 21.21595055412512
12 norm of gradients: 0.28837236463909305
12 norm of weights: 2.8346675242157877
-----
Iteration Number: 3084
Loss: 21.208146113612663
12 norm of gradients: 0.28824174244107
12 norm of weights: 2.834134236749896
-----
Iteration Number: 3085
Loss: 21.20034757022638
12 norm of gradients: 0.2881111816288246
12 norm of weights: 2.8336012855510373
-----
Iteration Number: 3086
Loss: 21.192554920006447
12 norm of gradients: 0.2879806821905462
12 norm of weights: 2.8330686704964667
-----
Iteration Number: 3087
Loss: 21.184768158994785
12 norm of gradients: 0.2878502441144441
12 norm of weights: 2.8325363914634103
-----
Iteration Number: 3088
Loss: 21.176987283235025
12 norm of gradients: 0.2877198673887475
12 norm of weights: 2.8320044483290636
-----

```

Iteration Number: 3089
Loss: 21.1692122887726
l2 norm of gradients: 0.28758955200170544
l2 norm of weights: 2.831472840970593

Iteration Number: 3090
Loss: 21.16144317165467
l2 norm of gradients: 0.2874592979415869
l2 norm of weights: 2.830941569265135

Iteration Number: 3091
Loss: 21.153679927930117
l2 norm of gradients: 0.2873291051966803
l2 norm of weights: 2.830410633089797

Iteration Number: 3092
Loss: 21.145922553649598
l2 norm of gradients: 0.28719897375529374
l2 norm of weights: 2.8298800323216566

Iteration Number: 3093
Loss: 21.138171044865487
l2 norm of gradients: 0.2870689036057546
l2 norm of weights: 2.829349766837762

Iteration Number: 3094
Loss: 21.130425397631914
l2 norm of gradients: 0.28693889473640943
l2 norm of weights: 2.8288198365151316

Iteration Number: 3095
Loss: 21.122685608004744
l2 norm of gradients: 0.2868089471356244
l2 norm of weights: 2.828290241230755

Iteration Number: 3096
Loss: 21.11495167204157
l2 norm of gradients: 0.2866790607917843
l2 norm of weights: 2.827760980861593

Iteration Number: 3097
Loss: 21.107223585801755
l2 norm of gradients: 0.2865492356932932
l2 norm of weights: 2.8272320552845756

Iteration Number: 3098
Loss: 21.099501345346358
l2 norm of gradients: 0.28641947182857397

```

12 norm of weights: 2.8267034643766054
-----
Iteration Number: 3099
Loss: 21.091784946738183
12 norm of gradients: 0.286289769186068
12 norm of weights: 2.8261752080145555
-----
Iteration Number: 3100
Loss: 21.084074386041788
12 norm of gradients: 0.2861601277542358
12 norm of weights: 2.825647286075269
-----
Iteration Number: 3101
Loss: 21.076369659323458
12 norm of gradients: 0.2860305475215559
12 norm of weights: 2.82511969843556
-----
Iteration Number: 3102
Loss: 21.068670762651163
12 norm of gradients: 0.28590102847652543
12 norm of weights: 2.8245924449722155
-----
Iteration Number: 3103
Loss: 21.060977692094685
12 norm of gradients: 0.28577157060766
12 norm of weights: 2.824065525561992
-----
Iteration Number: 3104
Loss: 21.053290443725448
12 norm of gradients: 0.2856421739034932
12 norm of weights: 2.823538940081618
-----
Iteration Number: 3105
Loss: 21.045609013616673
12 norm of gradients: 0.2855128383525767
12 norm of weights: 2.823012688407792
-----
Iteration Number: 3106
Loss: 21.037933397843265
12 norm of gradients: 0.28538356394348036
12 norm of weights: 2.8224867704171848
-----
Iteration Number: 3107
Loss: 21.03026359248186
12 norm of gradients: 0.28525435066479155
12 norm of weights: 2.821961185986439
-----
Iteration Number: 3108

```

```
Loss: 21.022599593610853
l2 norm of gradients: 0.28512519850511553
l2 norm of weights: 2.821435934992168
-----
Iteration Number: 3109
Loss: 21.014941397310295
l2 norm of gradients: 0.2849961074530755
l2 norm of weights: 2.8209110173109555
-----
Iteration Number: 3110
Loss: 21.007288999662034
l2 norm of gradients: 0.2848670774973117
l2 norm of weights: 2.820386432819359
-----
Iteration Number: 3111
Loss: 20.99964239674957
l2 norm of gradients: 0.284738108626482
l2 norm of weights: 2.819862181393906
-----
Iteration Number: 3112
Loss: 20.992001584658155
l2 norm of gradients: 0.28460920082926167
l2 norm of weights: 2.819338262911097
-----
Iteration Number: 3113
Loss: 20.98436655947479
l2 norm of gradients: 0.2844803540943428
l2 norm of weights: 2.818814677247402
-----
Iteration Number: 3114
Loss: 20.976737317288112
l2 norm of gradients: 0.28435156841043496
l2 norm of weights: 2.818291424279265
-----
Iteration Number: 3115
Loss: 20.96911385418855
l2 norm of gradients: 0.28422284376626444
l2 norm of weights: 2.8177685038831
-----
Iteration Number: 3116
Loss: 20.9614961662682
l2 norm of gradients: 0.2840941801505744
l2 norm of weights: 2.817245915935295
-----
Iteration Number: 3117
Loss: 20.953884249620906
l2 norm of gradients: 0.2839655775521248
l2 norm of weights: 2.8167236603122077
```

```

-----
Iteration Number: 3118
Loss: 20.9462781003422
l2 norm of gradients: 0.2838370359596921
l2 norm of weights: 2.8162017368901684
-----
Iteration Number: 3119
Loss: 20.93867771452932
l2 norm of gradients: 0.2837085553620693
l2 norm of weights: 2.815680145545481
-----
Iteration Number: 3120
Loss: 20.93108308828123
l2 norm of gradients: 0.2835801357480658
l2 norm of weights: 2.8151588861544194
-----
Iteration Number: 3121
Loss: 20.923494217698604
l2 norm of gradients: 0.2834517771065072
l2 norm of weights: 2.814637958593231
-----
Iteration Number: 3122
Loss: 20.915911098883807
l2 norm of gradients: 0.28332347942623537
l2 norm of weights: 2.8141173627381346
-----
Iteration Number: 3123
Loss: 20.908333727940928
l2 norm of gradients: 0.2831952426961082
l2 norm of weights: 2.8135970984653214
-----
Iteration Number: 3124
Loss: 20.900762100975765
l2 norm of gradients: 0.2830670669049993
l2 norm of weights: 2.8130771656509563
-----
Iteration Number: 3125
Loss: 20.89319621409577
l2 norm of gradients: 0.2829389520417984
l2 norm of weights: 2.8125575641711746
-----
Iteration Number: 3126
Loss: 20.885636063410175
l2 norm of gradients: 0.28281089809541066
l2 norm of weights: 2.812038293902085
-----
Iteration Number: 3127
Loss: 20.87808164502986

```

```

12 norm of gradients: 0.282682905054757
12 norm of weights: 2.8115193547197688
-----
Iteration Number: 3128
Loss: 20.870532955067407
12 norm of gradients: 0.2825549729087736
12 norm of weights: 2.8110007465002793
-----
Iteration Number: 3129
Loss: 20.862989989637125
12 norm of gradients: 0.2824271016464124
12 norm of weights: 2.8104824691196435
-----
Iteration Number: 3130
Loss: 20.85545274485501
12 norm of gradients: 0.28229929125664005
12 norm of weights: 2.8099645224538596
-----
Iteration Number: 3131
Loss: 20.84792121683875
12 norm of gradients: 0.2821715417284385
12 norm of weights: 2.8094469063789003
-----
Iteration Number: 3132
Loss: 20.840395401707745
12 norm of gradients: 0.28204385305080487
12 norm of weights: 2.8089296207707095
-----
Iteration Number: 3133
Loss: 20.83287529558305
12 norm of gradients: 0.28191622521275106
12 norm of weights: 2.8084126655052053
-----
Iteration Number: 3134
Loss: 20.825360894587455
12 norm of gradients: 0.2817886582033037
12 norm of weights: 2.807896040458277
-----
Iteration Number: 3135
Loss: 20.81785219484544
12 norm of gradients: 0.281661152011504
12 norm of weights: 2.8073797455057887
-----
Iteration Number: 3136
Loss: 20.81034919248317
12 norm of gradients: 0.28153370662640786
12 norm of weights: 2.8068637805235768
-----

```

Iteration Number: 3137
Loss: 20.802851883628474
l2 norm of gradients: 0.28140632203708565
l2 norm of weights: 2.806348145387451

Iteration Number: 3138
Loss: 20.795360264410917
l2 norm of gradients: 0.28127899823262176
l2 norm of weights: 2.805832839973193

Iteration Number: 3139
Loss: 20.787874330961728
l2 norm of gradients: 0.2811517352021149
l2 norm of weights: 2.8053178641565597

Iteration Number: 3140
Loss: 20.78039407941382
l2 norm of gradients: 0.2810245329346779
l2 norm of weights: 2.8048032178132796

Iteration Number: 3141
Loss: 20.77291950590181
l2 norm of gradients: 0.2808973914194376
l2 norm of weights: 2.804288900819056

Iteration Number: 3142
Loss: 20.765450606561977
l2 norm of gradients: 0.2807703106455344
l2 norm of weights: 2.803774913049564

Iteration Number: 3143
Loss: 20.757987377532324
l2 norm of gradients: 0.2806432906021229
l2 norm of weights: 2.8032612543804536

Iteration Number: 3144
Loss: 20.750529814952493
l2 norm of gradients: 0.2805163312783706
l2 norm of weights: 2.8027479246873477

Iteration Number: 3145
Loss: 20.74307791496384
l2 norm of gradients: 0.28038943266345917
l2 norm of weights: 2.8022349238458424

Iteration Number: 3146
Loss: 20.735631673709378
l2 norm of gradients: 0.28026259474658316

```

12 norm of weights: 2.8017222517315084
-----
Iteration Number: 3147
Loss: 20.728191087333837
12 norm of gradients: 0.28013581751695066
12 norm of weights: 2.801209908219889
-----
Iteration Number: 3148
Loss: 20.720756151983576
12 norm of gradients: 0.28000910096378273
12 norm of weights: 2.800697893186502
-----
Iteration Number: 3149
Loss: 20.71332686380667
12 norm of gradients: 0.2798824450763134
12 norm of weights: 2.8001862065068397
-----
Iteration Number: 3150
Loss: 20.705903218952884
12 norm of gradients: 0.2797558498437896
12 norm of weights: 2.7996748480563665
-----
Iteration Number: 3151
Loss: 20.69848521357362
12 norm of gradients: 0.2796293152554713
12 norm of weights: 2.799163817710522
-----
Iteration Number: 3152
Loss: 20.69107284382197
12 norm of gradients: 0.27950284130063074
12 norm of weights: 2.7986531153447194
-----
Iteration Number: 3153
Loss: 20.683666105852716
12 norm of gradients: 0.27937642796855305
12 norm of weights: 2.7981427408343467
-----
Iteration Number: 3154
Loss: 20.676264995822308
12 norm of gradients: 0.27925007524853535
12 norm of weights: 2.797632694054765
-----
Iteration Number: 3155
Loss: 20.66886950988886
12 norm of gradients: 0.2791237831298874
12 norm of weights: 2.7971229748813102
-----
Iteration Number: 3156

```


Loss: 20.661479644212154
l2 norm of gradients: 0.278997551601931
l2 norm of weights: 2.7966135831892927

Iteration Number: 3157
Loss: 20.654095394953657
l2 norm of gradients: 0.278871380654
l2 norm of weights: 2.796104518853997

Iteration Number: 3158
Loss: 20.646716758276504
l2 norm of gradients: 0.2787452702754403
l2 norm of weights: 2.7955957817506807

Iteration Number: 3159
Loss: 20.639343730345505
l2 norm of gradients: 0.2786192204556095
l2 norm of weights: 2.795087371754578

Iteration Number: 3160
Loss: 20.63197630732712
l2 norm of gradients: 0.2784932311838768
l2 norm of weights: 2.7945792887408962

Iteration Number: 3161
Loss: 20.624614485389472
l2 norm of gradients: 0.27836730244962327
l2 norm of weights: 2.7940715325848178

Iteration Number: 3162
Loss: 20.617258260702386
l2 norm of gradients: 0.2782414342422412
l2 norm of weights: 2.7935641031615

Iteration Number: 3163
Loss: 20.60990762943732
l2 norm of gradients: 0.2781156265511343
l2 norm of weights: 2.7930570003460744

Iteration Number: 3164
Loss: 20.602562587767395
l2 norm of gradients: 0.2779898793657174
l2 norm of weights: 2.792550224013647

Iteration Number: 3165
Loss: 20.595223131867442
l2 norm of gradients: 0.27786419267541673
l2 norm of weights: 2.7920437740392994

```

-----
Iteration Number: 3166
Loss: 20.58788925791389
l2 norm of gradients: 0.277738566469669
l2 norm of weights: 2.7915376502980878
-----
Iteration Number: 3167
Loss: 20.58056096208487
l2 norm of gradients: 0.2776130007379223
l2 norm of weights: 2.791031852665043
-----
Iteration Number: 3168
Loss: 20.573238240560162
l2 norm of gradients: 0.27748749546963497
l2 norm of weights: 2.790526381015171
-----
Iteration Number: 3169
Loss: 20.565921089521215
l2 norm of gradients: 0.2773620506542765
l2 norm of weights: 2.790021235223453
-----
Iteration Number: 3170
Loss: 20.558609505151132
l2 norm of gradients: 0.27723666628132637
l2 norm of weights: 2.7895164151648464
-----
Iteration Number: 3171
Loss: 20.551303483634662
l2 norm of gradients: 0.27711134234027474
l2 norm of weights: 2.789011920714281
-----
Iteration Number: 3172
Loss: 20.544003021158225
l2 norm of gradients: 0.2769860788206218
l2 norm of weights: 2.7885077517466654
-----
Iteration Number: 3173
Loss: 20.536708113909896
l2 norm of gradients: 0.2768608757118783
l2 norm of weights: 2.78800390813688
-----
Iteration Number: 3174
Loss: 20.529418758079416
l2 norm of gradients: 0.2767357330035643
l2 norm of weights: 2.787500389759784
-----
Iteration Number: 3175
Loss: 20.522134949858156

```

```

12 norm of gradients: 0.27661065068521046
12 norm of weights: 2.7869971964902094
-----
Iteration Number: 3176
Loss: 20.514856685439156
12 norm of gradients: 0.2764856287463568
12 norm of weights: 2.786494328202965
-----
Iteration Number: 3177
Loss: 20.507583961017115
12 norm of gradients: 0.27636066717655305
12 norm of weights: 2.785991784772835
-----
Iteration Number: 3178
Loss: 20.500316772788366
12 norm of gradients: 0.27623576596535865
12 norm of weights: 2.785489566074579
-----
Iteration Number: 3179
Loss: 20.493055116950917
12 norm of gradients: 0.27611092510234203
12 norm of weights: 2.7849876719829325
-----
Iteration Number: 3180
Loss: 20.485798989704403
12 norm of gradients: 0.27598614457708154
12 norm of weights: 2.784486102372607
-----
Iteration Number: 3181
Loss: 20.478548387250125
12 norm of gradients: 0.27586142437916406
12 norm of weights: 2.78398485711829
-----
Iteration Number: 3182
Loss: 20.47130330579102
12 norm of gradients: 0.27573676449818585
12 norm of weights: 2.7834839360946435
-----
Iteration Number: 3183
Loss: 20.464063741531703
12 norm of gradients: 0.27561216492375196
12 norm of weights: 2.7829833391763072
-----
Iteration Number: 3184
Loss: 20.456829690678394
12 norm of gradients: 0.27548762564547635
12 norm of weights: 2.782483066237897
-----

```

Iteration Number: 3185
Loss: 20.449601149438998
l2 norm of gradients: 0.27536314665298156
l2 norm of weights: 2.781983117154002

Iteration Number: 3186
Loss: 20.44237811402302
l2 norm of gradients: 0.2752387279358987
l2 norm of weights: 2.7814834917991913

Iteration Number: 3187
Loss: 20.435160580641664
l2 norm of gradients: 0.27511436948386736
l2 norm of weights: 2.780984190048008

Iteration Number: 3188
Loss: 20.427948545507743
l2 norm of gradients: 0.2749900712865354
l2 norm of weights: 2.7804852117749723

Iteration Number: 3189
Loss: 20.42074200483572
l2 norm of gradients: 0.27486583333355874
l2 norm of weights: 2.77998655685458

Iteration Number: 3190
Loss: 20.413540954841704
l2 norm of gradients: 0.2747416556146016
l2 norm of weights: 2.7794882251613036

Iteration Number: 3191
Loss: 20.406345391743454
l2 norm of gradients: 0.27461753811933604
l2 norm of weights: 2.778990216569593

Iteration Number: 3192
Loss: 20.399155311760353
l2 norm of gradients: 0.2744934808374418
l2 norm of weights: 2.7784925309538737

Iteration Number: 3193
Loss: 20.391970711113427
l2 norm of gradients: 0.2743694837586065
l2 norm of weights: 2.777995168188548

Iteration Number: 3194
Loss: 20.38479158602536
l2 norm of gradients: 0.2742455468725252

```

12 norm of weights: 2.7774981281479953
-----
Iteration Number: 3195
Loss: 20.377617932720455
12 norm of gradients: 0.27412167016890054
12 norm of weights: 2.7770014107065712
-----
Iteration Number: 3196
Loss: 20.370449747424676
12 norm of gradients: 0.27399785363744233
12 norm of weights: 2.776505015738608
-----
Iteration Number: 3197
Loss: 20.363287026365583
12 norm of gradients: 0.2738740972678677
12 norm of weights: 2.7760089431184163
-----
Iteration Number: 3198
Loss: 20.356129765772412
12 norm of gradients: 0.27375040104990084
12 norm of weights: 2.7755131927202816
-----
Iteration Number: 3199
Loss: 20.34897796187603
12 norm of gradients: 0.2736267649732729
12 norm of weights: 2.7750177644184686
-----
Iteration Number: 3200
Loss: 20.34183161090891
12 norm of gradients: 0.2735031890277217
12 norm of weights: 2.7745226580872173
-----
Iteration Number: 3201
Loss: 20.3346907091052
12 norm of gradients: 0.27337967320299206
12 norm of weights: 2.7740278736007453
-----
Iteration Number: 3202
Loss: 20.32755525270065
12 norm of gradients: 0.27325621748883505
12 norm of weights: 2.7735334108332483
-----
Iteration Number: 3203
Loss: 20.32042523793265
12 norm of gradients: 0.2731328218750086
12 norm of weights: 2.773039269658898
-----
Iteration Number: 3204

```

```

Loss: 20.313300661040245
l2 norm of gradients: 0.2730094863512766
l2 norm of weights: 2.772545449951845
-----
Iteration Number: 3205
Loss: 20.306181518264083
l2 norm of gradients: 0.27288621090740933
l2 norm of weights: 2.772051951586216
-----
Iteration Number: 3206
Loss: 20.29906780584646
l2 norm of gradients: 0.2727629955331833
l2 norm of weights: 2.771558774436115
-----
Iteration Number: 3207
Loss: 20.29195952003128
l2 norm of gradients: 0.2726398402183806
l2 norm of weights: 2.7710659183756254
-----
Iteration Number: 3208
Loss: 20.28485665706413
l2 norm of gradients: 0.27251674495278966
l2 norm of weights: 2.770573383278806
-----
Iteration Number: 3209
Loss: 20.27775921319215
l2 norm of gradients: 0.27239370972620414
l2 norm of weights: 2.770081169019696
-----
Iteration Number: 3210
Loss: 20.27066718466416
l2 norm of gradients: 0.27227073452842365
l2 norm of weights: 2.7695892754723084
-----
Iteration Number: 3211
Loss: 20.263580567730603
l2 norm of gradients: 0.27214781934925314
l2 norm of weights: 2.769097702510638
-----
Iteration Number: 3212
Loss: 20.256499358643538
l2 norm of gradients: 0.2720249641785029
l2 norm of weights: 2.768606450008656
-----
Iteration Number: 3213
Loss: 20.249423553656655
l2 norm of gradients: 0.2719021690059885
l2 norm of weights: 2.76811551784031

```

```
-----
Iteration Number: 3214
Loss: 20.242353149025266
l2 norm of gradients: 0.2717794338215306
l2 norm of weights: 2.7676249058795284
-----
Iteration Number: 3215
Loss: 20.235288141006308
l2 norm of gradients: 0.27165675861495464
l2 norm of weights: 2.767134614000215
-----
Iteration Number: 3216
Loss: 20.22822852585835
l2 norm of gradients: 0.2715341433760912
l2 norm of weights: 2.7666446420762547
-----
Iteration Number: 3217
Loss: 20.221174299841582
l2 norm of gradients: 0.2714115880947756
l2 norm of weights: 2.766154989981508
-----
Iteration Number: 3218
Loss: 20.214125459217794
l2 norm of gradients: 0.27128909276084745
l2 norm of weights: 2.7656656575898153
-----
Iteration Number: 3219
Loss: 20.20708200025045
l2 norm of gradients: 0.27116665736415124
l2 norm of weights: 2.765176644774995
-----
Iteration Number: 3220
Loss: 20.20004391920459
l2 norm of gradients: 0.27104428189453533
l2 norm of weights: 2.7646879514108433
-----
Iteration Number: 3221
Loss: 20.193011212346885
l2 norm of gradients: 0.2709219663418528
l2 norm of weights: 2.7641995773711354
-----
Iteration Number: 3222
Loss: 20.18598387594564
l2 norm of gradients: 0.2707997106959606
l2 norm of weights: 2.763711522529625
-----
Iteration Number: 3223
Loss: 20.178961906270786
```

```

12 norm of gradients: 0.27067751494671966
12 norm of weights: 2.7632237867600455
-----
Iteration Number: 3224
Loss: 20.17194529959384
12 norm of gradients: 0.2705553790839948
12 norm of weights: 2.7627363699361074
-----
Iteration Number: 3225
Loss: 20.16493405218797
12 norm of gradients: 0.27043330309765456
12 norm of weights: 2.7622492719315006
-----
Iteration Number: 3226
Loss: 20.15792816032795
12 norm of gradients: 0.27031128697757123
12 norm of weights: 2.761762492619894
-----
Iteration Number: 3227
Loss: 20.150927620290165
12 norm of gradients: 0.27018933071362045
12 norm of weights: 2.761276031874935
-----
Iteration Number: 3228
Loss: 20.143932428352635
12 norm of gradients: 0.27006743429568114
12 norm of weights: 2.760789889570251
-----
Iteration Number: 3229
Loss: 20.136942580794994
12 norm of gradients: 0.2699455977136357
12 norm of weights: 2.7603040655794473
-----
Iteration Number: 3230
Loss: 20.12995807389845
12 norm of gradients: 0.2698238209573693
12 norm of weights: 2.7598185597761096
-----
Iteration Number: 3231
Loss: 20.122978903945917
12 norm of gradients: 0.2697021040167705
12 norm of weights: 2.7593333720338005
-----
Iteration Number: 3232
Loss: 20.116005067221824
12 norm of gradients: 0.26958044688173044
12 norm of weights: 2.758848502226065
-----

```


Iteration Number: 3233
Loss: 20.10903656001228
l2 norm of gradients: 0.26945884954214316
l2 norm of weights: 2.758363950226425

Iteration Number: 3234
Loss: 20.102073378604977
l2 norm of gradients: 0.2693373119879052
l2 norm of weights: 2.7578797159083837

Iteration Number: 3235
Loss: 20.09511551928925
l2 norm of gradients: 0.26921583420891565
l2 norm of weights: 2.757395799145421

Iteration Number: 3236
Loss: 20.088162978356014
l2 norm of gradients: 0.26909441619507596
l2 norm of weights: 2.7569121998109996

Iteration Number: 3237
Loss: 20.0812157520978
l2 norm of gradients: 0.2689730579362899
l2 norm of weights: 2.7564289177785595

Iteration Number: 3238
Loss: 20.074273836808796
l2 norm of gradients: 0.26885175942246314
l2 norm of weights: 2.7559459529215213

Iteration Number: 3239
Loss: 20.06733722878474
l2 norm of gradients: 0.26873052064350356
l2 norm of weights: 2.755463305113286

Iteration Number: 3240
Loss: 20.06040592432301
l2 norm of gradients: 0.26860934158932087
l2 norm of weights: 2.754980974227233

Iteration Number: 3241
Loss: 20.0534799197226
l2 norm of gradients: 0.26848822224982644
l2 norm of weights: 2.754498960136722

Iteration Number: 3242
Loss: 20.046559211284112
l2 norm of gradients: 0.2683671626149334

```

12 norm of weights: 2.7540172627150943
-----
Iteration Number: 3243
Loss: 20.039643795309743
12 norm of gradients: 0.26824616267455614
12 norm of weights: 2.7535358818356683
-----
Iteration Number: 3244
Loss: 20.032733668103315
12 norm of gradients: 0.26812522241861075
12 norm of weights: 2.753054817371745
-----
Iteration Number: 3245
Loss: 20.025828825970247
12 norm of gradients: 0.2680043418370143
12 norm of weights: 2.752574069196605
-----
Iteration Number: 3246
Loss: 20.018929265217572
12 norm of gradients: 0.26788352091968515
12 norm of weights: 2.752093637183508
-----
Iteration Number: 3247
Loss: 20.012034982153935
12 norm of gradients: 0.2677627596565425
12 norm of weights: 2.751613521205696
-----
Iteration Number: 3248
Loss: 20.00514597308958
12 norm of gradients: 0.2676420580375065
12 norm of weights: 2.751133721136389
-----
Iteration Number: 3249
Loss: 19.998262234336355
12 norm of gradients: 0.26752141605249824
12 norm of weights: 2.75065423684879
-----
Iteration Number: 3250
Loss: 19.99138376220774
12 norm of gradients: 0.26740083369143913
12 norm of weights: 2.7501750682160804
-----
Iteration Number: 3251
Loss: 19.98451055301879
12 norm of gradients: 0.26728031094425114
12 norm of weights: 2.7496962151114235
-----
Iteration Number: 3252

```

Loss: 19.977642603086174
l2 norm of gradients: 0.2671598478008569
l2 norm of weights: 2.7492176774079624

Iteration Number: 3253
Loss: 19.970779908728165
l2 norm of gradients: 0.2670394442511789
l2 norm of weights: 2.748739454978822

Iteration Number: 3254
Loss: 19.96392246626465
l2 norm of gradients: 0.26691910028514015
l2 norm of weights: 2.748261547697108

Iteration Number: 3255
Loss: 19.957070272017113
l2 norm of gradients: 0.26679881589266335
l2 norm of weights: 2.7477839554359047

Iteration Number: 3256
Loss: 19.950223322308652
l2 norm of gradients: 0.26667859106367114
l2 norm of weights: 2.7473066780682807

Iteration Number: 3257
Loss: 19.94338161346396
l2 norm of gradients: 0.2665584257880861
l2 norm of weights: 2.746829715467283

Iteration Number: 3258
Loss: 19.93654514180931
l2 norm of gradients: 0.2664383200558303
l2 norm of weights: 2.7463530675059413

Iteration Number: 3259
Loss: 19.929713903672642
l2 norm of gradients: 0.26631827385682527
l2 norm of weights: 2.745876734057266

Iteration Number: 3260
Loss: 19.922887895383415
l2 norm of gradients: 0.26619828718099214
l2 norm of weights: 2.7454007149942488

Iteration Number: 3261
Loss: 19.916067113272753
l2 norm of gradients: 0.266078360018251
l2 norm of weights: 2.7449250101898617

```

-----
Iteration Number: 3262
Loss: 19.90925155367336
l2 norm of gradients: 0.26595849235852137
l2 norm of weights: 2.74444961951706
-----
Iteration Number: 3263
Loss: 19.902441212919534
l2 norm of gradients: 0.2658386841917217
l2 norm of weights: 2.7439745428487785
-----
Iteration Number: 3264
Loss: 19.895636087347192
l2 norm of gradients: 0.26571893550776915
l2 norm of weights: 2.743499780057936
-----
Iteration Number: 3265
Loss: 19.888836173293832
l2 norm of gradients: 0.26559924629657994
l2 norm of weights: 2.7430253310174293
-----
Iteration Number: 3266
Loss: 19.882041467098563
l2 norm of gradients: 0.2654796165480686
l2 norm of weights: 2.7425511956001407
-----
Iteration Number: 3267
Loss: 19.875251965102102
l2 norm of gradients: 0.2653600462521486
l2 norm of weights: 2.742077373678932
-----
Iteration Number: 3268
Loss: 19.868467663646737
l2 norm of gradients: 0.2652405353987314
l2 norm of weights: 2.741603865126647
-----
Iteration Number: 3269
Loss: 19.86168855907638
l2 norm of gradients: 0.265121083977727
l2 norm of weights: 2.741130669816113
-----
Iteration Number: 3270
Loss: 19.85491464773654
l2 norm of gradients: 0.26500169197904344
l2 norm of weights: 2.7406577876201363
-----
Iteration Number: 3271
Loss: 19.848145925974304

```

```

12 norm of gradients: 0.2648823593925868
12 norm of weights: 2.7401852184115083
-----
Iteration Number: 3272
Loss: 19.84138239013839
12 norm of gradients: 0.2647630862082611
12 norm of weights: 2.739712962063
-----
Iteration Number: 3273
Loss: 19.83462403657909
12 norm of gradients: 0.26464387241596815
12 norm of weights: 2.739241018447367
-----
Iteration Number: 3274
Loss: 19.827870861648297
12 norm of gradients: 0.2645247180056074
12 norm of weights: 2.7387693874373458
-----
Iteration Number: 3275
Loss: 19.8211228616995
12 norm of gradients: 0.2644056229670757
12 norm of weights: 2.738298068905654
-----
Iteration Number: 3276
Loss: 19.814380033087787
12 norm of gradients: 0.26428658729026755
12 norm of weights: 2.7378270627249948
-----
Iteration Number: 3277
Loss: 19.807642372169852
12 norm of gradients: 0.26416761096507463
12 norm of weights: 2.737356368768051
-----
Iteration Number: 3278
Loss: 19.800909875303976
12 norm of gradients: 0.2640486939813857
12 norm of weights: 2.736885986907489
-----
Iteration Number: 3279
Loss: 19.79418253885004
12 norm of gradients: 0.2639298363290868
12 norm of weights: 2.7364159170159583
-----
Iteration Number: 3280
Loss: 19.787460359169508
12 norm of gradients: 0.2638110379980605
12 norm of weights: 2.7359461589660907
-----

```

```

Iteration Number: 3281
Loss: 19.780743332625455
l2 norm of gradients: 0.2636922989781867
l2 norm of weights: 2.7354767126305
-----
Iteration Number: 3282
Loss: 19.77403145558253
l2 norm of gradients: 0.26357361925934153
l2 norm of weights: 2.7350075778817846
-----
Iteration Number: 3283
Loss: 19.767324724407025
l2 norm of gradients: 0.2634549988313979
l2 norm of weights: 2.734538754592524
-----
Iteration Number: 3284
Loss: 19.760623135466766
l2 norm of gradients: 0.26333643768422516
l2 norm of weights: 2.734070242635281
-----
Iteration Number: 3285
Loss: 19.75392668513121
l2 norm of gradients: 0.2632179358076888
l2 norm of weights: 2.7336020418826035
-----
Iteration Number: 3286
Loss: 19.747235369771406
l2 norm of gradients: 0.26309949319165055
l2 norm of weights: 2.73313415220702
-----
Iteration Number: 3287
Loss: 19.740549185759978
l2 norm of gradients: 0.2629811098259685
l2 norm of weights: 2.7326665734810427
-----
Iteration Number: 3288
Loss: 19.73386812947116
l2 norm of gradients: 0.2628627857004962
l2 norm of weights: 2.732199305577168
-----
Iteration Number: 3289
Loss: 19.727192197280782
l2 norm of gradients: 0.2627445208050834
l2 norm of weights: 2.7317323483678755
-----
Iteration Number: 3290
Loss: 19.720521385566254
l2 norm of gradients: 0.2626263151295755

```

```

l2 norm of weights: 2.7312657017256274
-----
Iteration Number: 3291
Loss: 19.713855690706595
l2 norm of gradients: 0.2625081686638134
l2 norm of weights: 2.7307993655228704
-----
Iteration Number: 3292
Loss: 19.7071951090824
l2 norm of gradients: 0.2623900813976333
l2 norm of weights: 2.730333339632035
-----
Iteration Number: 3293
Loss: 19.700539637075856
l2 norm of gradients: 0.26227205332086706
l2 norm of weights: 2.729867623925533
-----
Iteration Number: 3294
Loss: 19.693889271070766
l2 norm of gradients: 0.2621540844233416
l2 norm of weights: 2.7294022182757627
-----
Iteration Number: 3295
Loss: 19.6872440074525
l2 norm of gradients: 0.26203617469487894
l2 norm of weights: 2.728937122555105
-----
Iteration Number: 3296
Loss: 19.680603842608047
l2 norm of gradients: 0.2619183241252961
l2 norm of weights: 2.7284723366359245
-----
Iteration Number: 3297
Loss: 19.67396877292595
l2 norm of gradients: 0.26180053270440484
l2 norm of weights: 2.72800786039057
-----
Iteration Number: 3298
Loss: 19.66733879479637
l2 norm of gradients: 0.261682800422012
l2 norm of weights: 2.7275436936913753
-----
Iteration Number: 3299
Loss: 19.660713904611058
l2 norm of gradients: 0.26156512726791864
l2 norm of weights: 2.727079836410656
-----
Iteration Number: 3300

```

```

Loss: 19.654094098763363
l2 norm of gradients: 0.26144751323192045
l2 norm of weights: 2.7266162884207144
-----
Iteration Number: 3301
Loss: 19.647479373648196
l2 norm of gradients: 0.2613299583038076
l2 norm of weights: 2.726153049593835
-----
Iteration Number: 3302
Loss: 19.64086972566208
l2 norm of gradients: 0.2612124624733644
l2 norm of weights: 2.7256901198022874
-----
Iteration Number: 3303
Loss: 19.63426515120313
l2 norm of gradients: 0.2610950257303693
l2 norm of weights: 2.725227498918327
-----
Iteration Number: 3304
Loss: 19.627665646671062
l2 norm of gradients: 0.26097764806459484
l2 norm of weights: 2.7247651868141904
-----
Iteration Number: 3305
Loss: 19.62107120846715
l2 norm of gradients: 0.2608603294658074
l2 norm of weights: 2.7243031833621023
-----
Iteration Number: 3306
Loss: 19.614481832994276
l2 norm of gradients: 0.26074306992376717
l2 norm of weights: 2.7238414884342697
-----
Iteration Number: 3307
Loss: 19.607897516656934
l2 norm of gradients: 0.2606258694282279
l2 norm of weights: 2.723380101902885
-----
Iteration Number: 3308
Loss: 19.601318255861166
l2 norm of gradients: 0.2605087279689371
l2 norm of weights: 2.722919023640125
-----
Iteration Number: 3309
Loss: 19.594744047014643
l2 norm of gradients: 0.26039164553563543
l2 norm of weights: 2.7224582535181527

```



```

-----
Iteration Number: 3310
Loss: 19.588174886526605
l2 norm of gradients: 0.260274622118057
l2 norm of weights: 2.7219977914091142
-----
Iteration Number: 3311
Loss: 19.58161077080787
l2 norm of gradients: 0.2601576577059292
l2 norm of weights: 2.7215376371851416
-----
Iteration Number: 3312
Loss: 19.575051696270876
l2 norm of gradients: 0.26004075228897217
l2 norm of weights: 2.7210777907183523
-----
Iteration Number: 3313
Loss: 19.568497659329637
l2 norm of gradients: 0.25992390585689934
l2 norm of weights: 2.7206182518808477
-----
Iteration Number: 3314
Loss: 19.561948656399746
l2 norm of gradients: 0.2598071183994167
l2 norm of weights: 2.720159020544715
-----
Iteration Number: 3315
Loss: 19.555404683898402
l2 norm of gradients: 0.25969038990622323
l2 norm of weights: 2.7197000965820273
-----
Iteration Number: 3316
Loss: 19.548865738244384
l2 norm of gradients: 0.2595737203670102
l2 norm of weights: 2.719241479864843
-----
Iteration Number: 3317
Loss: 19.542331815858066
l2 norm of gradients: 0.2594571097714615
l2 norm of weights: 2.7187831702652043
-----
Iteration Number: 3318
Loss: 19.535802913161408
l2 norm of gradients: 0.2593405581092535
l2 norm of weights: 2.718325167655141
-----
Iteration Number: 3319
Loss: 19.529279026577957

```

```

12 norm of gradients: 0.25922406537005455
12 norm of weights: 2.7178674719066676
-----
Iteration Number: 3320
Loss: 19.522760152532847
12 norm of gradients: 0.25910763154352523
12 norm of weights: 2.7174100828917838
-----
Iteration Number: 3321
Loss: 19.516246287452812
12 norm of gradients: 0.2589912566193182
12 norm of weights: 2.7169530004824756
-----
Iteration Number: 3322
Loss: 19.50973742776617
12 norm of gradients: 0.25887494058707805
12 norm of weights: 2.7164962245507143
-----
Iteration Number: 3323
Loss: 19.503233569902807
12 norm of gradients: 0.25875868343644093
12 norm of weights: 2.7160397549684583
-----
Iteration Number: 3324
Loss: 19.49673471029424
12 norm of gradients: 0.2586424851570348
12 norm of weights: 2.7155835916076505
-----
Iteration Number: 3325
Loss: 19.490240845373545
12 norm of gradients: 0.25852634573847905
12 norm of weights: 2.7151277343402214
-----
Iteration Number: 3326
Loss: 19.483751971575387
12 norm of gradients: 0.25841026517038473
12 norm of weights: 2.7146721830380858
-----
Iteration Number: 3327
Loss: 19.47726808533604
12 norm of gradients: 0.2582942434423539
12 norm of weights: 2.7142169375731453
-----
Iteration Number: 3328
Loss: 19.47078918309333
12 norm of gradients: 0.2581782805439801
12 norm of weights: 2.7137619978172896
-----

```

Iteration Number: 3329
Loss: 19.464315261286714
l2 norm of gradients: 0.2580623764648476
l2 norm of weights: 2.7133073636423912

Iteration Number: 3330
Loss: 19.45784631635721
l2 norm of gradients: 0.25794653119453187
l2 norm of weights: 2.7128530349203124

Iteration Number: 3331
Loss: 19.451382344747447
l2 norm of gradients: 0.2578307447225993
l2 norm of weights: 2.7123990115229004

Iteration Number: 3332
Loss: 19.44492334290161
l2 norm of gradients: 0.2577150170386068
l2 norm of weights: 2.711945293321988

Iteration Number: 3333
Loss: 19.4384693072655
l2 norm of gradients: 0.25759934813210206
l2 norm of weights: 2.7114918801893975

Iteration Number: 3334
Loss: 19.432020234286504
l2 norm of gradients: 0.2574837379926232
l2 norm of weights: 2.711038771996935

Iteration Number: 3335
Loss: 19.425576120413584
l2 norm of gradients: 0.25736818660969857
l2 norm of weights: 2.710585968616395

Iteration Number: 3336
Loss: 19.419136962097298
l2 norm of gradients: 0.25725269397284717
l2 norm of weights: 2.7101334699195583

Iteration Number: 3337
Loss: 19.412702755789798
l2 norm of gradients: 0.2571372600715779
l2 norm of weights: 2.7096812757781925

Iteration Number: 3338
Loss: 19.406273497944824
l2 norm of gradients: 0.2570218848953897

```

12 norm of weights: 2.7092293860640533
-----
Iteration Number: 3339
Loss: 19.39984918501769
12 norm of gradients: 0.2569065684337716
12 norm of weights: 2.7087778006488823
-----
Iteration Number: 3340
Loss: 19.393429813465318
12 norm of gradients: 0.25679131067620226
12 norm of weights: 2.7083265194044084
-----
Iteration Number: 3341
Loss: 19.387015379746202
12 norm of gradients: 0.2566761116121502
12 norm of weights: 2.7078755422023484
-----
Iteration Number: 3342
Loss: 19.38060588032045
12 norm of gradients: 0.25656097123107346
12 norm of weights: 2.7074248689144054
-----
Iteration Number: 3343
Loss: 19.374201311649728
12 norm of gradients: 0.2564458895224197
12 norm of weights: 2.7069744994122713
-----
Iteration Number: 3344
Loss: 19.367801670197313
12 norm of gradients: 0.2563308664756257
12 norm of weights: 2.706524433567624
-----
Iteration Number: 3345
Loss: 19.361406952428055
12 norm of gradients: 0.25621590208011763
12 norm of weights: 2.70607467125213
-----
Iteration Number: 3346
Loss: 19.35501715480842
12 norm of gradients: 0.2561009963253109
12 norm of weights: 2.7056252123374427
-----
Iteration Number: 3347
Loss: 19.348632273806423
12 norm of gradients: 0.25598614920060986
12 norm of weights: 2.7051760566952043
-----
Iteration Number: 3348

```

Loss: 19.34225230589171
l2 norm of gradients: 0.2558713606954078
l2 norm of weights: 2.704727204197043

Iteration Number: 3349
Loss: 19.335877247535482
l2 norm of gradients: 0.2557566307990867
l2 norm of weights: 2.7042786547145763

Iteration Number: 3350
Loss: 19.329507095210552
l2 norm of gradients: 0.2556419595010176
l2 norm of weights: 2.7038304081194093

Iteration Number: 3351
Loss: 19.32314184539131
l2 norm of gradients: 0.2555273467905598
l2 norm of weights: 2.703382464283135

Iteration Number: 3352
Loss: 19.316781494553748
l2 norm of gradients: 0.2554127926570611
l2 norm of weights: 2.7029348230773342

Iteration Number: 3353
Loss: 19.31042603917543
l2 norm of gradients: 0.2552982970898579
l2 norm of weights: 2.702487484373576

Iteration Number: 3354
Loss: 19.304075475735527
l2 norm of gradients: 0.2551838600782746
l2 norm of weights: 2.7020404480434177

Iteration Number: 3355
Loss: 19.297729800714787
l2 norm of gradients: 0.2550694816116239
l2 norm of weights: 2.7015937139584056

Iteration Number: 3356
Loss: 19.291389010595562
l2 norm of gradients: 0.2549551616792067
l2 norm of weights: 2.7011472819900737

Iteration Number: 3357
Loss: 19.285053101861777
l2 norm of gradients: 0.2548409002703114
l2 norm of weights: 2.7007011520099438

```

-----
Iteration Number: 3358
Loss: 19.278722070998967
l2 norm of gradients: 0.25472669737421466
l2 norm of weights: 2.7002553238895275
-----
Iteration Number: 3359
Loss: 19.272395914494236
l2 norm of gradients: 0.25461255298018065
l2 norm of weights: 2.699809797500324
-----
Iteration Number: 3360
Loss: 19.26607462883629
l2 norm of gradients: 0.2544984670774612
l2 norm of weights: 2.6993645727138227
-----
Iteration Number: 3361
Loss: 19.25975821051543
l2 norm of gradients: 0.2543844396552956
l2 norm of weights: 2.6989196494014998
-----
Iteration Number: 3362
Loss: 19.25344665602354
l2 norm of gradients: 0.25427047070291064
l2 norm of weights: 2.6984750274348213
-----
Iteration Number: 3363
Loss: 19.247139961854103
l2 norm of gradients: 0.2541565602095202
l2 norm of weights: 2.698030706685242
-----
Iteration Number: 3364
Loss: 19.24083812450217
l2 norm of gradients: 0.25404270816432567
l2 norm of weights: 2.6975866870242062
-----
Iteration Number: 3365
Loss: 19.234541140464415
l2 norm of gradients: 0.25392891455651523
l2 norm of weights: 2.697142968323146
-----
Iteration Number: 3366
Loss: 19.228249006239086
l2 norm of gradients: 0.2538151793752639
l2 norm of weights: 2.696699550453484
-----
Iteration Number: 3367
Loss: 19.22196171832603

```

```

12 norm of gradients: 0.2537015026097341
12 norm of weights: 2.6962564332866314
-----
Iteration Number: 3368
Loss: 19.215679273226666
12 norm of gradients: 0.2535878842490746
12 norm of weights: 2.695813616693988
-----
Iteration Number: 3369
Loss: 19.209401667444038
12 norm of gradients: 0.2534743242824208
12 norm of weights: 2.6953711005469447
-----
Iteration Number: 3370
Loss: 19.203128897482756
12 norm of gradients: 0.2533608226988948
12 norm of weights: 2.6949288847168797
-----
Iteration Number: 3371
Loss: 19.19686095984903
12 norm of gradients: 0.25324737948760523
12 norm of weights: 2.6944869690751623
-----
Iteration Number: 3372
Loss: 19.190597851050654
12 norm of gradients: 0.25313399463764663
12 norm of weights: 2.6940453534931508
-----
Iteration Number: 3373
Loss: 19.184339567597032
12 norm of gradients: 0.2530206681381004
12 norm of weights: 2.6936040378421935
-----
Iteration Number: 3374
Loss: 19.178086105999153
12 norm of gradients: 0.2529073999780336
12 norm of weights: 2.6931630219936276
-----
Iteration Number: 3375
Loss: 19.171837462769595
12 norm of gradients: 0.2527941901464993
12 norm of weights: 2.6927223058187812
-----
Iteration Number: 3376
Loss: 19.16559363442253
12 norm of gradients: 0.25268103863253677
12 norm of weights: 2.6922818891889717
-----

```

Iteration Number: 3377
Loss: 19.159354617473728
l2 norm of gradients: 0.252567945425171
l2 norm of weights: 2.6918417719755063

Iteration Number: 3378
Loss: 19.153120408440543
l2 norm of gradients: 0.25245491051341257
l2 norm of weights: 2.6914019540496827

Iteration Number: 3379
Loss: 19.14689100384194
l2 norm of gradients: 0.2523419338862578
l2 norm of weights: 2.690962435282788

Iteration Number: 3380
Loss: 19.140666400198462
l2 norm of gradients: 0.25222901553268845
l2 norm of weights: 2.6905232155461007

Iteration Number: 3381
Loss: 19.13444659403224
l2 norm of gradients: 0.25211615544167176
l2 norm of weights: 2.6900842947108883

Iteration Number: 3382
Loss: 19.128231581867023
l2 norm of gradients: 0.2520033536021602
l2 norm of weights: 2.6896456726484095

Iteration Number: 3383
Loss: 19.122021360228143
l2 norm of gradients: 0.25189061000309154
l2 norm of weights: 2.689207349229913

Iteration Number: 3384
Loss: 19.115815925642526
l2 norm of gradients: 0.25177792463338844
l2 norm of weights: 2.688769324326638

Iteration Number: 3385
Loss: 19.109615274638674
l2 norm of gradients: 0.25166529748195887
l2 norm of weights: 2.688331597809815

Iteration Number: 3386
Loss: 19.103419403746727
l2 norm of gradients: 0.25155272853769545


```

12 norm of weights: 2.6878941695506633
-----
Iteration Number: 3387
Loss: 19.097228309498384
12 norm of gradients: 0.2514402177894757
12 norm of weights: 2.687457039420395
-----
Iteration Number: 3388
Loss: 19.091041988426962
12 norm of gradients: 0.2513277652261619
12 norm of weights: 2.6870202072902116
-----
Iteration Number: 3389
Loss: 19.08486043706736
12 norm of gradients: 0.2512153708366009
12 norm of weights: 2.686583673031307
-----
Iteration Number: 3390
Loss: 19.078683651956073
12 norm of gradients: 0.2511030346096239
12 norm of weights: 2.686147436514864
-----
Iteration Number: 3391
Loss: 19.0725116296312
12 norm of gradients: 0.2509907565340467
12 norm of weights: 2.6857114976120577
-----
Iteration Number: 3392
Loss: 19.066344366632432
12 norm of gradients: 0.2508785365986693
12 norm of weights: 2.685275856194054
-----
Iteration Number: 3393
Loss: 19.060181859501068
12 norm of gradients: 0.25076637479227615
12 norm of weights: 2.6848405121320114
-----
Iteration Number: 3394
Loss: 19.05402410477998
12 norm of gradients: 0.2506542711036353
12 norm of weights: 2.684405465297076
-----
Iteration Number: 3395
Loss: 19.04787109901367
12 norm of gradients: 0.2505422255214993
12 norm of weights: 2.683970715560389
-----
Iteration Number: 3396

```

Loss: 19.041722838748207
l2 norm of gradients: 0.25043023803460446
l2 norm of weights: 2.683536262793081

Iteration Number: 3397
Loss: 19.035579320531276
l2 norm of gradients: 0.25031830863167087
l2 norm of weights: 2.6831021068662744

Iteration Number: 3398
Loss: 19.02944054091216
l2 norm of gradients: 0.25020643730140224
l2 norm of weights: 2.682668247651084

Iteration Number: 3399
Loss: 19.02330649644173
l2 norm of gradients: 0.2500946240324861
l2 norm of weights: 2.6822346850186154

Iteration Number: 3400
Loss: 19.01717718367247
l2 norm of gradients: 0.24998286881359352
l2 norm of weights: 2.6818014188399655

Iteration Number: 3401
Loss: 19.01105259915845
l2 norm of gradients: 0.24987117163337885
l2 norm of weights: 2.681368448986224

Iteration Number: 3402
Loss: 19.004932739455352
l2 norm of gradients: 0.24975953248047994
l2 norm of weights: 2.680935775328472

Iteration Number: 3403
Loss: 18.99881760112045
l2 norm of gradients: 0.2496479513435177
l2 norm of weights: 2.680503397737782

Iteration Number: 3404
Loss: 18.992707180712607
l2 norm of gradients: 0.2495364282110963
l2 norm of weights: 2.6800713160852196

Iteration Number: 3405
Loss: 18.986601474792323
l2 norm of gradients: 0.24942496307180312
l2 norm of weights: 2.6796395302418414

```

-----
Iteration Number: 3406
Loss: 18.980500479921655
l2 norm of gradients: 0.2493135559142082
l2 norm of weights: 2.6792080400786973
-----
Iteration Number: 3407
Loss: 18.974404192664288
l2 norm of gradients: 0.24920220672686466
l2 norm of weights: 2.6787768454668277
-----
Iteration Number: 3408
Loss: 18.968312609585496
l2 norm of gradients: 0.24909091549830836
l2 norm of weights: 2.678345946277267
-----
Iteration Number: 3409
Loss: 18.962225727252168
l2 norm of gradients: 0.24897968221705782
l2 norm of weights: 2.6779153423810413
-----
Iteration Number: 3410
Loss: 18.95614354223278
l2 norm of gradients: 0.24886850687161416
l2 norm of weights: 2.6774850336491696
-----
Iteration Number: 3411
Loss: 18.95006605109743
l2 norm of gradients: 0.24875738945046097
l2 norm of weights: 2.6770550199526624
-----
Iteration Number: 3412
Loss: 18.94399325041779
l2 norm of gradients: 0.24864632994206434
l2 norm of weights: 2.676625301162524
-----
Iteration Number: 3413
Loss: 18.93792513676716
l2 norm of gradients: 0.24853532833487257
l2 norm of weights: 2.6761958771497496
-----
Iteration Number: 3414
Loss: 18.931861706720436
l2 norm of gradients: 0.24842438461731634
l2 norm of weights: 2.675766747785329
-----
Iteration Number: 3415
Loss: 18.92580295685412

```

```

12 norm of gradients: 0.24831349877780834
12 norm of weights: 2.6753379129402446
-----
Iteration Number: 3416
Loss: 18.9197488837463
12 norm of gradients: 0.24820267080474331
12 norm of weights: 2.6749093724854713
-----
Iteration Number: 3417
Loss: 18.913699483976693
12 norm of gradients: 0.24809190068649803
12 norm of weights: 2.674481126291976
-----
Iteration Number: 3418
Loss: 18.907654754126607
12 norm of gradients: 0.24798118841143105
12 norm of weights: 2.6740531742307203
-----
Iteration Number: 3419
Loss: 18.901614690778963
12 norm of gradients: 0.24787053396788286
12 norm of weights: 2.673625516172658
-----
Iteration Number: 3420
Loss: 18.895579290518274
12 norm of gradients: 0.24775993734417545
12 norm of weights: 2.6731981519887364
-----
Iteration Number: 3421
Loss: 18.88954854993068
12 norm of gradients: 0.2476493985286125
12 norm of weights: 2.6727710815498957
-----
Iteration Number: 3422
Loss: 18.883522465603914
12 norm of gradients: 0.24753891750947937
12 norm of weights: 2.67234430472707
-----
Iteration Number: 3423
Loss: 18.877501034127306
12 norm of gradients: 0.24742849427504246
12 norm of weights: 2.6719178213911867
-----
Iteration Number: 3424
Loss: 18.87148425209181
12 norm of gradients: 0.24731812881354998
12 norm of weights: 2.671491631413166
-----

```

Iteration Number: 3425
Loss: 18.865472116089983
l2 norm of gradients: 0.247207821113231
l2 norm of weights: 2.671065734663922

Iteration Number: 3426
Loss: 18.859464622715983
l2 norm of gradients: 0.24709757116229603
l2 norm of weights: 2.6706401310143644

Iteration Number: 3427
Loss: 18.853461768565584
l2 norm of gradients: 0.24698737894893652
l2 norm of weights: 2.6702148203353926

Iteration Number: 3428
Loss: 18.84746355023616
l2 norm of gradients: 0.24687724446132503
l2 norm of weights: 2.6697898024979034

Iteration Number: 3429
Loss: 18.841469964326713
l2 norm of gradients: 0.2467671676876149
l2 norm of weights: 2.669365077372786

Iteration Number: 3430
Loss: 18.835481007437828
l2 norm of gradients: 0.24665714861594037
l2 norm of weights: 2.6689406448309234

Iteration Number: 3431
Loss: 18.829496676171704
l2 norm of gradients: 0.2465471872344165
l2 norm of weights: 2.668516504743193

Iteration Number: 3432
Loss: 18.82351696713218
l2 norm of gradients: 0.2464372835311389
l2 norm of weights: 2.668092656980466

Iteration Number: 3433
Loss: 18.817541876924665
l2 norm of gradients: 0.2463274374941838
l2 norm of weights: 2.6676691014136087

Iteration Number: 3434
Loss: 18.8115714021562
l2 norm of gradients: 0.24621764911160793

```

12 norm of weights: 2.66724583791348
-----
Iteration Number: 3435
Loss: 18.805605539435444
12 norm of gradients: 0.24610791837144844
12 norm of weights: 2.666822866350934
-----
Iteration Number: 3436
Loss: 18.799644285372647
12 norm of gradients: 0.24599824526172281
12 norm of weights: 2.6664001865968197
-----
Iteration Number: 3437
Loss: 18.793687636579687
12 norm of gradients: 0.24588862977042886
12 norm of weights: 2.66597779852198
-----
Iteration Number: 3438
Loss: 18.787735589670042
12 norm of gradients: 0.24577907188554435
12 norm of weights: 2.665555701997252
-----
Iteration Number: 3439
Loss: 18.781788141258826
12 norm of gradients: 0.2456695715950273
12 norm of weights: 2.665133896893468
-----
Iteration Number: 3440
Loss: 18.77584528796273
12 norm of gradients: 0.24556012888681572
12 norm of weights: 2.6647123830814547
-----
Iteration Number: 3441
Loss: 18.769907026400098
12 norm of gradients: 0.24545074374882755
12 norm of weights: 2.6642911604320343
-----
Iteration Number: 3442
Loss: 18.763973353190856
12 norm of gradients: 0.2453414161689606
12 norm of weights: 2.663870228816022
-----
Iteration Number: 3443
Loss: 18.75804426495656
12 norm of gradients: 0.24523214613509234
12 norm of weights: 2.6634495881042293
-----
Iteration Number: 3444

```

```

Loss: 18.75211975832038
l2 norm of gradients: 0.24512293363508006
l2 norm of weights: 2.6630292381674634
-----
Iteration Number: 3445
Loss: 18.746199829907106
l2 norm of gradients: 0.24501377865676052
l2 norm of weights: 2.6626091788765245
-----
Iteration Number: 3446
Loss: 18.740284476343124
l2 norm of gradients: 0.2449046811879502
l2 norm of weights: 2.662189410102209
-----
Iteration Number: 3447
Loss: 18.734373694256462
l2 norm of gradients: 0.24479564121644487
l2 norm of weights: 2.6617699317153094
-----
Iteration Number: 3448
Loss: 18.72846748027675
l2 norm of gradients: 0.2446866587300198
l2 norm of weights: 2.6613507435866124
-----
Iteration Number: 3449
Loss: 18.722565831035237
l2 norm of gradients: 0.24457773371642952
l2 norm of weights: 2.660931845586899
-----
Iteration Number: 3450
Loss: 18.71666874316478
l2 norm of gradients: 0.24446886616340768
l2 norm of weights: 2.660513237586948
-----
Iteration Number: 3451
Loss: 18.71077621329988
l2 norm of gradients: 0.24436005605866729
l2 norm of weights: 2.6600949194575314
-----
Iteration Number: 3452
Loss: 18.704888238076634
l2 norm of gradients: 0.24425130338990023
l2 norm of weights: 2.6596768910694197
-----
Iteration Number: 3453
Loss: 18.699004814132763
l2 norm of gradients: 0.24414260814477753
l2 norm of weights: 2.6592591522933757

```

```

-----
Iteration Number: 3454
Loss: 18.693125938107613
l2 norm of gradients: 0.24403397031094912
l2 norm of weights: 2.6588417030001597
-----
Iteration Number: 3455
Loss: 18.687251606642143
l2 norm of gradients: 0.2439253898760436
l2 norm of weights: 2.6584245430605273
-----
Iteration Number: 3456
Loss: 18.681381816378934
l2 norm of gradients: 0.24381686682766865
l2 norm of weights: 2.658007672345231
-----
Iteration Number: 3457
Loss: 18.675516563962198
l2 norm of gradients: 0.24370840115341041
l2 norm of weights: 2.6575910907250178
-----
Iteration Number: 3458
Loss: 18.669655846037745
l2 norm of gradients: 0.2435999928408339
l2 norm of weights: 2.657174798070631
-----
Iteration Number: 3459
Loss: 18.66379965925303
l2 norm of gradients: 0.2434916418774823
l2 norm of weights: 2.65675879425281
-----
Iteration Number: 3460
Loss: 18.65794800025712
l2 norm of gradients: 0.24338334825087768
l2 norm of weights: 2.656343079142292
-----
Iteration Number: 3461
Loss: 18.652100865700707
l2 norm of gradients: 0.2432751119485204
l2 norm of weights: 2.6559276526098072
-----
Iteration Number: 3462
Loss: 18.646258252236095
l2 norm of gradients: 0.24316693295788902
l2 norm of weights: 2.655512514526085
-----
Iteration Number: 3463
Loss: 18.640420156517227

```



```

12 norm of gradients: 0.24305881126644055
12 norm of weights: 2.65509766476185
-----
Iteration Number: 3464
Loss: 18.634586575199677
12 norm of gradients: 0.24295074686161014
12 norm of weights: 2.654683103187823
-----
Iteration Number: 3465
Loss: 18.6287575049406
12 norm of gradients: 0.24284273973081105
12 norm of weights: 2.6542688296747214
-----
Iteration Number: 3466
Loss: 18.622932942398844
12 norm of gradients: 0.2427347898614346
12 norm of weights: 2.6538548440932597
-----
Iteration Number: 3467
Loss: 18.61711288423482
12 norm of gradients: 0.2426268972408502
12 norm of weights: 2.6534411463141487
-----
Iteration Number: 3468
Loss: 18.61129732711059
12 norm of gradients: 0.24251906185640515
12 norm of weights: 2.6530277362080956
-----
Iteration Number: 3469
Loss: 18.60548626768985
12 norm of gradients: 0.24241128369542445
12 norm of weights: 2.652614613645805
-----
Iteration Number: 3470
Loss: 18.599679702637918
12 norm of gradients: 0.2423035627452111
12 norm of weights: 2.652201778497978
-----
Iteration Number: 3471
Loss: 18.593877628621737
12 norm of gradients: 0.24219589899304572
12 norm of weights: 2.651789230635313
-----
Iteration Number: 3472
Loss: 18.588080042309866
12 norm of gradients: 0.24208829242618665
12 norm of weights: 2.6513769699285046
-----

```

Iteration Number: 3473
Loss: 18.582286940372512
l2 norm of gradients: 0.24198074303186973
l2 norm of weights: 2.650964996248245

Iteration Number: 3474
Loss: 18.57649831948151
l2 norm of gradients: 0.24187325079730845
l2 norm of weights: 2.650553309465224

Iteration Number: 3475
Loss: 18.570714176310304
l2 norm of gradients: 0.24176581570969366
l2 norm of weights: 2.650141909450128

Iteration Number: 3476
Loss: 18.564934507533994
l2 norm of gradients: 0.24165843775619358
l2 norm of weights: 2.649730796073641

Iteration Number: 3477
Loss: 18.559159309829287
l2 norm of gradients: 0.24155111692395392
l2 norm of weights: 2.649319969206444

Iteration Number: 3478
Loss: 18.553388579874536
l2 norm of gradients: 0.24144385320009742
l2 norm of weights: 2.6489094287192163

Iteration Number: 3479
Loss: 18.547622314349717
l2 norm of gradients: 0.24133664657172424
l2 norm of weights: 2.6484991744826334

Iteration Number: 3480
Loss: 18.541860509936452
l2 norm of gradients: 0.2412294970259116
l2 norm of weights: 2.6480892063673687

Iteration Number: 3481
Loss: 18.536103163317982
l2 norm of gradients: 0.24112240454971368
l2 norm of weights: 2.647679524244095

Iteration Number: 3482
Loss: 18.530350271179174
l2 norm of gradients: 0.24101536913016194

```
l2 norm of weights: 2.64727012798348
-----
Iteration Number: 3483
Loss: 18.52460183020655
l2 norm of gradients: 0.2409083907542645
l2 norm of weights: 2.646861017456192
-----
Iteration Number: 3484
Loss: 18.518857837088255
l2 norm of gradients: 0.24080146940900662
l2 norm of weights: 2.6464521925328945
-----
Iteration Number: 3485
Loss: 18.513118288514057
l2 norm of gradients: 0.24069460508135024
l2 norm of weights: 2.6460436530842513
-----
Iteration Number: 3486
Loss: 18.50738318117539
l2 norm of gradients: 0.24058779775823405
l2 norm of weights: 2.6456353989809225
-----
Iteration Number: 3487
Loss: 18.501652511765275
l2 norm of gradients: 0.24048104742657359
l2 norm of weights: 2.6452274300935676
-----
Iteration Number: 3488
Loss: 18.495926276978423
l2 norm of gradients: 0.24037435407326094
l2 norm of weights: 2.6448197462928427
-----
Iteration Number: 3489
Loss: 18.490204473511152
l2 norm of gradients: 0.2402677176851648
l2 norm of weights: 2.6444123474494035
-----
Iteration Number: 3490
Loss: 18.48448709806141
l2 norm of gradients: 0.24016113824913043
l2 norm of weights: 2.6440052334339033
-----
Iteration Number: 3491
Loss: 18.47877414732881
l2 norm of gradients: 0.24005461575197945
l2 norm of weights: 2.643598404116994
-----
Iteration Number: 3492
```

Loss: 18.473065618014576
l2 norm of gradients: 0.23994815018051008
l2 norm of weights: 2.643191859369326

Iteration Number: 3493
Loss: 18.467361506821582
l2 norm of gradients: 0.23984174152149676
l2 norm of weights: 2.6427855990615483

Iteration Number: 3494
Loss: 18.461661810454345
l2 norm of gradients: 0.23973538976169032
l2 norm of weights: 2.6423796230643077

Iteration Number: 3495
Loss: 18.45596652561902
l2 norm of gradients: 0.2396290948878177
l2 norm of weights: 2.6419739312482506

Iteration Number: 3496
Loss: 18.450275649023396
l2 norm of gradients: 0.23952285688658226
l2 norm of weights: 2.6415685234840214

Iteration Number: 3497
Loss: 18.444589177376912
l2 norm of gradients: 0.2394166757446633
l2 norm of weights: 2.6411633996422634

Iteration Number: 3498
Loss: 18.43890710739063
l2 norm of gradients: 0.23931055144871624
l2 norm of weights: 2.64075855959362

Iteration Number: 3499
Loss: 18.43322943577728
l2 norm of gradients: 0.23920448398537258
l2 norm of weights: 2.6403540032087305

Iteration Number: 3500
Loss: 18.427556159251232
l2 norm of gradients: 0.23909847334123976
l2 norm of weights: 2.639949730358237

Iteration Number: 3501
Loss: 18.42188727452846
l2 norm of gradients: 0.23899251950290107
l2 norm of weights: 2.6395457409127783

```

-----
Iteration Number: 3502
Loss: 18.416222778326627
l2 norm of gradients: 0.23888662245691586
l2 norm of weights: 2.639142034742992
-----
Iteration Number: 3503
Loss: 18.410562667365028
l2 norm of gradients: 0.23878078218981905
l2 norm of weights: 2.638738611719517
-----
Iteration Number: 3504
Loss: 18.404906938364586
l2 norm of gradients: 0.23867499868812153
l2 norm of weights: 2.638335471712989
-----
Iteration Number: 3505
Loss: 18.399255588047893
l2 norm of gradients: 0.23856927193830976
l2 norm of weights: 2.6379326145940447
-----
Iteration Number: 3506
Loss: 18.39360861313917
l2 norm of gradients: 0.23846360192684593
l2 norm of weights: 2.63753004023332
-----
Iteration Number: 3507
Loss: 18.38796601036428
l2 norm of gradients: 0.23835798864016775
l2 norm of weights: 2.63712774850145
-----
Iteration Number: 3508
Loss: 18.382327776450758
l2 norm of gradients: 0.23825243206468852
l2 norm of weights: 2.636725739269069
-----
Iteration Number: 3509
Loss: 18.376693908127752
l2 norm of gradients: 0.23814693218679725
l2 norm of weights: 2.636324012406811
-----
Iteration Number: 3510
Loss: 18.37106440212609
l2 norm of gradients: 0.23804148899285799
l2 norm of weights: 2.63592256778531
-----
Iteration Number: 3511
Loss: 18.365439255178238

```

```

12 norm of gradients: 0.2379361024692107
12 norm of weights: 2.635521405275201
-----
Iteration Number: 3512
Loss: 18.35981846401829
12 norm of gradients: 0.2378307726021702
12 norm of weights: 2.635120524747116
-----
Iteration Number: 3513
Loss: 18.354202025382023
12 norm of gradients: 0.23772549937802692
12 norm of weights: 2.634719926071689
-----
Iteration Number: 3514
Loss: 18.348589936006835
12 norm of gradients: 0.23762028278304656
12 norm of weights: 2.6343196091195535
-----
Iteration Number: 3515
Loss: 18.3429821926318
12 norm of gradients: 0.23751512280346998
12 norm of weights: 2.6339195737613426
-----
Iteration Number: 3516
Loss: 18.337378791997622
12 norm of gradients: 0.23741001942551307
12 norm of weights: 2.6335198198676903
-----
Iteration Number: 3517
Loss: 18.33177973084667
12 norm of gradients: 0.23730497263536712
12 norm of weights: 2.6331203473092293
-----
Iteration Number: 3518
Loss: 18.32618500592296
12 norm of gradients: 0.2371999824191983
12 norm of weights: 2.632721155956594
-----
Iteration Number: 3519
Loss: 18.320594613972165
12 norm of gradients: 0.23709504876314774
12 norm of weights: 2.632322245680419
-----
Iteration Number: 3520
Loss: 18.315008551741606
12 norm of gradients: 0.23699017165333186
12 norm of weights: 2.631923616351338
-----

```

Iteration Number: 3521
Loss: 18.309426815980267
l2 norm of gradients: 0.23688535107584174
l2 norm of weights: 2.6315252678399865

Iteration Number: 3522
Loss: 18.303849403438765
l2 norm of gradients: 0.2367805870167436
l2 norm of weights: 2.6311272000170005

Iteration Number: 3523
Loss: 18.2982763108694
l2 norm of gradients: 0.23667587946207824
l2 norm of weights: 2.630729412753015

Iteration Number: 3524
Loss: 18.292707535026114
l2 norm of gradients: 0.2365712283978615
l2 norm of weights: 2.630331905918667

Iteration Number: 3525
Loss: 18.287143072664506
l2 norm of gradients: 0.23646663381008381
l2 norm of weights: 2.6299346793845944

Iteration Number: 3526
Loss: 18.281582920541826
l2 norm of gradients: 0.2363620956847105
l2 norm of weights: 2.629537733021435

Iteration Number: 3527
Loss: 18.276027075416994
l2 norm of gradients: 0.23625761400768153
l2 norm of weights: 2.6291410666998276

Iteration Number: 3528
Loss: 18.270475534050576
l2 norm of gradients: 0.23615318876491145
l2 norm of weights: 2.628744680290413

Iteration Number: 3529
Loss: 18.26492829320481
l2 norm of gradients: 0.2360488199422895
l2 norm of weights: 2.628348573663831

Iteration Number: 3530
Loss: 18.259385349643587
l2 norm of gradients: 0.23594450752567947

```

12 norm of weights: 2.6279527466907244
-----
Iteration Number: 3531
Loss: 18.253846700132446
12 norm of gradients: 0.2358402515009195
12 norm of weights: 2.627557199241736
-----
Iteration Number: 3532
Loss: 18.248312341438606
12 norm of gradients: 0.23573605185382257
12 norm of weights: 2.62716193118751
-----
Iteration Number: 3533
Loss: 18.242782270330924
12 norm of gradients: 0.23563190857017577
12 norm of weights: 2.6267669423986915
-----
Iteration Number: 3534
Loss: 18.23725648357995
12 norm of gradients: 0.23552782163574085
12 norm of weights: 2.626372232745928
-----
Iteration Number: 3535
Loss: 18.231734977957867
12 norm of gradients: 0.23542379103625372
12 norm of weights: 2.6259778020998668
-----
Iteration Number: 3536
Loss: 18.22621775023853
12 norm of gradients: 0.23531981675742478
12 norm of weights: 2.6255836503311585
-----
Iteration Number: 3537
Loss: 18.22070479719746
12 norm of gradients: 0.23521589878493862
12 norm of weights: 2.6251897773104536
-----
Iteration Number: 3538
Loss: 18.215196115611842
12 norm of gradients: 0.23511203710445425
12 norm of weights: 2.624796182908405
-----
Iteration Number: 3539
Loss: 18.209691702260525
12 norm of gradients: 0.23500823170160468
12 norm of weights: 2.6244028669956676
-----
Iteration Number: 3540

```


Loss: 18.204191553924016
l2 norm of gradients: 0.23490448256199734
l2 norm of weights: 2.6240098294428966

Iteration Number: 3541
Loss: 18.1986956673845
l2 norm of gradients: 0.2348007896712137
l2 norm of weights: 2.62361707012075

Iteration Number: 3542
Loss: 18.193204039425815
l2 norm of gradients: 0.23469715301480928
l2 norm of weights: 2.623224588899888

Iteration Number: 3543
Loss: 18.187716666833474
l2 norm of gradients: 0.23459357257831387
l2 norm of weights: 2.622832385650972

Iteration Number: 3544
Loss: 18.182233546394656
l2 norm of gradients: 0.2344900483472311
l2 norm of weights: 2.622440460244665

Iteration Number: 3545
Loss: 18.1767546748982
l2 norm of gradients: 0.23438658030703882
l2 norm of weights: 2.6220488125516335

Iteration Number: 3546
Loss: 18.171280049134623
l2 norm of gradients: 0.2342831684431887
l2 norm of weights: 2.621657442442545

Iteration Number: 3547
Loss: 18.165809665896113
l2 norm of gradients: 0.2341798127411065
l2 norm of weights: 2.6212663497880695

Iteration Number: 3548
Loss: 18.160343521976515
l2 norm of gradients: 0.23407651318619185
l2 norm of weights: 2.6208755344588783

Iteration Number: 3549
Loss: 18.15488161417136
l2 norm of gradients: 0.23397326976381821
l2 norm of weights: 2.620484996325647

```
-----  
Iteration Number: 3550  
Loss: 18.149423939277824  
l2 norm of gradients: 0.23387008245933308  
l2 norm of weights: 2.6200947352590522  
-----  
Iteration Number: 3551  
Loss: 18.143970494094788  
l2 norm of gradients: 0.23376695125805755  
l2 norm of weights: 2.6197047511297726  
-----  
Iteration Number: 3552  
Loss: 18.13852127542278  
l2 norm of gradients: 0.23366387614528666  
l2 norm of weights: 2.61931504380849  
-----  
Iteration Number: 3553  
Loss: 18.13307628006401  
l2 norm of gradients: 0.23356085710628924  
l2 norm of weights: 2.6189256131658896  
-----  
Iteration Number: 3554  
Loss: 18.127635504822354  
l2 norm of gradients: 0.23345789412630785  
l2 norm of weights: 2.6185364590726574  
-----  
Iteration Number: 3555  
Loss: 18.122198946503374  
l2 norm of gradients: 0.2333549871905586  
l2 norm of weights: 2.618147581399484  
-----  
Iteration Number: 3556  
Loss: 18.1167666019143  
l2 norm of gradients: 0.23325213628423153  
l2 norm of weights: 2.617758980017061  
-----  
Iteration Number: 3557  
Loss: 18.111338467864034  
l2 norm of gradients: 0.23314934139249016  
l2 norm of weights: 2.617370654796084  
-----  
Iteration Number: 3558  
Loss: 18.105914541163145  
l2 norm of gradients: 0.2330466025004717  
l2 norm of weights: 2.6169826056072516  
-----  
Iteration Number: 3559  
Loss: 18.1004948186239
```

```

12 norm of gradients: 0.232943919593287
12 norm of weights: 2.6165948323212644
-----
Iteration Number: 3560
Loss: 18.09507929706023
12 norm of gradients: 0.23284129265602052
12 norm of weights: 2.616207334808826
-----
Iteration Number: 3561
Loss: 18.08966797328773
12 norm of gradients: 0.23273872167373008
12 norm of weights: 2.615820112940645
-----
Iteration Number: 3562
Loss: 18.0842608441237
12 norm of gradients: 0.23263620663144727
12 norm of weights: 2.6154331665874313
-----
Iteration Number: 3563
Loss: 18.0788579063871
12 norm of gradients: 0.232533747514177
12 norm of weights: 2.6150464956198975
-----
Iteration Number: 3564
Loss: 18.073459156898565
12 norm of gradients: 0.23243134430689769
12 norm of weights: 2.614660099908762
-----
Iteration Number: 3565
Loss: 18.06806459248043
12 norm of gradients: 0.2323289969945613
12 norm of weights: 2.614273979324744
-----
Iteration Number: 3566
Loss: 18.06267420995669
12 norm of gradients: 0.23222670556209324
12 norm of weights: 2.6138881337385675
-----
Iteration Number: 3567
Loss: 18.05728800615302
12 norm of gradients: 0.2321244699943921
12 norm of weights: 2.613502563020959
-----
Iteration Number: 3568
Loss: 18.051905977896805
12 norm of gradients: 0.23202229027633015
12 norm of weights: 2.6131172670426506
-----

```

Iteration Number: 3569
Loss: 18.046528122017065
l2 norm of gradients: 0.23192016639275284
l2 norm of weights: 2.6127322456743753

Iteration Number: 3570
Loss: 18.04115443534455
l2 norm of gradients: 0.23181809832847894
l2 norm of weights: 2.6123474987868716

Iteration Number: 3571
Loss: 18.03578491471165
l2 norm of gradients: 0.2317160860683007
l2 norm of weights: 2.611963026250881

Iteration Number: 3572
Loss: 18.03041955695247
l2 norm of gradients: 0.23161412959698346
l2 norm of weights: 2.6115788279371484

Iteration Number: 3573
Loss: 18.025058358902793
l2 norm of gradients: 0.231512228899266
l2 norm of weights: 2.6111949037164237

Iteration Number: 3574
Loss: 18.019701317400067
l2 norm of gradients: 0.23141038395986038
l2 norm of weights: 2.61081125345946

Iteration Number: 3575
Loss: 18.014348429283444
l2 norm of gradients: 0.23130859476345167
l2 norm of weights: 2.6104278770370146

Iteration Number: 3576
Loss: 18.008999691393768
l2 norm of gradients: 0.23120686129469845
l2 norm of weights: 2.610044774319849

Iteration Number: 3577
Loss: 18.003655100573553
l2 norm of gradients: 0.23110518353823223
l2 norm of weights: 2.609661945178728

Iteration Number: 3578
Loss: 17.99831465366699
l2 norm of gradients: 0.23100356147865794

```

12 norm of weights: 2.609279389484421
-----
Iteration Number: 3579
Loss: 17.992978347519998
12 norm of gradients: 0.23090199510055354
12 norm of weights: 2.608897107107703
-----
Iteration Number: 3580
Loss: 17.987646178980143
12 norm of gradients: 0.23080048438847015
12 norm of weights: 2.6085150979193505
-----
Iteration Number: 3581
Loss: 17.982318144896695
12 norm of gradients: 0.230699029326932
12 norm of weights: 2.608133361790147
-----
Iteration Number: 3582
Loss: 17.97699424212062
12 norm of gradients: 0.23059762990043656
12 norm of weights: 2.607751898590879
-----
Iteration Number: 3583
Loss: 17.971674467504563
12 norm of gradients: 0.2304962860934542
12 norm of weights: 2.6073707081923376
-----
Iteration Number: 3584
Loss: 17.966358817902872
12 norm of gradients: 0.23039499789042864
12 norm of weights: 2.6069897904653194
-----
Iteration Number: 3585
Loss: 17.961047290171567
12 norm of gradients: 0.23029376527577633
12 norm of weights: 2.6066091452806237
-----
Iteration Number: 3586
Loss: 17.955739881168373
12 norm of gradients: 0.23019258823388705
12 norm of weights: 2.606228772509057
-----
Iteration Number: 3587
Loss: 17.95043658775271
12 norm of gradients: 0.2300914667491235
12 norm of weights: 2.6058486720214287
-----
Iteration Number: 3588

```

```

Loss: 17.945137406785676
l2 norm of gradients: 0.2299904008058214
l2 norm of weights: 2.605468843688554
-----
Iteration Number: 3589
Loss: 17.939842335130074
l2 norm of gradients: 0.2298893903882896
l2 norm of weights: 2.6050892873812512
-----
Iteration Number: 3590
Loss: 17.934551369650396
l2 norm of gradients: 0.2297884354808097
l2 norm of weights: 2.604710002970346
-----
Iteration Number: 3591
Loss: 17.92926450721284
l2 norm of gradients: 0.22968753606763662
l2 norm of weights: 2.604330990326668
-----
Iteration Number: 3592
Loss: 17.92398174468528
l2 norm of gradients: 0.2295866921329979
l2 norm of weights: 2.6039522493210505
-----
Iteration Number: 3593
Loss: 17.918703078937295
l2 norm of gradients: 0.22948590366109428
l2 norm of weights: 2.603573779824335
-----
Iteration Number: 3594
Loss: 17.913428506840162
l2 norm of gradients: 0.22938517063609945
l2 norm of weights: 2.6031955817073644
-----
Iteration Number: 3595
Loss: 17.90815802526685
l2 norm of gradients: 0.22928449304215992
l2 norm of weights: 2.60281765484099
-----
Iteration Number: 3596
Loss: 17.90289163109203
l2 norm of gradients: 0.2291838708633952
l2 norm of weights: 2.602439999096067
-----
Iteration Number: 3597
Loss: 17.897629321192063
l2 norm of gradients: 0.22908330408389768
l2 norm of weights: 2.602062614343456

```

```

-----
Iteration Number: 3598
Loss: 17.892371092445025
l2 norm of gradients: 0.2289827926877328
l2 norm of weights: 2.601685500454023
-----
Iteration Number: 3599
Loss: 17.887116941730667
l2 norm of gradients: 0.2288823366589387
l2 norm of weights: 2.60130865729864
-----
Iteration Number: 3600
Loss: 17.881866865930462
l2 norm of gradients: 0.2287819359815265
l2 norm of weights: 2.600932084748184
-----
Iteration Number: 3601
Loss: 17.876620861927563
l2 norm of gradients: 0.22868159063948026
l2 norm of weights: 2.6005557826735375
-----
Iteration Number: 3602
Loss: 17.871378926606837
l2 norm of gradients: 0.22858130061675688
l2 norm of weights: 2.6001797509455895
-----
Iteration Number: 3603
Loss: 17.866141056854854
l2 norm of gradients: 0.22848106589728606
l2 norm of weights: 2.5998039894352334
-----
Iteration Number: 3604
Loss: 17.860907249559872
l2 norm of gradients: 0.22838088646497046
l2 norm of weights: 2.5994284980133697
-----
Iteration Number: 3605
Loss: 17.85567750161187
l2 norm of gradients: 0.22828076230368552
l2 norm of weights: 2.5990532765509045
-----
Iteration Number: 3606
Loss: 17.85045180990251
l2 norm of gradients: 0.22818069339727964
l2 norm of weights: 2.5986783249187493
-----
Iteration Number: 3607
Loss: 17.845230171325177

```

```

12 norm of gradients: 0.22808067972957397
12 norm of weights: 2.5983036429878212
-----
Iteration Number: 3608
Loss: 17.840012582774936
12 norm of gradients: 0.22798072128436253
12 norm of weights: 2.5979292306290445
-----
Iteration Number: 3609
Loss: 17.834799041148578
12 norm of gradients: 0.2278808180454121
12 norm of weights: 2.597555087713349
-----
Iteration Number: 3610
Loss: 17.8295895433446
12 norm of gradients: 0.22778096999646247
12 norm of weights: 2.59718121411167
-----
Iteration Number: 3611
Loss: 17.82438408626318
12 norm of gradients: 0.22768117712122607
12 norm of weights: 2.5968076096949506
-----
Iteration Number: 3612
Loss: 17.819182666806235
12 norm of gradients: 0.22758143940338826
12 norm of weights: 2.596434274334139
-----
Iteration Number: 3613
Loss: 17.81398528187736
12 norm of gradients: 0.2274817568266072
12 norm of weights: 2.5960612079001892
-----
Iteration Number: 3614
Loss: 17.808791928381872
12 norm of gradients: 0.22738212937451388
12 norm of weights: 2.595688410264063
-----
Iteration Number: 3615
Loss: 17.8036026032268
12 norm of gradients: 0.2272825570307121
12 norm of weights: 2.5953158812967283
-----
Iteration Number: 3616
Loss: 17.798417303320868
12 norm of gradients: 0.22718303977877835
12 norm of weights: 2.5949436208691594
-----

```


Iteration Number: 3617
Loss: 17.79323602557452
l2 norm of gradients: 0.22708357760226214
l2 norm of weights: 2.5945716288523357

Iteration Number: 3618
Loss: 17.788058766899898
l2 norm of gradients: 0.22698417048468555
l2 norm of weights: 2.5941999051172457

Iteration Number: 3619
Loss: 17.782885524210872
l2 norm of gradients: 0.22688481840954383
l2 norm of weights: 2.5938284495348833

Iteration Number: 3620
Loss: 17.77771629442301
l2 norm of gradients: 0.22678552136030466
l2 norm of weights: 2.593457261976249

Iteration Number: 3621
Loss: 17.772551074453588
l2 norm of gradients: 0.22668627932040866
l2 norm of weights: 2.5930863423123505

Iteration Number: 3622
Loss: 17.7673898612216
l2 norm of gradients: 0.22658709227326942
l2 norm of weights: 2.592715690414202

Iteration Number: 3623
Loss: 17.76223265164775
l2 norm of gradients: 0.22648796020227313
l2 norm of weights: 2.5923453061528257

Iteration Number: 3624
Loss: 17.757079442654465
l2 norm of gradients: 0.2263888830907789
l2 norm of weights: 2.591975189399249

Iteration Number: 3625
Loss: 17.75193023116587
l2 norm of gradients: 0.22628986092211859
l2 norm of weights: 2.591605340024508

Iteration Number: 3626
Loss: 17.746785014107807
l2 norm of gradients: 0.2261908936795969

```

12 norm of weights: 2.5912357578996454
-----
Iteration Number: 3627
Loss: 17.74164378840784
12 norm of gradients: 0.22609198134649142
12 norm of weights: 2.59086644289571
-----
Iteration Number: 3628
Loss: 17.736506550995244
12 norm of gradients: 0.22599312390605242
12 norm of weights: 2.59049739488376
-----
Iteration Number: 3629
Loss: 17.731373298801014
12 norm of gradients: 0.2258943213415031
12 norm of weights: 2.590128613734859
-----
Iteration Number: 3630
Loss: 17.726244028757847
12 norm of gradients: 0.22579557363603955
12 norm of weights: 2.589760099320078
-----
Iteration Number: 3631
Loss: 17.721118737800168
12 norm of gradients: 0.22569688077283045
12 norm of weights: 2.5893918515104972
-----
Iteration Number: 3632
Loss: 17.71599742286412
12 norm of gradients: 0.2255982427350176
12 norm of weights: 2.589023870177202
-----
Iteration Number: 3633
Loss: 17.71088008088757
12 norm of gradients: 0.22549965950571557
12 norm of weights: 2.588656155191287
-----
Iteration Number: 3634
Loss: 17.705766708810074
12 norm of gradients: 0.22540113106801152
12 norm of weights: 2.5882887064238536
-----
Iteration Number: 3635
Loss: 17.700657303572942
12 norm of gradients: 0.2253026574049658
12 norm of weights: 2.587921523746011
-----
Iteration Number: 3636

```

```

Loss: 17.695551862119185
l2 norm of gradients: 0.22520423849961146
l2 norm of weights: 2.587554607028876
-----
Iteration Number: 3637
Loss: 17.690450381393532
l2 norm of gradients: 0.22510587433495446
l2 norm of weights: 2.5871879561435733
-----
Iteration Number: 3638
Loss: 17.68535285834244
l2 norm of gradients: 0.22500756489397353
l2 norm of weights: 2.5868215709612357
-----
Iteration Number: 3639
Loss: 17.680259289914087
l2 norm of gradients: 0.2249093101596204
l2 norm of weights: 2.586455451353003
-----
Iteration Number: 3640
Loss: 17.67516967305837
l2 norm of gradients: 0.22481111011481966
l2 norm of weights: 2.5860895971900235
-----
Iteration Number: 3641
Loss: 17.6700840047269
l2 norm of gradients: 0.22471296474246882
l2 norm of weights: 2.5857240083434547
-----
Iteration Number: 3642
Loss: 17.665002281873022
l2 norm of gradients: 0.22461487402543812
l2 norm of weights: 2.585358684684459
-----
Iteration Number: 3643
Loss: 17.659924501451805
l2 norm of gradients: 0.22451683794657099
l2 norm of weights: 2.58499362608421
-----
Iteration Number: 3644
Loss: 17.65485066042002
l2 norm of gradients: 0.22441885648868354
l2 norm of weights: 2.5846288324138884
-----
Iteration Number: 3645
Loss: 17.649780755736206
l2 norm of gradients: 0.2243209296345649
l2 norm of weights: 2.5842643035446824

```

```

-----
Iteration Number: 3646
Loss: 17.64471478436057
l2 norm of gradients: 0.22422305736697723
l2 norm of weights: 2.58390003934779
-----
Iteration Number: 3647
Loss: 17.639652743255088
l2 norm of gradients: 0.2241252396686554
l2 norm of weights: 2.5835360396944154
-----
Iteration Number: 3648
Loss: 17.63459462938344
l2 norm of gradients: 0.22402747652230753
l2 norm of weights: 2.583172304455773
-----
Iteration Number: 3649
Loss: 17.629540439711043
l2 norm of gradients: 0.22392976791061456
l2 norm of weights: 2.582808833503085
-----
Iteration Number: 3650
Loss: 17.624490171205036
l2 norm of gradients: 0.22383211381623033
l2 norm of weights: 2.582445626707583
-----
Iteration Number: 3651
Loss: 17.619443820834277
l2 norm of gradients: 0.2237345142217819
l2 norm of weights: 2.5820826839405053
-----
Iteration Number: 3652
Loss: 17.61440138556937
l2 norm of gradients: 0.22363696910986905
l2 norm of weights: 2.5817200050731004
-----
Iteration Number: 3653
Loss: 17.609362862382635
l2 norm of gradients: 0.2235394784630649
l2 norm of weights: 2.5813575899766246
-----
Iteration Number: 3654
Loss: 17.604328248248116
l2 norm of gradients: 0.2234420422639153
l2 norm of weights: 2.580995438522344
-----
Iteration Number: 3655
Loss: 17.599297540141595

```

```

12 norm of gradients: 0.2233446604949393
12 norm of weights: 2.5806335505815325
-----
Iteration Number: 3656
Loss: 17.594270735040585
12 norm of gradients: 0.22324733313862902
12 norm of weights: 2.5802719260254725
-----
Iteration Number: 3657
Loss: 17.58924782992432
12 norm of gradients: 0.22315006017744962
12 norm of weights: 2.579910564725457
-----
Iteration Number: 3658
Loss: 17.58422882177377
12 norm of gradients: 0.2230528415938393
12 norm of weights: 2.5795494665527867
-----
Iteration Number: 3659
Loss: 17.57921370757164
12 norm of gradients: 0.2229556773702095
12 norm of weights: 2.579188631378771
-----
Iteration Number: 3660
Loss: 17.57420248430236
12 norm of gradients: 0.22285856748894453
12 norm of weights: 2.5788280590747292
-----
Iteration Number: 3661
Loss: 17.569195148952083
12 norm of gradients: 0.22276151193240207
12 norm of weights: 2.57846774951199
-----
Iteration Number: 3662
Loss: 17.56419169850873
12 norm of gradients: 0.22266451068291282
12 norm of weights: 2.5781077025618906
-----
Iteration Number: 3663
Loss: 17.55919212996191
12 norm of gradients: 0.22256756372278072
12 norm of weights: 2.577747918095777
-----
Iteration Number: 3664
Loss: 17.554196440303002
12 norm of gradients: 0.2224706710342827
12 norm of weights: 2.5773883959850066
-----

```

Iteration Number: 3665
Loss: 17.54920462652509
l2 norm of gradients: 0.22237383259966917
l2 norm of weights: 2.577029136100943

Iteration Number: 3666
Loss: 17.544216685623017
l2 norm of gradients: 0.22227704840116358
l2 norm of weights: 2.5766701383149626

Iteration Number: 3667
Loss: 17.53923261459335
l2 norm of gradients: 0.22218031842096256
l2 norm of weights: 2.5763114024984484

Iteration Number: 3668
Loss: 17.53425241043438
l2 norm of gradients: 0.22208364264123615
l2 norm of weights: 2.5759529285227942

Iteration Number: 3669
Loss: 17.529276070146153
l2 norm of gradients: 0.22198702104412743
l2 norm of weights: 2.575594716259405

Iteration Number: 3670
Loss: 17.52430359073045
l2 norm of gradients: 0.22189045361175302
l2 norm of weights: 2.5752367655796915

Iteration Number: 3671
Loss: 17.51933496919078
l2 norm of gradients: 0.22179394032620267
l2 norm of weights: 2.574879076355078

Iteration Number: 3672
Loss: 17.514370202532383
l2 norm of gradients: 0.22169748116953944
l2 norm of weights: 2.5745216484569964

Iteration Number: 3673
Loss: 17.509409287762267
l2 norm of gradients: 0.22160107612379984
l2 norm of weights: 2.574164481756889

Iteration Number: 3674
Loss: 17.50445222188914
l2 norm of gradients: 0.22150472517099362

```
l2 norm of weights: 2.573807576126209
-----
Iteration Number: 3675
Loss: 17.499499001923475
l2 norm of gradients: 0.2214084282931041
l2 norm of weights: 2.573450931436417
-----
Iteration Number: 3676
Loss: 17.49454962487747
l2 norm of gradients: 0.22131218547208772
l2 norm of weights: 2.5730945475589855
-----
Iteration Number: 3677
Loss: 17.489604087765077
l2 norm of gradients: 0.22121599668987454
l2 norm of weights: 2.5727384243653977
-----
Iteration Number: 3678
Loss: 17.484662387601976
l2 norm of gradients: 0.22111986192836802
l2 norm of weights: 2.5723825617271445
-----
Iteration Number: 3679
Loss: 17.479724521405597
l2 norm of gradients: 0.22102378116944504
l2 norm of weights: 2.572026959515729
-----
Iteration Number: 3680
Loss: 17.47479048619509
l2 norm of gradients: 0.22092775439495596
l2 norm of weights: 2.5716716176026635
-----
Iteration Number: 3681
Loss: 17.46986027899139
l2 norm of gradients: 0.22083178158672476
l2 norm of weights: 2.5713165358594714
-----
Iteration Number: 3682
Loss: 17.464933896817122
l2 norm of gradients: 0.2207358627265488
l2 norm of weights: 2.570961714157685
-----
Iteration Number: 3683
Loss: 17.460011336696695
l2 norm of gradients: 0.22063999779619906
l2 norm of weights: 2.570607152368849
-----
Iteration Number: 3684
```

Loss: 17.455092595656232
l2 norm of gradients: 0.22054418677742002
l2 norm of weights: 2.570252850364516

Iteration Number: 3685
Loss: 17.450177670723622
l2 norm of gradients: 0.22044842965192996
l2 norm of weights: 2.569898808016252

Iteration Number: 3686
Loss: 17.445266558928484
l2 norm of gradients: 0.22035272640142045
l2 norm of weights: 2.569545025195631

Iteration Number: 3687
Loss: 17.44035925730219
l2 norm of gradients: 0.2202570770075571
l2 norm of weights: 2.569191501774239

Iteration Number: 3688
Loss: 17.43545576287784
l2 norm of gradients: 0.22016148145197875
l2 norm of weights: 2.5688382376236722

Iteration Number: 3689
Loss: 17.430556072690308
l2 norm of gradients: 0.22006593971629837
l2 norm of weights: 2.5684852326155374

Iteration Number: 3690
Loss: 17.425660183776188
l2 norm of gradients: 0.2199704517821023
l2 norm of weights: 2.5681324866214528

Iteration Number: 3691
Loss: 17.42076809317383
l2 norm of gradients: 0.21987501763095077
l2 norm of weights: 2.567779999513047

Iteration Number: 3692
Loss: 17.415879797923328
l2 norm of gradients: 0.2197796372443778
l2 norm of weights: 2.5674277711619586

Iteration Number: 3693
Loss: 17.410995295066535
l2 norm of gradients: 0.21968431060389107
l2 norm of weights: 2.567075801439839


```

-----
Iteration Number: 3694
Loss: 17.40611458164703
l2 norm of gradients: 0.21958903769097227
l2 norm of weights: 2.566724090218349
-----
Iteration Number: 3695
Loss: 17.401237654710158
l2 norm of gradients: 0.2194938184870768
l2 norm of weights: 2.5663726373691613
-----
Iteration Number: 3696
Loss: 17.39636451130301
l2 norm of gradients: 0.21939865297363395
l2 norm of weights: 2.5660214427639594
-----
Iteration Number: 3697
Loss: 17.39149514847441
l2 norm of gradients: 0.2193035411320469
l2 norm of weights: 2.5656705062744374
-----
Iteration Number: 3698
Loss: 17.386629563274948
l2 norm of gradients: 0.21920848294369277
l2 norm of weights: 2.565319827772302
-----
Iteration Number: 3699
Loss: 17.381767752756964
l2 norm of gradients: 0.21911347838992265
l2 norm of weights: 2.5649694071292695
-----
Iteration Number: 3700
Loss: 17.376909713974538
l2 norm of gradients: 0.21901852745206166
l2 norm of weights: 2.5646192442170688
-----
Iteration Number: 3701
Loss: 17.372055443983506
l2 norm of gradients: 0.21892363011140878
l2 norm of weights: 2.56426933890744
-----
Iteration Number: 3702
Loss: 17.367204939841457
l2 norm of gradients: 0.21882878634923708
l2 norm of weights: 2.563919691072133
-----
Iteration Number: 3703
Loss: 17.36235819860772

```

```

12 norm of gradients: 0.21873399614679384
12 norm of weights: 2.5635703005829122
-----
Iteration Number: 3704
Loss: 17.357515217343398
12 norm of gradients: 0.21863925948530025
12 norm of weights: 2.563221167311551
-----
Iteration Number: 3705
Loss: 17.35267599311132
12 norm of gradients: 0.21854457634595162
12 norm of weights: 2.5628722911298345
-----
Iteration Number: 3706
Loss: 17.34784052297608
12 norm of gradients: 0.21844994670991755
12 norm of weights: 2.5625236719095614
-----
Iteration Number: 3707
Loss: 17.343008804004036
12 norm of gradients: 0.21835537055834178
12 norm of weights: 2.5621753095225404
-----
Iteration Number: 3708
Loss: 17.338180833263284
12 norm of gradients: 0.21826084787234215
12 norm of weights: 2.5618272038405925
-----
Iteration Number: 3709
Loss: 17.333356607823674
12 norm of gradients: 0.2181663786330109
12 norm of weights: 2.5614793547355497
-----
Iteration Number: 3710
Loss: 17.32853612475682
12 norm of gradients: 0.2180719628214144
12 norm of weights: 2.561131762079257
-----
Iteration Number: 3711
Loss: 17.323719381136076
12 norm of gradients: 0.21797760041859343
12 norm of weights: 2.5607844257435715
-----
Iteration Number: 3712
Loss: 17.31890637403657
12 norm of gradients: 0.21788329140556303
12 norm of weights: 2.560437345600361
-----

```

Iteration Number: 3713
Loss: 17.31409710053517
l2 norm of gradients: 0.21778903576331268
l2 norm of weights: 2.5600905215215057

Iteration Number: 3714
Loss: 17.30929155771051
l2 norm of gradients: 0.2176948334728061
l2 norm of weights: 2.5597439533788986

Iteration Number: 3715
Loss: 17.304489742642968
l2 norm of gradients: 0.2176006845149817
l2 norm of weights: 2.559397641044444

Iteration Number: 3716
Loss: 17.299691652414698
l2 norm of gradients: 0.21750658887075214
l2 norm of weights: 2.559051584390059

Iteration Number: 3717
Loss: 17.294897284109588
l2 norm of gradients: 0.2174125465210047
l2 norm of weights: 2.5587057832876727

Iteration Number: 3718
Loss: 17.290106634813306
l2 norm of gradients: 0.217318557446601
l2 norm of weights: 2.5583602376092265

Iteration Number: 3719
Loss: 17.285319701613254
l2 norm of gradients: 0.2172246216283774
l2 norm of weights: 2.558014947226674

Iteration Number: 3720
Loss: 17.280536481598617
l2 norm of gradients: 0.2171307390471448
l2 norm of weights: 2.5576699120119812

Iteration Number: 3721
Loss: 17.27575697186031
l2 norm of gradients: 0.21703690968368872
l2 norm of weights: 2.557325131837127

Iteration Number: 3722
Loss: 17.270981169491037
l2 norm of gradients: 0.21694313351876926

```

12 norm of weights: 2.556980606574102
-----
Iteration Number: 3723
Loss: 17.266209071585244
12 norm of gradients: 0.21684941053312143
12 norm of weights: 2.55663633609491
-----
Iteration Number: 3724
Loss: 17.26144067523914
12 norm of gradients: 0.21675574070745476
12 norm of weights: 2.5562923202715675
-----
Iteration Number: 3725
Loss: 17.256675977550692
12 norm of gradients: 0.21666212402245363
12 norm of weights: 2.555948558976103
-----
Iteration Number: 3726
Loss: 17.251914975619634
12 norm of gradients: 0.21656856045877723
12 norm of weights: 2.555605052080558
-----
Iteration Number: 3727
Loss: 17.247157666547455
12 norm of gradients: 0.21647504999705947
12 norm of weights: 2.555261799456988
-----
Iteration Number: 3728
Loss: 17.242404047437407
12 norm of gradients: 0.21638159261790926
12 norm of weights: 2.554918800977459
-----
Iteration Number: 3729
Loss: 17.2376541153945
12 norm of gradients: 0.21628818830191038
12 norm of weights: 2.554576056514051
-----
Iteration Number: 3730
Loss: 17.232907867525512
12 norm of gradients: 0.21619483702962147
12 norm of weights: 2.5542335659388584
-----
Iteration Number: 3731
Loss: 17.228165300938972
12 norm of gradients: 0.21610153878157626
12 norm of weights: 2.553891329123986
-----
Iteration Number: 3732

```

Loss: 17.223426412745198
l2 norm of gradients: 0.21600829353828346
l2 norm of weights: 2.553549345941553

Iteration Number: 3733
Loss: 17.218691200056238
l2 norm of gradients: 0.2159151012802267
l2 norm of weights: 2.5532076162636916

Iteration Number: 3734
Loss: 17.21395965998592
l2 norm of gradients: 0.2158219619878649
l2 norm of weights: 2.552866139962548

Iteration Number: 3735
Loss: 17.209231789649838
l2 norm of gradients: 0.21572887564163196
l2 norm of weights: 2.5525249169102797

Iteration Number: 3736
Loss: 17.204507586165345
l2 norm of gradients: 0.21563584222193696
l2 norm of weights: 2.552183946979059

Iteration Number: 3737
Loss: 17.199787046651558
l2 norm of gradients: 0.2155428617091642
l2 norm of weights: 2.5518432300410705

Iteration Number: 3738
Loss: 17.195070168229364
l2 norm of gradients: 0.21544993408367322
l2 norm of weights: 2.5515027659685123

Iteration Number: 3739
Loss: 17.190356948021407
l2 norm of gradients: 0.21535705932579885
l2 norm of weights: 2.5511625546335974

Iteration Number: 3740
Loss: 17.185647383152094
l2 norm of gradients: 0.21526423741585113
l2 norm of weights: 2.5508225959085506

Iteration Number: 3741
Loss: 17.180941470747626
l2 norm of gradients: 0.21517146833411555
l2 norm of weights: 2.55048288966561

```

-----
Iteration Number: 3742
Loss: 17.176239207935925
l2 norm of gradients: 0.21507875206085295
l2 norm of weights: 2.5501434357770294
-----
Iteration Number: 3743
Loss: 17.171540591846718
l2 norm of gradients: 0.21498608857629953
l2 norm of weights: 2.5498042341150735
-----
Iteration Number: 3744
Loss: 17.166845619611472
l2 norm of gradients: 0.21489347786066715
l2 norm of weights: 2.549465284552023
-----
Iteration Number: 3745
Loss: 17.162154288363443
l2 norm of gradients: 0.21480091989414296
l2 norm of weights: 2.54912658696017
-----
Iteration Number: 3746
Loss: 17.157466595237644
l2 norm of gradients: 0.2147084146568897
l2 norm of weights: 2.548788141211823
-----
Iteration Number: 3747
Loss: 17.152782537370847
l2 norm of gradients: 0.21461596212904582
l2 norm of weights: 2.5484499471793027
-----
Iteration Number: 3748
Loss: 17.148102111901608
l2 norm of gradients: 0.21452356229072525
l2 norm of weights: 2.548112004734944
-----
Iteration Number: 3749
Loss: 17.143425315970244
l2 norm of gradients: 0.21443121512201763
l2 norm of weights: 2.5477743137510958
-----
Iteration Number: 3750
Loss: 17.138752146718836
l2 norm of gradients: 0.2143389206029883
l2 norm of weights: 2.5474368741001205
-----
Iteration Number: 3751
Loss: 17.134082601291244

```

```

12 norm of gradients: 0.2142466787136784
12 norm of weights: 2.5470996856543957
-----
Iteration Number: 3752
Loss: 17.129416676833088
12 norm of gradients: 0.21415448943410473
12 norm of weights: 2.546762748286312
-----
Iteration Number: 3753
Loss: 17.124754370491768
12 norm of gradients: 0.21406235274426
12 norm of weights: 2.546426061868275
-----
Iteration Number: 3754
Loss: 17.12009567941644
12 norm of gradients: 0.21397026862411286
12 norm of weights: 2.5460896262727037
-----
Iteration Number: 3755
Loss: 17.11544060075805
12 norm of gradients: 0.21387823705360778
12 norm of weights: 2.5457534413720317
-----
Iteration Number: 3756
Loss: 17.11078913166929
12 norm of gradients: 0.2137862580126651
12 norm of weights: 2.545417507038707
-----
Iteration Number: 3757
Loss: 17.106141269304644
12 norm of gradients: 0.21369433148118128
12 norm of weights: 2.5450818231451926
-----
Iteration Number: 3758
Loss: 17.10149701082035
12 norm of gradients: 0.21360245743902875
12 norm of weights: 2.5447463895639646
-----
Iteration Number: 3759
Loss: 17.09685635337444
12 norm of gradients: 0.21351063586605604
12 norm of weights: 2.544411206167514
-----
Iteration Number: 3760
Loss: 17.09221929412669
12 norm of gradients: 0.21341886674208782
12 norm of weights: 2.544076272828347
-----

```

Iteration Number: 3761
Loss: 17.087585830238677
l2 norm of gradients: 0.21332715004692485
l2 norm of weights: 2.5437415894189837

Iteration Number: 3762
Loss: 17.082955958873722
l2 norm of gradients: 0.2132354857603442
l2 norm of weights: 2.5434071558119586

Iteration Number: 3763
Loss: 17.078329677196937
l2 norm of gradients: 0.2131438738620991
l2 norm of weights: 2.5430729718798224

Iteration Number: 3764
Loss: 17.0737069823752
l2 norm of gradients: 0.21305231433191904
l2 norm of weights: 2.5427390374951386

Iteration Number: 3765
Loss: 17.069087871577167
l2 norm of gradients: 0.21296080714951002
l2 norm of weights: 2.5424053525304866

Iteration Number: 3766
Loss: 17.064472341973264
l2 norm of gradients: 0.21286935229455412
l2 norm of weights: 2.5420719168584602

Iteration Number: 3767
Loss: 17.059860390735693
l2 norm of gradients: 0.21277794974671016
l2 norm of weights: 2.541738730351668

Iteration Number: 3768
Loss: 17.055252015038423
l2 norm of gradients: 0.21268659948561328
l2 norm of weights: 2.5414057928827347

Iteration Number: 3769
Loss: 17.05064721205721
l2 norm of gradients: 0.21259530149087502
l2 norm of weights: 2.5410731043242984

Iteration Number: 3770
Loss: 17.04604597896957
l2 norm of gradients: 0.21250405574208375


```
l2 norm of weights: 2.540740664549013
-----
Iteration Number: 3771
Loss: 17.04144831295481
l2 norm of gradients: 0.21241286221880412
l2 norm of weights: 2.540408473429548
-----
Iteration Number: 3772
Loss: 17.03685421119399
l2 norm of gradients: 0.21232172090057763
l2 norm of weights: 2.5400765308385855
-----
Iteration Number: 3773
Loss: 17.03226367086997
l2 norm of gradients: 0.21223063176692242
l2 norm of weights: 2.5397448366488273
-----
Iteration Number: 3774
Loss: 17.02767668916737
l2 norm of gradients: 0.21213959479733335
l2 norm of weights: 2.539413390732987
-----
Iteration Number: 3775
Loss: 17.023093263272592
l2 norm of gradients: 0.21204860997128214
l2 norm of weights: 2.539082192963794
-----
Iteration Number: 3776
Loss: 17.01851339037381
l2 norm of gradients: 0.21195767726821715
l2 norm of weights: 2.538751243213994
-----
Iteration Number: 3777
Loss: 17.013937067660976
l2 norm of gradients: 0.21186679666756375
l2 norm of weights: 2.538420541356347
-----
Iteration Number: 3778
Loss: 17.009364292325824
l2 norm of gradients: 0.2117759681487242
l2 norm of weights: 2.5380900872636296
-----
Iteration Number: 3779
Loss: 17.004795061561854
l2 norm of gradients: 0.21168519169107775
l2 norm of weights: 2.537759880808634
-----
Iteration Number: 3780
```

Loss: 17.000229372564355
l2 norm of gradients: 0.2115944672739805
l2 norm of weights: 2.537429921864166

Iteration Number: 3781
Loss: 16.99566722253039
l2 norm of gradients: 0.21150379487676593
l2 norm of weights: 2.537100210303049

Iteration Number: 3782
Loss: 16.991108608658784
l2 norm of gradients: 0.2114131744787442
l2 norm of weights: 2.5367707459981217

Iteration Number: 3783
Loss: 16.98655352815017
l2 norm of gradients: 0.21132260605920297
l2 norm of weights: 2.5364415288222384

Iteration Number: 3784
Loss: 16.982001978206934
l2 norm of gradients: 0.2112320895974069
l2 norm of weights: 2.5361125586482682

Iteration Number: 3785
Loss: 16.977453956033248
l2 norm of gradients: 0.21114162507259798
l2 norm of weights: 2.5357838353490973

Iteration Number: 3786
Loss: 16.972909458835062
l2 norm of gradients: 0.21105121246399544
l2 norm of weights: 2.5354553587976283

Iteration Number: 3787
Loss: 16.968368483820115
l2 norm of gradients: 0.2109608517507959
l2 norm of weights: 2.5351271288667774

Iteration Number: 3788
Loss: 16.963831028197905
l2 norm of gradients: 0.21087054291217328
l2 norm of weights: 2.5347991454294783

Iteration Number: 3789
Loss: 16.959297089179728
l2 norm of gradients: 0.210780285927279
l2 norm of weights: 2.534471408358682

```
-----
Iteration Number: 3790
Loss: 16.95476666397865
l2 norm of gradients: 0.21069008077524184
l2 norm of weights: 2.5341439175273526
-----
Iteration Number: 3791
Loss: 16.950239749809516
l2 norm of gradients: 0.2105999274351683
l2 norm of weights: 2.533816672808473
-----
Iteration Number: 3792
Loss: 16.945716343888964
l2 norm of gradients: 0.21050982588614234
l2 norm of weights: 2.53348967407504
-----
Iteration Number: 3793
Loss: 16.941196443435384
l2 norm of gradients: 0.21041977610722556
l2 norm of weights: 2.5331629212000695
-----
Iteration Number: 3794
Loss: 16.936680045668982
l2 norm of gradients: 0.21032977807745729
l2 norm of weights: 2.532836414056591
-----
Iteration Number: 3795
Loss: 16.932167147811718
l2 norm of gradients: 0.21023983177585434
l2 norm of weights: 2.532510152517651
-----
Iteration Number: 3796
Loss: 16.92765774708734
l2 norm of gradients: 0.21014993718141162
l2 norm of weights: 2.5321841364563147
-----
Iteration Number: 3797
Loss: 16.923151840721385
l2 norm of gradients: 0.21006009427310166
l2 norm of weights: 2.53185836574566
-----
Iteration Number: 3798
Loss: 16.918649425941158
l2 norm of gradients: 0.2099703030298749
l2 norm of weights: 2.531532840258784
-----
Iteration Number: 3799
Loss: 16.914150499975758
```

```

12 norm of gradients: 0.20988056343065975
12 norm of weights: 2.531207559868799
-----
Iteration Number: 3800
Loss: 16.909655060056053
12 norm of gradients: 0.20979087545436248
12 norm of weights: 2.530882524448835
-----
Iteration Number: 3801
Loss: 16.905163103414704
12 norm of gradients: 0.2097012390798674
12 norm of weights: 2.5305577338720378
-----
Iteration Number: 3802
Loss: 16.90067462728615
12 norm of gradients: 0.209611654286037
12 norm of weights: 2.5302331880115703
-----
Iteration Number: 3803
Loss: 16.89618962890661
12 norm of gradients: 0.20952212105171175
12 norm of weights: 2.529908886740612
-----
Iteration Number: 3804
Loss: 16.891708105514084
12 norm of gradients: 0.20943263935571035
12 norm of weights: 2.529584829932359
-----
Iteration Number: 3805
Loss: 16.88723005434836
12 norm of gradients: 0.2093432091768297
12 norm of weights: 2.5292610174600254
-----
Iteration Number: 3806
Loss: 16.88275547265101
12 norm of gradients: 0.20925383049384494
12 norm of weights: 2.52893744919684
-----
Iteration Number: 3807
Loss: 16.87828435766538
12 norm of gradients: 0.20916450328550953
12 norm of weights: 2.5286141250160514
-----
Iteration Number: 3808
Loss: 16.8738167066366
12 norm of gradients: 0.20907522753055535
12 norm of weights: 2.528291044790923
-----

```

Iteration Number: 3809
Loss: 16.8693525168116
l2 norm of gradients: 0.20898600320769273
l2 norm of weights: 2.5279682083947352

Iteration Number: 3810
Loss: 16.86489178543907
l2 norm of gradients: 0.2088968302956103
l2 norm of weights: 2.5276456157007883

Iteration Number: 3811
Loss: 16.860434509769497
l2 norm of gradients: 0.20880770877297528
l2 norm of weights: 2.527323266582396

Iteration Number: 3812
Loss: 16.85598068705515
l2 norm of gradients: 0.20871863861843357
l2 norm of weights: 2.527001160912892

Iteration Number: 3813
Loss: 16.851530314550086
l2 norm of gradients: 0.20862961981060954
l2 norm of weights: 2.5266792985656257

Iteration Number: 3814
Loss: 16.847083389510132
l2 norm of gradients: 0.20854065232810634
l2 norm of weights: 2.5263576794139655

Iteration Number: 3815
Loss: 16.84263990919291
l2 norm of gradients: 0.20845173614950582
l2 norm of weights: 2.526036303331295

Iteration Number: 3816
Loss: 16.83819987085783
l2 norm of gradients: 0.20836287125336853
l2 norm of weights: 2.5257151701910168

Iteration Number: 3817
Loss: 16.833763271766085
l2 norm of gradients: 0.20827405761823395
l2 norm of weights: 2.5253942798665507

Iteration Number: 3818
Loss: 16.829330109180635
l2 norm of gradients: 0.20818529522262028

```

12 norm of weights: 2.525073632231334
-----
Iteration Number: 3819
Loss: 16.82490038036625
12 norm of gradients: 0.20809658404502487
12 norm of weights: 2.524753227158821
-----
Iteration Number: 3820
Loss: 16.820474082589467
12 norm of gradients: 0.20800792406392393
12 norm of weights: 2.5244330645224844
-----
Iteration Number: 3821
Loss: 16.816051213118623
12 norm of gradients: 0.20791931525777263
12 norm of weights: 2.5241131441958142
-----
Iteration Number: 3822
Loss: 16.81163176922383
12 norm of gradients: 0.20783075760500552
12 norm of weights: 2.5237934660523185
-----
Iteration Number: 3823
Loss: 16.80721574817698
12 norm of gradients: 0.20774225108403582
12 norm of weights: 2.5234740299655227
-----
Iteration Number: 3824
Loss: 16.802803147251765
12 norm of gradients: 0.20765379567325637
12 norm of weights: 2.5231548358089704
-----
Iteration Number: 3825
Loss: 16.79839396372366
12 norm of gradients: 0.20756539135103896
12 norm of weights: 2.5228358834562226
-----
Iteration Number: 3826
Loss: 16.793988194869915
12 norm of gradients: 0.20747703809573495
12 norm of weights: 2.5225171727808595
-----
Iteration Number: 3827
Loss: 16.78958583796958
12 norm of gradients: 0.2073887358856748
12 norm of weights: 2.522198703656478
-----
Iteration Number: 3828

```

Loss: 16.78518689030347
l2 norm of gradients: 0.20730048469916848
l2 norm of weights: 2.5218804759566926

Iteration Number: 3829
Loss: 16.780791349154214
l2 norm of gradients: 0.2072122845145054
l2 norm of weights: 2.5215624895551385

Iteration Number: 3830
Loss: 16.776399211806215
l2 norm of gradients: 0.2071241353099546
l2 norm of weights: 2.521244744325466

Iteration Number: 3831
Loss: 16.772010475545645
l2 norm of gradients: 0.20703603706376444
l2 norm of weights: 2.520927240141346

Iteration Number: 3832
Loss: 16.767625137660495
l2 norm of gradients: 0.20694798975416312
l2 norm of weights: 2.5206099768764654

Iteration Number: 3833
Loss: 16.763243195440523
l2 norm of gradients: 0.20685999335935834
l2 norm of weights: 2.520292954404532

Iteration Number: 3834
Loss: 16.758864646177273
l2 norm of gradients: 0.20677204785753772
l2 norm of weights: 2.5199761725992693

Iteration Number: 3835
Loss: 16.754489487164076
l2 norm of gradients: 0.20668415322686837
l2 norm of weights: 2.5196596313344215

Iteration Number: 3836
Loss: 16.75011771569607
l2 norm of gradients: 0.20659630944549756
l2 norm of weights: 2.51934333048375

Iteration Number: 3837
Loss: 16.745749329070144
l2 norm of gradients: 0.2065085164915522
l2 norm of weights: 2.5190272699210348

```

-----
Iteration Number: 3838
Loss: 16.741384324585006
l2 norm of gradients: 0.20642077434313916
l2 norm of weights: 2.5187114495200746
-----
Iteration Number: 3839
Loss: 16.737022699541143
l2 norm of gradients: 0.20633308297834552
l2 norm of weights: 2.5183958691546873
-----
Iteration Number: 3840
Loss: 16.73266445124082
l2 norm of gradients: 0.20624544237523812
l2 norm of weights: 2.518080528698708
-----
Iteration Number: 3841
Loss: 16.728309576988096
l2 norm of gradients: 0.20615785251186414
l2 norm of weights: 2.5177654280259922
-----
Iteration Number: 3842
Loss: 16.72395807408882
l2 norm of gradients: 0.20607031336625076
l2 norm of weights: 2.517450567010414
-----
Iteration Number: 3843
Loss: 16.719609939850624
l2 norm of gradients: 0.20598282491640535
l2 norm of weights: 2.5171359455258644
-----
Iteration Number: 3844
Loss: 16.71526517158293
l2 norm of gradients: 0.2058953871403157
l2 norm of weights: 2.516821563446255
-----
Iteration Number: 3845
Loss: 16.71092376659695
l2 norm of gradients: 0.2058080000159498
l2 norm of weights: 2.5165074206455174
-----
Iteration Number: 3846
Loss: 16.70658572220568
l2 norm of gradients: 0.2057206635212561
l2 norm of weights: 2.516193516997599
-----
Iteration Number: 3847
Loss: 16.7022510357239

```



```

12 norm of gradients: 0.20563337763416337
12 norm of weights: 2.5158798523764685
-----
Iteration Number: 3848
Loss: 16.697919704468198
12 norm of gradients: 0.20554614233258095
12 norm of weights: 2.5155664266561137
-----
Iteration Number: 3849
Loss: 16.693591725756924
12 norm of gradients: 0.20545895759439867
12 norm of weights: 2.5152532397105407
-----
Iteration Number: 3850
Loss: 16.689267096910232
12 norm of gradients: 0.20537182339748702
12 norm of weights: 2.5149402914137746
-----
Iteration Number: 3851
Loss: 16.68494581525006
12 norm of gradients: 0.20528473971969705
12 norm of weights: 2.514627581639861
-----
Iteration Number: 3852
Loss: 16.680627878100132
12 norm of gradients: 0.20519770653886063
12 norm of weights: 2.5143151102628636
-----
Iteration Number: 3853
Loss: 16.67631328278597
12 norm of gradients: 0.20511072383279028
12 norm of weights: 2.514002877156866
-----
Iteration Number: 3854
Loss: 16.67200202663488
12 norm of gradients: 0.20502379157927944
12 norm of weights: 2.513690882195971
-----
Iteration Number: 3855
Loss: 16.66769410697594
12 norm of gradients: 0.20493690975610232
12 norm of weights: 2.5133791252543007
-----
Iteration Number: 3856
Loss: 16.663389521140044
12 norm of gradients: 0.2048500783410141
12 norm of weights: 2.5130676062059973
-----

```

Iteration Number: 3857
Loss: 16.659088266459857
l2 norm of gradients: 0.204763297311751
l2 norm of weights: 2.512756324925222

Iteration Number: 3858
Loss: 16.65479034026984
l2 norm of gradients: 0.20467656664603012
l2 norm of weights: 2.5124452812861553

Iteration Number: 3859
Loss: 16.650495739906237
l2 norm of gradients: 0.2045898863215499
l2 norm of weights: 2.5121344751629984

Iteration Number: 3860
Loss: 16.64620446270709
l2 norm of gradients: 0.2045032563159897
l2 norm of weights: 2.5118239064299708

Iteration Number: 3861
Loss: 16.64191650601222
l2 norm of gradients: 0.2044166766070102
l2 norm of weights: 2.5115135749613136

Iteration Number: 3862
Loss: 16.637631867163236
l2 norm of gradients: 0.2043301471722534
l2 norm of weights: 2.511203480631286

Iteration Number: 3863
Loss: 16.63335054350355
l2 norm of gradients: 0.20424366798934251
l2 norm of weights: 2.510893623314168

Iteration Number: 3864
Loss: 16.62907253237835
l2 norm of gradients: 0.20415723903588207
l2 norm of weights: 2.5105840028842588

Iteration Number: 3865
Loss: 16.62479783113461
l2 norm of gradients: 0.20407086028945826
l2 norm of weights: 2.510274619215878

Iteration Number: 3866
Loss: 16.620526437121114
l2 norm of gradients: 0.20398453172763858

```

12 norm of weights: 2.5099654721833655
-----
Iteration Number: 3867
Loss: 16.616258347688408
12 norm of gradients: 0.20389825332797215
12 norm of weights: 2.5096565616610804
-----
Iteration Number: 3868
Loss: 16.611993560188846
12 norm of gradients: 0.20381202506798965
12 norm of weights: 2.509347887523403
-----
Iteration Number: 3869
Loss: 16.607732071976567
12 norm of gradients: 0.20372584692520346
12 norm of weights: 2.509039449644733
-----
Iteration Number: 3870
Loss: 16.603473880407495
12 norm of gradients: 0.20363971887710766
12 norm of weights: 2.5087312478994908
-----
Iteration Number: 3871
Loss: 16.599218982839343
12 norm of gradients: 0.20355364090117806
12 norm of weights: 2.5084232821621164
-----
Iteration Number: 3872
Loss: 16.594967376631615
12 norm of gradients: 0.20346761297487242
12 norm of weights: 2.50811555230707
-----
Iteration Number: 3873
Loss: 16.59071905914561
12 norm of gradients: 0.2033816350756302
12 norm of weights: 2.507808058208834
-----
Iteration Number: 3874
Loss: 16.58647402774441
12 norm of gradients: 0.203295707180873
12 norm of weights: 2.5075007997419085
-----
Iteration Number: 3875
Loss: 16.582232279792883
12 norm of gradients: 0.20320982926800438
12 norm of weights: 2.5071937767808166
-----
Iteration Number: 3876

```

Loss: 16.577993812657688
l2 norm of gradients: 0.2031240013144098
l2 norm of weights: 2.5068869892001

Iteration Number: 3877
Loss: 16.573758623707285
l2 norm of gradients: 0.20303822329745697
l2 norm of weights: 2.506580436874322

Iteration Number: 3878
Loss: 16.569526710311912
l2 norm of gradients: 0.20295249519449585
l2 norm of weights: 2.5062741196780665

Iteration Number: 3879
Loss: 16.565298069843593
l2 norm of gradients: 0.20286681698285847
l2 norm of weights: 2.505968037485938

Iteration Number: 3880
Loss: 16.561072699676146
l2 norm of gradients: 0.20278118863985917
l2 norm of weights: 2.5056621901725613

Iteration Number: 3881
Loss: 16.55685059718518
l2 norm of gradients: 0.20269561014279477
l2 norm of weights: 2.5053565776125817

Iteration Number: 3882
Loss: 16.55263175974809
l2 norm of gradients: 0.20261008146894435
l2 norm of weights: 2.505051199680667

Iteration Number: 3883
Loss: 16.548416184744067
l2 norm of gradients: 0.2025246025955696
l2 norm of weights: 2.5047460562515043

Iteration Number: 3884
Loss: 16.544203869554078
l2 norm of gradients: 0.20243917349991453
l2 norm of weights: 2.5044411471998016

Iteration Number: 3885
Loss: 16.539994811560895
l2 norm of gradients: 0.20235379415920587
l2 norm of weights: 2.5041364724002895

```

-----
Iteration Number: 3886
Loss: 16.535789008149067
l2 norm of gradients: 0.20226846455065295
l2 norm of weights: 2.503832031727718
-----
Iteration Number: 3887
Loss: 16.53158645670493
l2 norm of gradients: 0.20218318465144774
l2 norm of weights: 2.5035278250568584
-----
Iteration Number: 3888
Loss: 16.527387154616623
l2 norm of gradients: 0.20209795443876516
l2 norm of weights: 2.503223852262504
-----
Iteration Number: 3889
Loss: 16.52319109927407
l2 norm of gradients: 0.2020127738897627
l2 norm of weights: 2.5029201132194685
-----
Iteration Number: 3890
Loss: 16.518998288068964
l2 norm of gradients: 0.2019276429815807
l2 norm of weights: 2.502616607802587
-----
Iteration Number: 3891
Loss: 16.514808718394814
l2 norm of gradients: 0.20184256169134268
l2 norm of weights: 2.5023133358867167
-----
Iteration Number: 3892
Loss: 16.51062238764691
l2 norm of gradients: 0.20175752999615498
l2 norm of weights: 2.5020102973467346
-----
Iteration Number: 3893
Loss: 16.50643929322232
l2 norm of gradients: 0.20167254787310693
l2 norm of weights: 2.5017074920575406
-----
Iteration Number: 3894
Loss: 16.502259432519914
l2 norm of gradients: 0.20158761529927105
l2 norm of weights: 2.501404919894055
-----
Iteration Number: 3895
Loss: 16.498082802940345

```

```

12 norm of gradients: 0.20150273225170304
12 norm of weights: 2.5011025807312204
-----
Iteration Number: 3896
Loss: 16.49390940188605
12 norm of gradients: 0.20141789870744178
12 norm of weights: 2.500800474444
-----
Iteration Number: 3897
Loss: 16.489739226761262
12 norm of gradients: 0.2013331146435093
12 norm of weights: 2.5004986009073793
-----
Iteration Number: 3898
Loss: 16.485572274972004
12 norm of gradients: 0.20124838003691115
12 norm of weights: 2.500196959996366
-----
Iteration Number: 3899
Loss: 16.48140854392608
12 norm of gradients: 0.20116369486463617
12 norm of weights: 2.499895551585988
-----
Iteration Number: 3900
Loss: 16.47724803103309
12 norm of gradients: 0.20107905910365675
12 norm of weights: 2.499594375551296
-----
Iteration Number: 3901
Loss: 16.473090733704414
12 norm of gradients: 0.20099447273092869
12 norm of weights: 2.499293431767363
-----
Iteration Number: 3902
Loss: 16.468936649353232
12 norm of gradients: 0.2009099357233913
12 norm of weights: 2.498992720109282
-----
Iteration Number: 3903
Loss: 16.464785775394496
12 norm of gradients: 0.20082544805796768
12 norm of weights: 2.4986922404521694
-----
Iteration Number: 3904
Loss: 16.460638109244968
12 norm of gradients: 0.2007410097115646
12 norm of weights: 2.4983919926711637
-----

```

Iteration Number: 3905
Loss: 16.45649364832318
l2 norm of gradients: 0.20065662066107234
l2 norm of weights: 2.4980919766414247

Iteration Number: 3906
Loss: 16.45235239004946
l2 norm of gradients: 0.20057228088336532
l2 norm of weights: 2.4977921922381343

Iteration Number: 3907
Loss: 16.448214331845914
l2 norm of gradients: 0.20048799035530152
l2 norm of weights: 2.497492639336497

Iteration Number: 3908
Loss: 16.444079471136455
l2 norm of gradients: 0.20040374905372305
l2 norm of weights: 2.4971933178117394

Iteration Number: 3909
Loss: 16.43994780534677
l2 norm of gradients: 0.20031955695545592
l2 norm of weights: 2.4968942275391095

Iteration Number: 3910
Loss: 16.435819331904334
l2 norm of gradients: 0.20023541403731013
l2 norm of weights: 2.4965953683938786

Iteration Number: 3911
Loss: 16.431694048238416
l2 norm of gradients: 0.20015132027607987
l2 norm of weights: 2.49629674025134

Iteration Number: 3912
Loss: 16.427571951780063
l2 norm of gradients: 0.2000672756485434
l2 norm of weights: 2.495998342986809

Iteration Number: 3913
Loss: 16.42345303996213
l2 norm of gradients: 0.1999832801314633
l2 norm of weights: 2.4957001764756233

Iteration Number: 3914
Loss: 16.419337310219227
l2 norm of gradients: 0.19989933370158633

```
l2 norm of weights: 2.495402240593144
-----
Iteration Number: 3915
Loss: 16.415224759987787
l2 norm of gradients: 0.19981543633564358
l2 norm of weights: 2.4951045352147543
-----
Iteration Number: 3916
Loss: 16.411115386706
l2 norm of gradients: 0.19973158801035065
l2 norm of weights: 2.494807060215859
-----
Iteration Number: 3917
Loss: 16.407009187813863
l2 norm of gradients: 0.19964778870240746
l2 norm of weights: 2.494509815471887
-----
Iteration Number: 3918
Loss: 16.402906160753147
l2 norm of gradients: 0.19956403838849843
l2 norm of weights: 2.494212800858289
-----
Iteration Number: 3919
Loss: 16.398806302967422
l2 norm of gradients: 0.19948033704529275
l2 norm of weights: 2.4939160162505387
-----
Iteration Number: 3920
Loss: 16.39470961190204
l2 norm of gradients: 0.199396684649444
l2 norm of weights: 2.493619461524132
-----
Iteration Number: 3921
Loss: 16.39061608500413
l2 norm of gradients: 0.1993130811775906
l2 norm of weights: 2.493323136554589
-----
Iteration Number: 3922
Loss: 16.38652571972262
l2 norm of gradients: 0.19922952660635562
l2 norm of weights: 2.493027041217452
-----
Iteration Number: 3923
Loss: 16.382438513508227
l2 norm of gradients: 0.19914602091234712
l2 norm of weights: 2.492731175388285
-----
Iteration Number: 3924
```


Loss: 16.37835446381344
l2 norm of gradients: 0.19906256407215775
l2 norm of weights: 2.492435538942676

Iteration Number: 3925
Loss: 16.374273568092548
l2 norm of gradients: 0.19897915606236527
l2 norm of weights: 2.492140131756238

Iteration Number: 3926
Loss: 16.370195823801616
l2 norm of gradients: 0.19889579685953243
l2 norm of weights: 2.4918449537046037

Iteration Number: 3927
Loss: 16.366121228398498
l2 norm of gradients: 0.19881248644020694
l2 norm of weights: 2.4915500046634307

Iteration Number: 3928
Loss: 16.36204977934284
l2 norm of gradients: 0.1987292247809217
l2 norm of weights: 2.4912552845083997

Iteration Number: 3929
Loss: 16.357981474096064
l2 norm of gradients: 0.1986460118581947
l2 norm of weights: 2.4909607931152147

Iteration Number: 3930
Loss: 16.353916310121384
l2 norm of gradients: 0.1985628476485293
l2 norm of weights: 2.4906665303596025

Iteration Number: 3931
Loss: 16.349854284883804
l2 norm of gradients: 0.1984797321284139
l2 norm of weights: 2.4903724961173137

Iteration Number: 3932
Loss: 16.3457953958501
l2 norm of gradients: 0.19839666527432248
l2 norm of weights: 2.4900786902641223

Iteration Number: 3933
Loss: 16.341739640488843
l2 norm of gradients: 0.19831364706271426
l2 norm of weights: 2.489785112675826

```

-----
Iteration Number: 3934
Loss: 16.337687016270383
l2 norm of gradients: 0.198230677470034
l2 norm of weights: 2.4894917632282443
-----
Iteration Number: 3935
Loss: 16.333637520666866
l2 norm of gradients: 0.19814775647271207
l2 norm of weights: 2.489198641797223
-----
Iteration Number: 3936
Loss: 16.329591151152208
l2 norm of gradients: 0.1980648840471642
l2 norm of weights: 2.4889057482586296
-----
Iteration Number: 3937
Loss: 16.325547905202118
l2 norm of gradients: 0.197982060169792
l2 norm of weights: 2.4886130824883557
-----
Iteration Number: 3938
Loss: 16.32150778029409
l2 norm of gradients: 0.1978992848169826
l2 norm of weights: 2.4883206443623163
-----
Iteration Number: 3939
Loss: 16.317470773907402
l2 norm of gradients: 0.19781655796510902
l2 norm of weights: 2.4880284337564516
-----
Iteration Number: 3940
Loss: 16.31343688352311
l2 norm of gradients: 0.19773387959053004
l2 norm of weights: 2.4877364505467234
-----
Iteration Number: 3941
Loss: 16.30940610662406
l2 norm of gradients: 0.1976512496695903
l2 norm of weights: 2.487444694609119
-----
Iteration Number: 3942
Loss: 16.30537844069488
l2 norm of gradients: 0.19756866817862048
l2 norm of weights: 2.4871531658196497
-----
Iteration Number: 3943
Loss: 16.30135388322198

```

```

12 norm of gradients: 0.19748613509393717
12 norm of weights: 2.486861864054349
-----
Iteration Number: 3944
Loss: 16.297332431693565
12 norm of gradients: 0.19740365039184304
12 norm of weights: 2.4865707891892765
-----
Iteration Number: 3945
Loss: 16.293314083599604
12 norm of gradients: 0.19732121404862699
12 norm of weights: 2.4862799411005145
-----
Iteration Number: 3946
Loss: 16.289298836431858
12 norm of gradients: 0.19723882604056386
12 norm of weights: 2.4859893196641702
-----
Iteration Number: 3947
Loss: 16.285286687683875
12 norm of gradients: 0.19715648634391503
12 norm of weights: 2.4856989247563748
-----
Iteration Number: 3948
Loss: 16.281277634850976
12 norm of gradients: 0.19707419493492795
12 norm of weights: 2.485408756253283
-----
Iteration Number: 3949
Loss: 16.277271675430285
12 norm of gradients: 0.19699195178983653
12 norm of weights: 2.4851188140310754
-----
Iteration Number: 3950
Loss: 16.27326880692068
12 norm of gradients: 0.19690975688486118
12 norm of weights: 2.4848290979659544
-----
Iteration Number: 3951
Loss: 16.26926902682284
12 norm of gradients: 0.19682761019620856
12 norm of weights: 2.48453960793415
-----
Iteration Number: 3952
Loss: 16.265272332639224
12 norm of gradients: 0.19674551170007212
12 norm of weights: 2.4842503438119135
-----

```

Iteration Number: 3953
Loss: 16.261278721874067
l2 norm of gradients: 0.19666346137263188
l2 norm of weights: 2.483961305475523

Iteration Number: 3954
Loss: 16.25728819203339
l2 norm of gradients: 0.19658145919005432
l2 norm of weights: 2.4836724928012797

Iteration Number: 3955
Loss: 16.253300740624994
l2 norm of gradients: 0.1964995051284929
l2 norm of weights: 2.4833839056655105

Iteration Number: 3956
Loss: 16.249316365158457
l2 norm of gradients: 0.19641759916408763
l2 norm of weights: 2.483095543944565

Iteration Number: 3957
Loss: 16.24533506314515
l2 norm of gradients: 0.19633574127296557
l2 norm of weights: 2.4828074075148208

Iteration Number: 3958
Loss: 16.241356832098212
l2 norm of gradients: 0.1962539314312406
l2 norm of weights: 2.4825194962526766

Iteration Number: 3959
Loss: 16.237381669532567
l2 norm of gradients: 0.1961721696150135
l2 norm of weights: 2.482231810034558

Iteration Number: 3960
Loss: 16.233409572964923
l2 norm of gradients: 0.19609045580037224
l2 norm of weights: 2.4819443487369157

Iteration Number: 3961
Loss: 16.229440539913764
l2 norm of gradients: 0.19600878996339166
l2 norm of weights: 2.4816571122362237

Iteration Number: 3962
Loss: 16.225474567899358
l2 norm of gradients: 0.19592717208013397

```
12 norm of weights: 2.481370100408982
-----
Iteration Number: 3963
Loss: 16.22151165444374
12 norm of gradients: 0.19584560212664837
12 norm of weights: 2.4810833131317156
-----
Iteration Number: 3964
Loss: 16.217551797070747
12 norm of gradients: 0.1957640800789715
12 norm of weights: 2.4807967502809745
-----
Iteration Number: 3965
Loss: 16.21359499330598
12 norm of gradients: 0.19568260591312725
12 norm of weights: 2.4805104117333334
-----
Iteration Number: 3966
Loss: 16.20964124067681
12 norm of gradients: 0.19560117960512693
12 norm of weights: 2.4802242973653925
-----
Iteration Number: 3967
Loss: 16.20569053671241
12 norm of gradients: 0.19551980113096923
12 norm of weights: 2.479938407053777
-----
Iteration Number: 3968
Loss: 16.201742878943715
12 norm of gradients: 0.19543847046664042
12 norm of weights: 2.479652740675137
-----
Iteration Number: 3969
Loss: 16.197798264903447
12 norm of gradients: 0.19535718758811427
12 norm of weights: 2.47936729810615
-----
Iteration Number: 3970
Loss: 16.193856692126104
12 norm of gradients: 0.19527595247135227
12 norm of weights: 2.4790820792235144
-----
Iteration Number: 3971
Loss: 16.189918158147954
12 norm of gradients: 0.19519476509230352
12 norm of weights: 2.4787970839039595
-----
Iteration Number: 3972
```

Loss: 16.185982660507054
l2 norm of gradients: 0.19511362542690489
l2 norm of weights: 2.478512312024236

Iteration Number: 3973
Loss: 16.18205019674323
l2 norm of gradients: 0.19503253345108107
l2 norm of weights: 2.478227763461122

Iteration Number: 3974
Loss: 16.178120764398088
l2 norm of gradients: 0.19495148914074467
l2 norm of weights: 2.4779434380914194

Iteration Number: 3975
Loss: 16.174194361015022
l2 norm of gradients: 0.19487049247179616
l2 norm of weights: 2.4776593357919583

Iteration Number: 3976
Loss: 16.17027098413918
l2 norm of gradients: 0.19478954342012392
l2 norm of weights: 2.4773754564395927

Iteration Number: 3977
Loss: 16.166350631317503
l2 norm of gradients: 0.1947086419616046
l2 norm of weights: 2.4770917999112023

Iteration Number: 3978
Loss: 16.1624333000987
l2 norm of gradients: 0.19462778807210276
l2 norm of weights: 2.4768083660836937

Iteration Number: 3979
Loss: 16.15851898803327
l2 norm of gradients: 0.19454698172747123
l2 norm of weights: 2.4765251548339973

Iteration Number: 3980
Loss: 16.154607692673462
l2 norm of gradients: 0.19446622290355117
l2 norm of weights: 2.476242166039072

Iteration Number: 3981
Loss: 16.150699411573328
l2 norm of gradients: 0.19438551157617173
l2 norm of weights: 2.4759593995759

```

-----
Iteration Number: 3982
Loss: 16.14679414228868
l2 norm of gradients: 0.1943048477211507
l2 norm of weights: 2.475676855321491
-----
Iteration Number: 3983
Loss: 16.142891882377103
l2 norm of gradients: 0.1942242313142942
l2 norm of weights: 2.475394533152881
-----
Iteration Number: 3984
Loss: 16.13899262939797
l2 norm of gradients: 0.19414366233139685
l2 norm of weights: 2.4751124329471312
-----
Iteration Number: 3985
Loss: 16.13509638091241
l2 norm of gradients: 0.19406314074824157
l2 norm of weights: 2.4748305545813287
-----
Iteration Number: 3986
Loss: 16.131203134483336
l2 norm of gradients: 0.19398266654060028
l2 norm of weights: 2.4745488979325874
-----
Iteration Number: 3987
Loss: 16.12731288767544
l2 norm of gradients: 0.19390223968423326
l2 norm of weights: 2.4742674628780477
-----
Iteration Number: 3988
Loss: 16.12342563805518
l2 norm of gradients: 0.19382186015488964
l2 norm of weights: 2.473986249294876
-----
Iteration Number: 3989
Loss: 16.119541383190786
l2 norm of gradients: 0.1937415279283072
l2 norm of weights: 2.4737052570602636
-----
Iteration Number: 3990
Loss: 16.115660120652265
l2 norm of gradients: 0.19366124298021267
l2 norm of weights: 2.4734244860514307
-----
Iteration Number: 3991
Loss: 16.111781848011397

```

```

12 norm of gradients: 0.19358100528632166
12 norm of weights: 2.473143936145622
-----
Iteration Number: 3992
Loss: 16.107906562841727
12 norm of gradients: 0.19350081482233875
12 norm of weights: 2.472863607220109
-----
Iteration Number: 3993
Loss: 16.104034262718585
12 norm of gradients: 0.19342067156395745
12 norm of weights: 2.472583499152191
-----
Iteration Number: 3994
Loss: 16.10016494521906
12 norm of gradients: 0.19334057548686046
12 norm of weights: 2.4723036118191923
-----
Iteration Number: 3995
Loss: 16.096298607922012
12 norm of gradients: 0.1932605265667195
12 norm of weights: 2.4720239450984645
-----
Iteration Number: 3996
Loss: 16.092435248408094
12 norm of gradients: 0.19318052477919562
12 norm of weights: 2.471744498867386
-----
Iteration Number: 3997
Loss: 16.088574864259698
12 norm of gradients: 0.19310057009993906
12 norm of weights: 2.471465273003361
-----
Iteration Number: 3998
Loss: 16.08471745306101
12 norm of gradients: 0.19302066250458935
12 norm of weights: 2.471186267383822
-----
Iteration Number: 3999
Loss: 16.080863012397977
12 norm of gradients: 0.19294080196877547
12 norm of weights: 2.4709074818862273
-----
Iteration Number: 4000
Loss: 16.077011539858315
12 norm of gradients: 0.19286098846811586
12 norm of weights: 2.4706289163880624
-----

```


Iteration Number: 4001
Loss: 16.073163033031516
l2 norm of gradients: 0.19278122197821834
l2 norm of weights: 2.4703505707668394

Iteration Number: 4002
Loss: 16.06931748950883
l2 norm of gradients: 0.1927015024746804
l2 norm of weights: 2.4700724449000977

Iteration Number: 4003
Loss: 16.06547490688329
l2 norm of gradients: 0.19262182993308916
l2 norm of weights: 2.4697945386654037

Iteration Number: 4004
Loss: 16.061635282749684
l2 norm of gradients: 0.19254220432902128
l2 norm of weights: 2.4695168519403508

Iteration Number: 4005
Loss: 16.05779861470458
l2 norm of gradients: 0.1924626256380434
l2 norm of weights: 2.4692393846025595

Iteration Number: 4006
Loss: 16.053964900346305
l2 norm of gradients: 0.1923830938357117
l2 norm of weights: 2.468962136529678

Iteration Number: 4007
Loss: 16.05013413727496
l2 norm of gradients: 0.19230360889757245
l2 norm of weights: 2.4686851075993808

Iteration Number: 4008
Loss: 16.04630632309241
l2 norm of gradients: 0.19222417079916174
l2 norm of weights: 2.4684082976893698

Iteration Number: 4009
Loss: 16.04248145540229
l2 norm of gradients: 0.19214477951600567
l2 norm of weights: 2.4681317066773754

Iteration Number: 4010
Loss: 16.03865953180999
l2 norm of gradients: 0.19206543502362033

```
l2 norm of weights: 2.467855334441154
-----
Iteration Number: 4011
Loss: 16.034840549922684
l2 norm of gradients: 0.19198613729751202
l2 norm of weights: 2.4675791808584906
-----
Iteration Number: 4012
Loss: 16.031024507349308
l2 norm of gradients: 0.19190688631317715
l2 norm of weights: 2.4673032458071966
-----
Iteration Number: 4013
Loss: 16.027211401700544
l2 norm of gradients: 0.1918276820461023
l2 norm of weights: 2.4670275291651107
-----
Iteration Number: 4014
Loss: 16.02340123058887
l2 norm of gradients: 0.19174852447176444
l2 norm of weights: 2.4667520308101016
-----
Iteration Number: 4015
Loss: 16.019593991628504
l2 norm of gradients: 0.19166941356563086
l2 norm of weights: 2.466476750620063
-----
Iteration Number: 4016
Loss: 16.015789682435443
l2 norm of gradients: 0.19159034930315919
l2 norm of weights: 2.4662016884729177
-----
Iteration Number: 4017
Loss: 16.01198830062744
l2 norm of gradients: 0.19151133165979764
l2 norm of weights: 2.465926844246615
-----
Iteration Number: 4018
Loss: 16.008189843824013
l2 norm of gradients: 0.19143236061098487
l2 norm of weights: 2.4656522178191334
-----
Iteration Number: 4019
Loss: 16.00439430964645
l2 norm of gradients: 0.19135343613215017
l2 norm of weights: 2.4653778090684777
-----
Iteration Number: 4020
```

Loss: 16.000601695717798
l2 norm of gradients: 0.19127455819871345
l2 norm of weights: 2.4651036178726824

Iteration Number: 4021
Loss: 15.996811999662864
l2 norm of gradients: 0.19119572678608532
l2 norm of weights: 2.4648296441098094

Iteration Number: 4022
Loss: 15.993025219108219
l2 norm of gradients: 0.19111694186966727
l2 norm of weights: 2.464555887657947

Iteration Number: 4023
Loss: 15.989241351682198
l2 norm of gradients: 0.19103820342485145
l2 norm of weights: 2.4642823483952134

Iteration Number: 4024
Loss: 15.985460395014895
l2 norm of gradients: 0.190959511427021
l2 norm of weights: 2.4640090261997543

Iteration Number: 4025
Loss: 15.981682346738175
l2 norm of gradients: 0.19088086585155012
l2 norm of weights: 2.4637359209497434

Iteration Number: 4026
Loss: 15.977907204485644
l2 norm of gradients: 0.1908022666738038
l2 norm of weights: 2.4634630325233826

Iteration Number: 4027
Loss: 15.974134965892691
l2 norm of gradients: 0.19072371386913825
l2 norm of weights: 2.463190360798902

Iteration Number: 4028
Loss: 15.970365628596454
l2 norm of gradients: 0.1906452074129007
l2 norm of weights: 2.4629179056545603

Iteration Number: 4029
Loss: 15.966599190235819
l2 norm of gradients: 0.19056674728042972
l2 norm of weights: 2.462645666968644

```

-----
Iteration Number: 4030
Loss: 15.96283564845146
l2 norm of gradients: 0.19048833344705512
l2 norm of weights: 2.4623736446194693
-----
Iteration Number: 4031
Loss: 15.959075000885784
l2 norm of gradients: 0.19040996588809783
l2 norm of weights: 2.462101838485378
-----
Iteration Number: 4032
Loss: 15.955317245182966
l2 norm of gradients: 0.1903316445788704
l2 norm of weights: 2.461830248444744
-----
Iteration Number: 4033
Loss: 15.951562378988948
l2 norm of gradients: 0.1902533694946767
l2 norm of weights: 2.4615588743759673
-----
Iteration Number: 4034
Loss: 15.947810399951416
l2 norm of gradients: 0.19017514061081195
l2 norm of weights: 2.4612877161574773
-----
Iteration Number: 4035
Loss: 15.944061305719817
l2 norm of gradients: 0.19009695790256328
l2 norm of weights: 2.461016773667732
-----
Iteration Number: 4036
Loss: 15.94031509394536
l2 norm of gradients: 0.19001882134520912
l2 norm of weights: 2.460746046785218
-----
Iteration Number: 4037
Loss: 15.936571762281012
l2 norm of gradients: 0.18994073091401972
l2 norm of weights: 2.4604755353884498
-----
Iteration Number: 4038
Loss: 15.932831308381484
l2 norm of gradients: 0.18986268658425706
l2 norm of weights: 2.4602052393559726
-----
Iteration Number: 4039
Loss: 15.929093729903258

```

```

12 norm of gradients: 0.1897846883311749
12 norm of weights: 2.4599351585663594
-----
Iteration Number: 4040
Loss: 15.925359024504564
12 norm of gradients: 0.18970673613001876
12 norm of weights: 2.4596652928982112
-----
Iteration Number: 4041
Loss: 15.921627189845383
12 norm of gradients: 0.1896288299560263
12 norm of weights: 2.459395642230159
-----
Iteration Number: 4042
Loss: 15.917898223587455
12 norm of gradients: 0.18955096978442698
12 norm of weights: 2.4591262064408634
-----
Iteration Number: 4043
Loss: 15.914172123394284
12 norm of gradients: 0.18947315559044228
12 norm of weights: 2.4588569854090125
-----
Iteration Number: 4044
Loss: 15.910448886931105
12 norm of gradients: 0.1893953873492859
12 norm of weights: 2.4585879790133243
-----
Iteration Number: 4045
Loss: 15.90672851186493
12 norm of gradients: 0.1893176650361636
12 norm of weights: 2.458319187132546
-----
Iteration Number: 4046
Loss: 15.903010995864506
12 norm of gradients: 0.18923998862627348
12 norm of weights: 2.4580506096454537
-----
Iteration Number: 4047
Loss: 15.899296336600345
12 norm of gradients: 0.18916235809480572
12 norm of weights: 2.457782246430853
-----
Iteration Number: 4048
Loss: 15.895584531744703
12 norm of gradients: 0.1890847734169431
12 norm of weights: 2.457514097367578
-----

```

Iteration Number: 4049
Loss: 15.891875578971593
l2 norm of gradients: 0.18900723456786048
l2 norm of weights: 2.457246162334494

Iteration Number: 4050
Loss: 15.888169475956772
l2 norm of gradients: 0.1889297415227255
l2 norm of weights: 2.4569784412104934

Iteration Number: 4051
Loss: 15.884466220377764
l2 norm of gradients: 0.18885229425669806
l2 norm of weights: 2.4567109338744997

Iteration Number: 4052
Loss: 15.88076580991382
l2 norm of gradients: 0.18877489274493087
l2 norm of weights: 2.456443640205465

Iteration Number: 4053
Loss: 15.877068242245963
l2 norm of gradients: 0.18869753696256902
l2 norm of weights: 2.4561765600823713

Iteration Number: 4054
Loss: 15.873373515056944
l2 norm of gradients: 0.18862022688475058
l2 norm of weights: 2.45590969338423

Iteration Number: 4055
Loss: 15.86968162603129
l2 norm of gradients: 0.1885429624866061
l2 norm of weights: 2.4556430399900826

Iteration Number: 4056
Loss: 15.865992572855244
l2 norm of gradients: 0.1884657437432592
l2 norm of weights: 2.455376599778999

Iteration Number: 4057
Loss: 15.86230635321683
l2 norm of gradients: 0.18838857062982622
l2 norm of weights: 2.455110372630081

Iteration Number: 4058
Loss: 15.858622964805797
l2 norm of gradients: 0.18831144312141643

```

12 norm of weights: 2.4548443584224575
-----
Iteration Number: 4059
Loss: 15.854942405313647
12 norm of gradients: 0.18823436119313228
12 norm of weights: 2.45457855703529
-----
Iteration Number: 4060
Loss: 15.851264672433631
12 norm of gradients: 0.18815732482006894
12 norm of weights: 2.4543129683477676
-----
Iteration Number: 4061
Loss: 15.847589763860748
12 norm of gradients: 0.18808033397731516
12 norm of weights: 2.4540475922391107
-----
Iteration Number: 4062
Loss: 15.843917677291737
12 norm of gradients: 0.1880033886399525
12 norm of weights: 2.453782428588569
-----
Iteration Number: 4063
Loss: 15.840248410425088
12 norm of gradients: 0.1879264887830559
12 norm of weights: 2.4535174772754225
-----
Iteration Number: 4064
Loss: 15.836581960961034
12 norm of gradients: 0.18784963438169358
12 norm of weights: 2.453252738178981
-----
Iteration Number: 4065
Loss: 15.83291832660155
12 norm of gradients: 0.1877728254109271
12 norm of weights: 2.452988211178585
-----
Iteration Number: 4066
Loss: 15.82925750505036
12 norm of gradients: 0.18769606184581145
12 norm of weights: 2.452723896153604
-----
Iteration Number: 4067
Loss: 15.825599494012927
12 norm of gradients: 0.1876193436613951
12 norm of weights: 2.45245979298344
-----
Iteration Number: 4068

```

```

Loss: 15.821944291196456
l2 norm of gradients: 0.18754267083272008
l2 norm of weights: 2.452195901547522
-----
Iteration Number: 4069
Loss: 15.818291894309901
l2 norm of gradients: 0.187466043334822
l2 norm of weights: 2.4519322217253126
-----
Iteration Number: 4070
Loss: 15.814642301063957
l2 norm of gradients: 0.1873894611427302
l2 norm of weights: 2.4516687533963024
-----
Iteration Number: 4071
Loss: 15.81099550917105
l2 norm of gradients: 0.18731292423146756
l2 norm of weights: 2.451405496440014
-----
Iteration Number: 4072
Loss: 15.807351516345364
l2 norm of gradients: 0.18723643257605094
l2 norm of weights: 2.4511424507359982
-----
Iteration Number: 4073
Loss: 15.803710320302809
l2 norm of gradients: 0.18715998615149088
l2 norm of weights: 2.4508796161638395
-----
Iteration Number: 4074
Loss: 15.800071918761043
l2 norm of gradients: 0.1870835849327919
l2 norm of weights: 2.450616992603151
-----
Iteration Number: 4075
Loss: 15.79643630943946
l2 norm of gradients: 0.18700722889495247
l2 norm of weights: 2.450354579933576
-----
Iteration Number: 4076
Loss: 15.792803490059198
l2 norm of gradients: 0.18693091801296502
l2 norm of weights: 2.4500923780347903
-----
Iteration Number: 4077
Loss: 15.789173458343129
l2 norm of gradients: 0.18685465226181622
l2 norm of weights: 2.449830386786498

```



```

-----
Iteration Number: 4078
Loss: 15.785546212015861
l2 norm of gradients: 0.1867784316164866
l2 norm of weights: 2.449568606068436
-----
Iteration Number: 4079
Loss: 15.781921748803754
l2 norm of gradients: 0.18670225605195115
l2 norm of weights: 2.449307035760371
-----
Iteration Number: 4080
Loss: 15.77830006643488
l2 norm of gradients: 0.18662612554317903
l2 norm of weights: 2.4490456757421004
-----
Iteration Number: 4081
Loss: 15.774681162639078
l2 norm of gradients: 0.18655004006513365
l2 norm of weights: 2.448784525893454
-----
Iteration Number: 4082
Loss: 15.771065035147894
l2 norm of gradients: 0.18647399959277286
l2 norm of weights: 2.4485235860942898
-----
Iteration Number: 4083
Loss: 15.767451681694626
l2 norm of gradients: 0.18639800410104887
l2 norm of weights: 2.4482628562245
-----
Iteration Number: 4084
Loss: 15.763841100014313
l2 norm of gradients: 0.18632205356490855
l2 norm of weights: 2.4480023361640053
-----
Iteration Number: 4085
Loss: 15.76023328784371
l2 norm of gradients: 0.18624614795929306
l2 norm of weights: 2.4477420257927593
-----
Iteration Number: 4086
Loss: 15.756628242921318
l2 norm of gradients: 0.18617028725913848
l2 norm of weights: 2.4474819249907447
-----
Iteration Number: 4087
Loss: 15.753025962987374

```

```

12 norm of gradients: 0.18609447143937535
12 norm of weights: 2.4472220336379773
-----
Iteration Number: 4088
Loss: 15.749426445783843
12 norm of gradients: 0.186018700474929
12 norm of weights: 2.4469623516145034
-----
Iteration Number: 4089
Loss: 15.74582968905442
12 norm of gradients: 0.18594297434071955
12 norm of weights: 2.4467028788004006
-----
Iteration Number: 4090
Loss: 15.742235690544536
12 norm of gradients: 0.18586729301166202
12 norm of weights: 2.4464436150757773
-----
Iteration Number: 4091
Loss: 15.73864444800135
12 norm of gradients: 0.18579165646266624
12 norm of weights: 2.4461845603207744
-----
Iteration Number: 4092
Loss: 15.735055959173762
12 norm of gradients: 0.18571606466863713
12 norm of weights: 2.4459257144155635
-----
Iteration Number: 4093
Loss: 15.731470221812389
12 norm of gradients: 0.18564051760447453
12 norm of weights: 2.4456670772403477
-----
Iteration Number: 4094
Loss: 15.727887233669586
12 norm of gradients: 0.18556501524507343
12 norm of weights: 2.445408648675362
-----
Iteration Number: 4095
Loss: 15.724306992499438
12 norm of gradients: 0.18548955756532398
12 norm of weights: 2.4451504286008716
-----
Iteration Number: 4096
Loss: 15.720729496057753
12 norm of gradients: 0.18541414454011151
12 norm of weights: 2.444892416897176
-----

```

Iteration Number: 4097
Loss: 15.717154742102071
l2 norm of gradients: 0.1853387761443166
l2 norm of weights: 2.4446346134446038

Iteration Number: 4098
Loss: 15.713582728391664
l2 norm of gradients: 0.1852634523528152
l2 norm of weights: 2.4443770181235163

Iteration Number: 4099
Loss: 15.71001345268752
l2 norm of gradients: 0.18518817314047864
l2 norm of weights: 2.4441196308143067

Iteration Number: 4100
Loss: 15.706446912752362
l2 norm of gradients: 0.18511293848217367
l2 norm of weights: 2.4438624513973997

Iteration Number: 4101
Loss: 15.702883106350638
l2 norm of gradients: 0.18503774835276252
l2 norm of weights: 2.443605479753252

Iteration Number: 4102
Loss: 15.699322031248522
l2 norm of gradients: 0.18496260272710308
l2 norm of weights: 2.4433487157623524

Iteration Number: 4103
Loss: 15.695763685213912
l2 norm of gradients: 0.18488750158004877
l2 norm of weights: 2.4430921593052215

Iteration Number: 4104
Loss: 15.692208066016425
l2 norm of gradients: 0.18481244488644874
l2 norm of weights: 2.4428358102624115

Iteration Number: 4105
Loss: 15.688655171427415
l2 norm of gradients: 0.18473743262114792
l2 norm of weights: 2.442579668514507

Iteration Number: 4106
Loss: 15.685104999219947
l2 norm of gradients: 0.1846624647589869

```

12 norm of weights: 2.442323733942125
-----
Iteration Number: 4107
Loss: 15.68155754716881
12 norm of gradients: 0.18458754127480226
12 norm of weights: 2.442068006425914
-----
Iteration Number: 4108
Loss: 15.678012813050527
12 norm of gradients: 0.18451266214342651
12 norm of weights: 2.441812485846555
-----
Iteration Number: 4109
Loss: 15.67447079464333
12 norm of gradients: 0.18443782733968797
12 norm of weights: 2.4415571720847615
-----
Iteration Number: 4110
Loss: 15.670931489727176
12 norm of gradients: 0.1843630368384112
12 norm of weights: 2.4413020650212784
-----
Iteration Number: 4111
Loss: 15.667394896083737
12 norm of gradients: 0.18428829061441684
12 norm of weights: 2.441047164536884
-----
Iteration Number: 4112
Loss: 15.663861011496417
12 norm of gradients: 0.18421358864252146
12 norm of weights: 2.4407924705123882
-----
Iteration Number: 4113
Loss: 15.660329833750328
12 norm of gradients: 0.18413893089753808
12 norm of weights: 2.440537982828634
-----
Iteration Number: 4114
Loss: 15.656801360632315
12 norm of gradients: 0.184064317354276
12 norm of weights: 2.440283701366496
-----
Iteration Number: 4115
Loss: 15.65327558993092
12 norm of gradients: 0.18398974798754061
12 norm of weights: 2.4400296260068823
-----
Iteration Number: 4116

```

Loss: 15.64975251943642
l2 norm of gradients: 0.18391522277213396
l2 norm of weights: 2.439775756630732

Iteration Number: 4117
Loss: 15.646232146940802
l2 norm of gradients: 0.18384074168285444
l2 norm of weights: 2.439522093119019

Iteration Number: 4118
Loss: 15.642714470237772
l2 norm of gradients: 0.18376630469449692
l2 norm of weights: 2.439268635352748

Iteration Number: 4119
Loss: 15.63919948712275
l2 norm of gradients: 0.18369191178185287
l2 norm of weights: 2.4390153832129573

Iteration Number: 4120
Loss: 15.635687195392869
l2 norm of gradients: 0.18361756291971043
l2 norm of weights: 2.4387623365807176

Iteration Number: 4121
Loss: 15.632177592846984
l2 norm of gradients: 0.18354325808285432
l2 norm of weights: 2.4385094953371325

Iteration Number: 4122
Loss: 15.628670677285651
l2 norm of gradients: 0.1834689972460661
l2 norm of weights: 2.4382568593633387

Iteration Number: 4123
Loss: 15.625166446511155
l2 norm of gradients: 0.18339478038412413
l2 norm of weights: 2.4380044285405047

Iteration Number: 4124
Loss: 15.621664898327486
l2 norm of gradients: 0.18332060747180362
l2 norm of weights: 2.437752202749834

Iteration Number: 4125
Loss: 15.618166030540344
l2 norm of gradients: 0.18324647848387662
l2 norm of weights: 2.4375001818725606

```

-----
Iteration Number: 4126
Loss: 15.614669840957145
l2 norm of gradients: 0.1831723933951123
l2 norm of weights: 2.4372483657899537
-----
Iteration Number: 4127
Loss: 15.611176327387012
l2 norm of gradients: 0.1830983521802767
l2 norm of weights: 2.436996754383314
-----
Iteration Number: 4128
Loss: 15.60768548764078
l2 norm of gradients: 0.18302435481413312
l2 norm of weights: 2.4367453475339764
-----
Iteration Number: 4129
Loss: 15.604197319530995
l2 norm of gradients: 0.182950401271442
l2 norm of weights: 2.4364941451233078
-----
Iteration Number: 4130
Loss: 15.600711820871918
l2 norm of gradients: 0.18287649152696087
l2 norm of weights: 2.436243147032709
-----
Iteration Number: 4131
Loss: 15.597228989479504
l2 norm of gradients: 0.18280262555544463
l2 norm of weights: 2.4359923531436145
-----
Iteration Number: 4132
Loss: 15.593748823171426
l2 norm of gradients: 0.18272880333164548
l2 norm of weights: 2.4357417633374916
-----
Iteration Number: 4133
Loss: 15.590271319767064
l2 norm of gradients: 0.1826550248303131
l2 norm of weights: 2.4354913774958407
-----
Iteration Number: 4134
Loss: 15.586796477087503
l2 norm of gradients: 0.18258129002619447
l2 norm of weights: 2.4352411955001956
-----
Iteration Number: 4135
Loss: 15.583324292955533

```

```

12 norm of gradients: 0.1825075988940342
12 norm of weights: 2.434991217232124
-----
Iteration Number: 4136
Loss: 15.579854765195648
12 norm of gradients: 0.18243395140857435
12 norm of weights: 2.434741442573227
-----
Iteration Number: 4137
Loss: 15.576387891634056
12 norm of gradients: 0.18236034754455477
12 norm of weights: 2.4344918714051387
-----
Iteration Number: 4138
Loss: 15.572923670098655
12 norm of gradients: 0.18228678727671285
12 norm of weights: 2.4342425036095277
-----
Iteration Number: 4139
Loss: 15.56946209841906
12 norm of gradients: 0.1822132705797837
12 norm of weights: 2.4339933390680946
-----
Iteration Number: 4140
Loss: 15.566003174426573
12 norm of gradients: 0.1821397974285004
12 norm of weights: 2.433744377662576
-----
Iteration Number: 4141
Loss: 15.56254689595422
12 norm of gradients: 0.18206636779759378
12 norm of weights: 2.4334956192747397
-----
Iteration Number: 4142
Loss: 15.559093260836711
12 norm of gradients: 0.18199298166179254
12 norm of weights: 2.4332470637863888
-----
Iteration Number: 4143
Loss: 15.555642266910462
12 norm of gradients: 0.18191963899582345
12 norm of weights: 2.43299871107936
-----
Iteration Number: 4144
Loss: 15.55219391201359
12 norm of gradients: 0.18184633977441125
12 norm of weights: 2.432750561035523
-----

```

Iteration Number: 4145
Loss: 15.548748193985913
l2 norm of gradients: 0.18177308397227881
l2 norm of weights: 2.432502613536783

Iteration Number: 4146
Loss: 15.545305110668945
l2 norm of gradients: 0.18169987156414713
l2 norm of weights: 2.432254868465078

Iteration Number: 4147
Loss: 15.5418646599059
l2 norm of gradients: 0.18162670252473545
l2 norm of weights: 2.4320073257023793

Iteration Number: 4148
Loss: 15.53842683954169
l2 norm of gradients: 0.1815535768287612
l2 norm of weights: 2.431759985130694

Iteration Number: 4149
Loss: 15.534991647422924
l2 norm of gradients: 0.18148049445094025
l2 norm of weights: 2.431512846632061

Iteration Number: 4150
Loss: 15.531559081397907
l2 norm of gradients: 0.18140745536598676
l2 norm of weights: 2.4312659100885563

Iteration Number: 4151
Loss: 15.528129139316642
l2 norm of gradients: 0.18133445954861332
l2 norm of weights: 2.431019175382287

Iteration Number: 4152
Loss: 15.52470181903082
l2 norm of gradients: 0.18126150697353116
l2 norm of weights: 2.430772642395396

Iteration Number: 4153
Loss: 15.521277118393833
l2 norm of gradients: 0.1811885976154499
l2 norm of weights: 2.430526311010061

Iteration Number: 4154
Loss: 15.517855035260764
l2 norm of gradients: 0.18111573144907783


```

12 norm of weights: 2.430280181108492
-----
Iteration Number: 4155
Loss: 15.514435567488393
12 norm of gradients: 0.1810429084491219
12 norm of weights: 2.430034252572935
-----
Iteration Number: 4156
Loss: 15.511018712935188
12 norm of gradients: 0.18097012859028794
12 norm of weights: 2.4297885252856704
-----
Iteration Number: 4157
Loss: 15.507604469461311
12 norm of gradients: 0.18089739184728038
12 norm of weights: 2.429542999129011
-----
Iteration Number: 4158
Loss: 15.504192834928613
12 norm of gradients: 0.18082469819480254
12 norm of weights: 2.4292976739853067
-----
Iteration Number: 4159
Loss: 15.500783807200635
12 norm of gradients: 0.18075204760755667
12 norm of weights: 2.4290525497369404
-----
Iteration Number: 4160
Loss: 15.497377384142611
12 norm of gradients: 0.18067944006024403
12 norm of weights: 2.42880762626633
-----
Iteration Number: 4161
Loss: 15.493973563621461
12 norm of gradients: 0.18060687552756477
12 norm of weights: 2.4285629034559273
-----
Iteration Number: 4162
Loss: 15.4905723435058
12 norm of gradients: 0.18053435398421833
12 norm of weights: 2.42831838118822
-----
Iteration Number: 4163
Loss: 15.487173721665915
12 norm of gradients: 0.180461875404903
12 norm of weights: 2.4280740593457293
-----
Iteration Number: 4164

```

```

Loss: 15.4837776959738
l2 norm of gradients: 0.1803894397643166
l2 norm of weights: 2.4278299378110115
-----
Iteration Number: 4165
Loss: 15.48038426430312
l2 norm of gradients: 0.1803170470371559
l2 norm of weights: 2.427586016466658
-----
Iteration Number: 4166
Loss: 15.476993424529232
l2 norm of gradients: 0.18024469719811712
l2 norm of weights: 2.4273422951952943
-----
Iteration Number: 4167
Loss: 15.473605174529176
l2 norm of gradients: 0.18017239022189588
l2 norm of weights: 2.4270987738795817
-----
Iteration Number: 4168
Loss: 15.470219512181677
l2 norm of gradients: 0.18010012608318715
l2 norm of weights: 2.4268554524022155
-----
Iteration Number: 4169
Loss: 15.466836435367146
l2 norm of gradients: 0.18002790475668543
l2 norm of weights: 2.426612330645926
-----
Iteration Number: 4170
Loss: 15.46345594196767
l2 norm of gradients: 0.17995572621708464
l2 norm of weights: 2.4263694084934797
-----
Iteration Number: 4171
Loss: 15.460078029867026
l2 norm of gradients: 0.17988359043907853
l2 norm of weights: 2.426126685827677
-----
Iteration Number: 4172
Loss: 15.456702696950668
l2 norm of gradients: 0.17981149739736033
l2 norm of weights: 2.4258841625313528
-----
Iteration Number: 4173
Loss: 15.453329941105729
l2 norm of gradients: 0.179739447066623
l2 norm of weights: 2.4256418384873784

```

```

-----
Iteration Number: 4174
Loss: 15.449959760221024
l2 norm of gradients: 0.17966743942155927
l2 norm of weights: 2.4253997135786594
-----
Iteration Number: 4175
Loss: 15.446592152187051
l2 norm of gradients: 0.1795954744368618
l2 norm of weights: 2.425157787688138
-----
Iteration Number: 4176
Loss: 15.44322711489598
l2 norm of gradients: 0.17952355208722307
l2 norm of weights: 2.4249160606987887
-----
Iteration Number: 4177
Loss: 15.439864646241661
l2 norm of gradients: 0.1794516723473354
l2 norm of weights: 2.4246745324936243
-----
Iteration Number: 4178
Loss: 15.436504744119624
l2 norm of gradients: 0.1793798351918912
l2 norm of weights: 2.424433202955692
-----
Iteration Number: 4179
Loss: 15.433147406427075
l2 norm of gradients: 0.179308040595583
l2 norm of weights: 2.424192071968074
-----
Iteration Number: 4180
Loss: 15.429792631062892
l2 norm of gradients: 0.1792362885331034
l2 norm of weights: 2.423951139413887
-----
Iteration Number: 4181
Loss: 15.426440415927624
l2 norm of gradients: 0.1791645789791451
l2 norm of weights: 2.4237104051762848
-----
Iteration Number: 4182
Loss: 15.42309075892351
l2 norm of gradients: 0.17909291190840113
l2 norm of weights: 2.4234698691384566
-----
Iteration Number: 4183
Loss: 15.419743657954445

```

```

12 norm of gradients: 0.17902128729556466
12 norm of weights: 2.4232295311836256
-----
Iteration Number: 4184
Loss: 15.416399110926006
12 norm of gradients: 0.17894970511532945
12 norm of weights: 2.422989391195052
-----
Iteration Number: 4185
Loss: 15.413057115745444
12 norm of gradients: 0.17887816534238934
12 norm of weights: 2.422749449056032
-----
Iteration Number: 4186
Loss: 15.40971767032167
12 norm of gradients: 0.17880666795143899
12 norm of weights: 2.422509704649895
-----
Iteration Number: 4187
Loss: 15.406380772565281
12 norm of gradients: 0.17873521291717323
12 norm of weights: 2.4222701578600088
-----
Iteration Number: 4188
Loss: 15.40304642038853
12 norm of gradients: 0.17866380021428763
12 norm of weights: 2.4220308085697755
-----
Iteration Number: 4189
Loss: 15.399714611705347
12 norm of gradients: 0.17859242981747844
12 norm of weights: 2.4217916566626334
-----
Iteration Number: 4190
Loss: 15.396385344431327
12 norm of gradients: 0.17852110170144256
12 norm of weights: 2.421552702022057
-----
Iteration Number: 4191
Loss: 15.393058616483733
12 norm of gradients: 0.1784498158408775
12 norm of weights: 2.4213139445315557
-----
Iteration Number: 4192
Loss: 15.389734425781503
12 norm of gradients: 0.17837857221048167
12 norm of weights: 2.4210753840746753
-----

```

Iteration Number: 4193
Loss: 15.38641277024522
l2 norm of gradients: 0.17830737078495437
l2 norm of weights: 2.420837020534998

Iteration Number: 4194
Loss: 15.383093647797157
l2 norm of gradients: 0.17823621153899577
l2 norm of weights: 2.4205988537961414

Iteration Number: 4195
Loss: 15.379777056361238
l2 norm of gradients: 0.17816509444730705
l2 norm of weights: 2.4203608837417594

Iteration Number: 4196
Loss: 15.37646299386305
l2 norm of gradients: 0.1780940194845904
l2 norm of weights: 2.4201231102555414

Iteration Number: 4197
Loss: 15.373151458229849
l2 norm of gradients: 0.17802298662554894
l2 norm of weights: 2.4198855332212137

Iteration Number: 4198
Loss: 15.36984244739055
l2 norm of gradients: 0.17795199584488722
l2 norm of weights: 2.419648152522539

Iteration Number: 4199
Loss: 15.366535959275735
l2 norm of gradients: 0.17788104711731076
l2 norm of weights: 2.419410968043315

Iteration Number: 4200
Loss: 15.363231991817635
l2 norm of gradients: 0.17781014041752644
l2 norm of weights: 2.419173979667377

Iteration Number: 4201
Loss: 15.35993054295015
l2 norm of gradients: 0.17773927572024234
l2 norm of weights: 2.4189371872785945

Iteration Number: 4202
Loss: 15.356631610608837
l2 norm of gradients: 0.177668453000168

```

12 norm of weights: 2.4187005907608756
-----
Iteration Number: 4203
Loss: 15.353335192730917
12 norm of gradients: 0.17759767223201448
12 norm of weights: 2.418464189998164
-----
Iteration Number: 4204
Loss: 15.350041287255252
12 norm of gradients: 0.17752693339049405
12 norm of weights: 2.4182279848744397
-----
Iteration Number: 4205
Loss: 15.346749892122387
12 norm of gradients: 0.17745623645032077
12 norm of weights: 2.4179919752737185
-----
Iteration Number: 4206
Loss: 15.3434610052745
12 norm of gradients: 0.17738558138621008
12 norm of weights: 2.4177561610800535
-----
Iteration Number: 4207
Loss: 15.340174624655434
12 norm of gradients: 0.17731496817287926
12 norm of weights: 2.417520542177534
-----
Iteration Number: 4208
Loss: 15.336890748210683
12 norm of gradients: 0.17724439678504716
12 norm of weights: 2.4172851184502866
-----
Iteration Number: 4209
Loss: 15.333609373887404
12 norm of gradients: 0.17717386719743453
12 norm of weights: 2.417049889782473
-----
Iteration Number: 4210
Loss: 15.330330499634396
12 norm of gradients: 0.17710337938476373
12 norm of weights: 2.4168148560582923
-----
Iteration Number: 4211
Loss: 15.327054123402116
12 norm of gradients: 0.1770329333217592
12 norm of weights: 2.4165800171619813
-----
Iteration Number: 4212

```

Loss: 15.323780243142673
l2 norm of gradients: 0.17696252898314715
l2 norm of weights: 2.416345372977812

Iteration Number: 4213
Loss: 15.32050885680982
l2 norm of gradients: 0.17689216634365584
l2 norm of weights: 2.4161109233900935

Iteration Number: 4214
Loss: 15.317239962358972
l2 norm of gradients: 0.17682184537801562
l2 norm of weights: 2.415876668283172

Iteration Number: 4215
Loss: 15.31397355774718
l2 norm of gradients: 0.17675156606095888
l2 norm of weights: 2.415642607541431

Iteration Number: 4216
Loss: 15.310709640933151
l2 norm of gradients: 0.17668132836722017
l2 norm of weights: 2.41540874104929

Iteration Number: 4217
Loss: 15.30744820987724
l2 norm of gradients: 0.17661113227153624
l2 norm of weights: 2.4151750686912057

Iteration Number: 4218
Loss: 15.304189262541446
l2 norm of gradients: 0.17654097774864613
l2 norm of weights: 2.414941590351672

Iteration Number: 4219
Loss: 15.300932796889413
l2 norm of gradients: 0.17647086477329124
l2 norm of weights: 2.4147083059152203

Iteration Number: 4220
Loss: 15.297678810886435
l2 norm of gradients: 0.1764007933202152
l2 norm of weights: 2.414475215266417

Iteration Number: 4221
Loss: 15.294427302499441
l2 norm of gradients: 0.17633076336416423
l2 norm of weights: 2.4142423182898676

```

-----
Iteration Number: 4222
Loss: 15.291178269697012
l2 norm of gradients: 0.17626077487988692
l2 norm of weights: 2.414009614870214
-----
Iteration Number: 4223
Loss: 15.287931710449374
l2 norm of gradients: 0.17619082784213455
l2 norm of weights: 2.4137771048921355
-----
Iteration Number: 4224
Loss: 15.284687622728386
l2 norm of gradients: 0.17612092222566086
l2 norm of weights: 2.4135447882403485
-----
Iteration Number: 4225
Loss: 15.281446004507554
l2 norm of gradients: 0.17605105800522222
l2 norm of weights: 2.413312664799606
-----
Iteration Number: 4226
Loss: 15.278206853762024
l2 norm of gradients: 0.17598123515557793
l2 norm of weights: 2.413080734454699
-----
Iteration Number: 4227
Loss: 15.274970168468576
l2 norm of gradients: 0.17591145365148983
l2 norm of weights: 2.4128489970904563
-----
Iteration Number: 4228
Loss: 15.271735946605633
l2 norm of gradients: 0.17584171346772268
l2 norm of weights: 2.4126174525917428
-----
Iteration Number: 4229
Loss: 15.268504186153258
l2 norm of gradients: 0.1757720145790441
l2 norm of weights: 2.412386100843461
-----
Iteration Number: 4230
Loss: 15.265274885093149
l2 norm of gradients: 0.17570235696022463
l2 norm of weights: 2.412154941730552
-----
Iteration Number: 4231
Loss: 15.26204804140864

```



```

12 norm of gradients: 0.17563274058603784
12 norm of weights: 2.4119239751379924
-----
Iteration Number: 4232
Loss: 15.2588236530847
12 norm of gradients: 0.17556316543126033
12 norm of weights: 2.4116932009507988
-----
Iteration Number: 4233
Loss: 15.255601718107933
12 norm of gradients: 0.17549363147067173
12 norm of weights: 2.411462619054023
-----
Iteration Number: 4234
Loss: 15.252382234466573
12 norm of gradients: 0.17542413867905493
12 norm of weights: 2.4112322293327555
-----
Iteration Number: 4235
Loss: 15.249165200150493
12 norm of gradients: 0.17535468703119594
12 norm of weights: 2.411002031672125
-----
Iteration Number: 4236
Loss: 15.245950613151198
12 norm of gradients: 0.17528527650188414
12 norm of weights: 2.410772025957296
-----
Iteration Number: 4237
Loss: 15.242738471461816
12 norm of gradients: 0.175215907065912
12 norm of weights: 2.4105422120734725
-----
Iteration Number: 4238
Loss: 15.23952877307712
12 norm of gradients: 0.1751465786980757
12 norm of weights: 2.4103125899058955
-----
Iteration Number: 4239
Loss: 15.236321515993493
12 norm of gradients: 0.17507729137317463
12 norm of weights: 2.4100831593398437
-----
Iteration Number: 4240
Loss: 15.23311669820897
12 norm of gradients: 0.1750080450660116
12 norm of weights: 2.409853920260634
-----

```

Iteration Number: 4241
Loss: 15.229914317723187
l2 norm of gradients: 0.1749388397513932
l2 norm of weights: 2.40962487255362

Iteration Number: 4242
Loss: 15.226714372537433
l2 norm of gradients: 0.1748696754041295
l2 norm of weights: 2.4093960161041954

Iteration Number: 4243
Loss: 15.223516860654605
l2 norm of gradients: 0.17480055199903405
l2 norm of weights: 2.409167350797789

Iteration Number: 4244
Loss: 15.220321780079237
l2 norm of gradients: 0.17473146951092444
l2 norm of weights: 2.40893887651987

Iteration Number: 4245
Loss: 15.217129128817481
l2 norm of gradients: 0.17466242791462172
l2 norm of weights: 2.4087105931559445

Iteration Number: 4246
Loss: 15.213938904877107
l2 norm of gradients: 0.17459342718495083
l2 norm of weights: 2.408482500591556

Iteration Number: 4247
Loss: 15.210751106267525
l2 norm of gradients: 0.17452446729674073
l2 norm of weights: 2.4082545987122876

Iteration Number: 4248
Loss: 15.207565730999754
l2 norm of gradients: 0.1744555482248241
l2 norm of weights: 2.4080268874037594

Iteration Number: 4249
Loss: 15.204382777086435
l2 norm of gradients: 0.17438666994403776
l2 norm of weights: 2.4077993665516297

Iteration Number: 4250
Loss: 15.201202242541832
l2 norm of gradients: 0.17431783242922239

```

12 norm of weights: 2.4075720360415955
-----
Iteration Number: 4251
Loss: 15.198024125381826
12 norm of gradients: 0.17424903565522296
12 norm of weights: 2.4073448957593917
-----
Iteration Number: 4252
Loss: 15.194848423623926
12 norm of gradients: 0.1741802795968884
12 norm of weights: 2.4071179455907914
-----
Iteration Number: 4253
Loss: 15.191675135287241
12 norm of gradients: 0.1741115642290719
12 norm of weights: 2.4068911854216055
-----
Iteration Number: 4254
Loss: 15.188504258392515
12 norm of gradients: 0.17404288952663097
12 norm of weights: 2.406664615137685
-----
Iteration Number: 4255
Loss: 15.18533579096209
12 norm of gradients: 0.1739742554644273
12 norm of weights: 2.406438234624917
-----
Iteration Number: 4256
Loss: 15.18216973101994
12 norm of gradients: 0.17390566201732705
12 norm of weights: 2.406212043769228
-----
Iteration Number: 4257
Loss: 15.179006076591648
12 norm of gradients: 0.17383710916020073
12 norm of weights: 2.4059860424565835
-----
Iteration Number: 4258
Loss: 15.175844825704404
12 norm of gradients: 0.17376859686792326
12 norm of weights: 2.4057602305729864
-----
Iteration Number: 4259
Loss: 15.172685976387013
12 norm of gradients: 0.17370012511537422
12 norm of weights: 2.405534608004479
-----
Iteration Number: 4260

```

Loss: 15.1695295266699
l2 norm of gradients: 0.17363169387743765
l2 norm of weights: 2.4053091746371416

Iteration Number: 4261
Loss: 15.166375474585086
l2 norm of gradients: 0.1735633031290022
l2 norm of weights: 2.4050839303570934

Iteration Number: 4262
Loss: 15.163223818166216
l2 norm of gradients: 0.1734949528449614
l2 norm of weights: 2.4048588750504916

Iteration Number: 4263
Loss: 15.160074555448535
l2 norm of gradients: 0.1734266430002133
l2 norm of weights: 2.4046340086035327

Iteration Number: 4264
Loss: 15.1569276844689
l2 norm of gradients: 0.17335837356966075
l2 norm of weights: 2.4044093309024515

Iteration Number: 4265
Loss: 15.153783203265776
l2 norm of gradients: 0.1732901445282116
l2 norm of weights: 2.404184841833522

Iteration Number: 4266
Loss: 15.15064110987923
l2 norm of gradients: 0.17322195585077851
l2 norm of weights: 2.4039605412830563

Iteration Number: 4267
Loss: 15.147501402350938
l2 norm of gradients: 0.17315380751227896
l2 norm of weights: 2.403736429137406

Iteration Number: 4268
Loss: 15.144364078724182
l2 norm of gradients: 0.17308569948763566
l2 norm of weights: 2.4035125052829605

Iteration Number: 4269
Loss: 15.141229137043837
l2 norm of gradients: 0.17301763175177626
l2 norm of weights: 2.4032887696061493

```

-----
Iteration Number: 4270
Loss: 15.138096575356393
l2 norm of gradients: 0.17294960427963352
l2 norm of weights: 2.4030652219934394
-----
Iteration Number: 4271
Loss: 15.134966391709945
l2 norm of gradients: 0.17288161704614535
l2 norm of weights: 2.4028418623313383
-----
Iteration Number: 4272
Loss: 15.131838584154172
l2 norm of gradients: 0.17281367002625492
l2 norm of weights: 2.402618690506391
-----
Iteration Number: 4273
Loss: 15.128713150740369
l2 norm of gradients: 0.17274576319491072
l2 norm of weights: 2.402395706405183
-----
Iteration Number: 4274
Loss: 15.125590089521422
l2 norm of gradients: 0.1726778965270664
l2 norm of weights: 2.402172909914337
-----
Iteration Number: 4275
Loss: 15.122469398551818
l2 norm of gradients: 0.17261006999768116
l2 norm of weights: 2.4019503009205163
-----
Iteration Number: 4276
Loss: 15.119351075887636
l2 norm of gradients: 0.17254228358171947
l2 norm of weights: 2.4017278793104224
-----
Iteration Number: 4277
Loss: 15.116235119586564
l2 norm of gradients: 0.17247453725415152
l2 norm of weights: 2.401505644970796
-----
Iteration Number: 4278
Loss: 15.113121527707875
l2 norm of gradients: 0.1724068309899528
l2 norm of weights: 2.401283597788418
-----
Iteration Number: 4279
Loss: 15.110010298312435

```

```

12 norm of gradients: 0.17233916476410444
12 norm of weights: 2.4010617376501076
-----
Iteration Number: 4280
Loss: 15.106901429462715
12 norm of gradients: 0.17227153855159333
12 norm of weights: 2.4008400644427232
-----
Iteration Number: 4281
Loss: 15.103794919222768
12 norm of gradients: 0.17220395232741198
12 norm of weights: 2.400618578053162
-----
Iteration Number: 4282
Loss: 15.100690765658243
12 norm of gradients: 0.17213640606655856
12 norm of weights: 2.400397278368362
-----
Iteration Number: 4283
Loss: 15.097588966836383
12 norm of gradients: 0.17206889974403722
12 norm of weights: 2.4001761652752993
-----
Iteration Number: 4284
Loss: 15.094489520826016
12 norm of gradients: 0.1720014333348578
12 norm of weights: 2.3999552386609895
-----
Iteration Number: 4285
Loss: 15.091392425697563
12 norm of gradients: 0.17193400681403612
12 norm of weights: 2.3997344984124886
-----
Iteration Number: 4286
Loss: 15.08829767952303
12 norm of gradients: 0.17186662015659396
12 norm of weights: 2.3995139444168907
-----
Iteration Number: 4287
Loss: 15.085205280376014
12 norm of gradients: 0.1717992733375591
12 norm of weights: 2.3992935765613295
-----
Iteration Number: 4288
Loss: 15.082115226331695
12 norm of gradients: 0.1717319663319654
12 norm of weights: 2.3990733947329796
-----

```

Iteration Number: 4289
Loss: 15.079027515466837
l2 norm of gradients: 0.17166469911485283
l2 norm of weights: 2.398853398819054

Iteration Number: 4290
Loss: 15.075942145859797
l2 norm of gradients: 0.17159747166126751
l2 norm of weights: 2.398633588706805

Iteration Number: 4291
Loss: 15.07285911559051
l2 norm of gradients: 0.17153028394626182
l2 norm of weights: 2.398413964283525

Iteration Number: 4292
Loss: 15.069778422740495
l2 norm of gradients: 0.17146313594489437
l2 norm of weights: 2.3981945254365464

Iteration Number: 4293
Loss: 15.066700065392846
l2 norm of gradients: 0.17139602763223008
l2 norm of weights: 2.39797527205324

Iteration Number: 4294
Loss: 15.063624041632247
l2 norm of gradients: 0.17132895898334027
l2 norm of weights: 2.3977562040210185

Iteration Number: 4295
Loss: 15.060550349544965
l2 norm of gradients: 0.1712619299733028
l2 norm of weights: 2.397537321227332

Iteration Number: 4296
Loss: 15.057478987218829
l2 norm of gradients: 0.17119494057720178
l2 norm of weights: 2.3973186235596717

Iteration Number: 4297
Loss: 15.054409952743265
l2 norm of gradients: 0.17112799077012805
l2 norm of weights: 2.3971001109055683

Iteration Number: 4298
Loss: 15.051343244209262
l2 norm of gradients: 0.1710610805271789

```
l2 norm of weights: 2.3968817831525917
-----
Iteration Number: 4299
Loss: 15.0482788597094
l2 norm of gradients: 0.17099420982345842
l2 norm of weights: 2.3966636401883528
-----
Iteration Number: 4300
Loss: 15.045216797337817
l2 norm of gradients: 0.17092737863407712
l2 norm of weights: 2.3964456819005022
-----
Iteration Number: 4301
Loss: 15.042157055190241
l2 norm of gradients: 0.17086058693415263
l2 norm of weights: 2.396227908176729
-----
Iteration Number: 4302
Loss: 15.03909963136396
l2 norm of gradients: 0.17079383469880902
l2 norm of weights: 2.3960103189047643
-----
Iteration Number: 4303
Loss: 15.036044523957845
l2 norm of gradients: 0.17072712190317738
l2 norm of weights: 2.395792913972378
-----
Iteration Number: 4304
Loss: 15.032991731072332
l2 norm of gradients: 0.17066044852239562
l2 norm of weights: 2.3955756932673804
-----
Iteration Number: 4305
Loss: 15.029941250809433
l2 norm of gradients: 0.1705938145316087
l2 norm of weights: 2.395358656677621
-----
Iteration Number: 4306
Loss: 15.026893081272723
l2 norm of gradients: 0.17052721990596842
l2 norm of weights: 2.3951418040909904
-----
Iteration Number: 4307
Loss: 15.023847220567355
l2 norm of gradients: 0.1704606646206338
l2 norm of weights: 2.394925135395419
-----
Iteration Number: 4308
```


Loss: 15.020803666800038
l2 norm of gradients: 0.17039414865077077
l2 norm of weights: 2.394708650478878

Iteration Number: 4309
Loss: 15.017762418079053
l2 norm of gradients: 0.17032767197155257
l2 norm of weights: 2.3944923492293775

Iteration Number: 4310
Loss: 15.014723472514255
l2 norm of gradients: 0.1702612345581596
l2 norm of weights: 2.3942762315349686

Iteration Number: 4311
Loss: 15.011686828217048
l2 norm of gradients: 0.17019483638577948
l2 norm of weights: 2.394060297283742

Iteration Number: 4312
Loss: 15.008652483300413
l2 norm of gradients: 0.1701284774296071
l2 norm of weights: 2.3938445463638303

Iteration Number: 4313
Loss: 15.005620435878889
l2 norm of gradients: 0.1700621576648449
l2 norm of weights: 2.3936289786634037

Iteration Number: 4314
Loss: 15.002590684068576
l2 norm of gradients: 0.16999587706670238
l2 norm of weights: 2.393413594070676

Iteration Number: 4315
Loss: 14.999563225987137
l2 norm of gradients: 0.16992963561039684
l2 norm of weights: 2.3931983924738987

Iteration Number: 4316
Loss: 14.996538059753794
l2 norm of gradients: 0.16986343327115283
l2 norm of weights: 2.3929833737613646

Iteration Number: 4317
Loss: 14.993515183489327
l2 norm of gradients: 0.16979727002420264
l2 norm of weights: 2.392768537821407

```

-----
Iteration Number: 4318
Loss: 14.990494595316077
l2 norm of gradients: 0.16973114584478602
l2 norm of weights: 2.3925538845424006
-----
Iteration Number: 4319
Loss: 14.987476293357945
l2 norm of gradients: 0.16966506070815054
l2 norm of weights: 2.3923394138127585
-----
Iteration Number: 4320
Loss: 14.984460275740375
l2 norm of gradients: 0.1695990145895513
l2 norm of weights: 2.3921251255209364
-----
Iteration Number: 4321
Loss: 14.98144654059038
l2 norm of gradients: 0.16953300746425126
l2 norm of weights: 2.3919110195554287
-----
Iteration Number: 4322
Loss: 14.978435086036521
l2 norm of gradients: 0.1694670393075212
l2 norm of weights: 2.391697095804772
-----
Iteration Number: 4323
Loss: 14.975425910208921
l2 norm of gradients: 0.16940111009463973
l2 norm of weights: 2.3914833541575424
-----
Iteration Number: 4324
Loss: 14.972419011239234
l2 norm of gradients: 0.16933521980089325
l2 norm of weights: 2.3912697945023575
-----
Iteration Number: 4325
Loss: 14.969414387260692
l2 norm of gradients: 0.16926936840157636
l2 norm of weights: 2.3910564167278743
-----
Iteration Number: 4326
Loss: 14.966412036408057
l2 norm of gradients: 0.16920355587199154
l2 norm of weights: 2.3908432207227923
-----
Iteration Number: 4327
Loss: 14.963411956817648

```

```

12 norm of gradients: 0.16913778218744918
12 norm of weights: 2.39063020637585
-----
Iteration Number: 4328
Loss: 14.960414146627336
12 norm of gradients: 0.16907204732326808
12 norm of weights: 2.390417373575828
-----
Iteration Number: 4329
Loss: 14.957418603976535
12 norm of gradients: 0.169006351254775
12 norm of weights: 2.3902047222115463
-----
Iteration Number: 4330
Loss: 14.954425327006202
12 norm of gradients: 0.16894069395730493
12 norm of weights: 2.3899922521718677
-----
Iteration Number: 4331
Loss: 14.951434313858847
12 norm of gradients: 0.1688750754062011
12 norm of weights: 2.3897799633456938
-----
Iteration Number: 4332
Loss: 14.948445562678526
12 norm of gradients: 0.1688094955768151
12 norm of weights: 2.389567855621968
-----
Iteration Number: 4333
Loss: 14.945459071610824
12 norm of gradients: 0.16874395444450688
12 norm of weights: 2.389355928889674
-----
Iteration Number: 4334
Loss: 14.942474838802882
12 norm of gradients: 0.16867845198464468
12 norm of weights: 2.3891441830378386
-----
Iteration Number: 4335
Loss: 14.939492862403378
12 norm of gradients: 0.16861298817260537
12 norm of weights: 2.388932617955527
-----
Iteration Number: 4336
Loss: 14.936513140562536
12 norm of gradients: 0.16854756298377416
12 norm of weights: 2.3887212335318457
-----

```

Iteration Number: 4337
Loss: 14.933535671432105
l2 norm of gradients: 0.16848217639354499
l2 norm of weights: 2.388510029655944

Iteration Number: 4338
Loss: 14.930560453165391
l2 norm of gradients: 0.1684168283773201
l2 norm of weights: 2.3882990062170104

Iteration Number: 4339
Loss: 14.927587483917224
l2 norm of gradients: 0.16835151891051073
l2 norm of weights: 2.388088163104275

Iteration Number: 4340
Loss: 14.924616761843975
l2 norm of gradients: 0.16828624796853667
l2 norm of weights: 2.3878775002070105

Iteration Number: 4341
Loss: 14.92164828510355
l2 norm of gradients: 0.16822101552682625
l2 norm of weights: 2.3876670174145276

Iteration Number: 4342
Loss: 14.91868205185539
l2 norm of gradients: 0.168155821560817
l2 norm of weights: 2.3874567146161816

Iteration Number: 4343
Loss: 14.91571806026047
l2 norm of gradients: 0.1680906660459549
l2 norm of weights: 2.3872465917013663

Iteration Number: 4344
Loss: 14.9127563084813
l2 norm of gradients: 0.16802554895769514
l2 norm of weights: 2.3870366485595182

Iteration Number: 4345
Loss: 14.909796794681915
l2 norm of gradients: 0.16796047027150163
l2 norm of weights: 2.3868268850801146

Iteration Number: 4346
Loss: 14.906839517027887
l2 norm of gradients: 0.1678954299628474

```

12 norm of weights: 2.3866173011526746
-----
Iteration Number: 4347
Loss: 14.90388447368631
12 norm of gradients: 0.16783042800721445
12 norm of weights: 2.3864078966667575
-----
Iteration Number: 4348
Loss: 14.900931662825812
12 norm of gradients: 0.16776546438009382
12 norm of weights: 2.3861986715119645
-----
Iteration Number: 4349
Loss: 14.897981082616552
12 norm of gradients: 0.16770053905698587
12 norm of weights: 2.3859896255779387
-----
Iteration Number: 4350
Loss: 14.895032731230208
12 norm of gradients: 0.16763565201339986
12 norm of weights: 2.385780758754364
-----
Iteration Number: 4351
Loss: 14.892086606839985
12 norm of gradients: 0.16757080322485457
12 norm of weights: 2.385572070930966
-----
Iteration Number: 4352
Loss: 14.889142707620616
12 norm of gradients: 0.16750599266687785
12 norm of weights: 2.3853635619975107
-----
Iteration Number: 4353
Loss: 14.886201031748357
12 norm of gradients: 0.16744122031500688
12 norm of weights: 2.3851552318438074
-----
Iteration Number: 4354
Loss: 14.88326157740098
12 norm of gradients: 0.16737648614478842
12 norm of weights: 2.384947080359706
-----
Iteration Number: 4355
Loss: 14.880324342757786
12 norm of gradients: 0.16731179013177838
12 norm of weights: 2.3847391074350965
-----
Iteration Number: 4356

```

Loss: 14.877389325999598
l2 norm of gradients: 0.1672471322515423
l2 norm of weights: 2.3845313129599135

Iteration Number: 4357
Loss: 14.874456525308744
l2 norm of gradients: 0.16718251247965518
l2 norm of weights: 2.3843236968241306

Iteration Number: 4358
Loss: 14.871525938869087
l2 norm of gradients: 0.16711793079170165
l2 norm of weights: 2.3841162589177642

Iteration Number: 4359
Loss: 14.868597564866004
l2 norm of gradients: 0.16705338716327583
l2 norm of weights: 2.383908999130872

Iteration Number: 4360
Loss: 14.865671401486376
l2 norm of gradients: 0.1669888815699816
l2 norm of weights: 2.383701917353553

Iteration Number: 4361
Loss: 14.862747446918615
l2 norm of gradients: 0.16692441398743238
l2 norm of weights: 2.3834950134759487

Iteration Number: 4362
Loss: 14.859825699352637
l2 norm of gradients: 0.16685998439125163
l2 norm of weights: 2.3832882873882415

Iteration Number: 4363
Loss: 14.856906156979875
l2 norm of gradients: 0.16679559275707234
l2 norm of weights: 2.3830817389806564

Iteration Number: 4364
Loss: 14.853988817993276
l2 norm of gradients: 0.16673123906053744
l2 norm of weights: 2.382875368143459

Iteration Number: 4365
Loss: 14.851073680587296
l2 norm of gradients: 0.16666692327729973
l2 norm of weights: 2.3826691747669586

```

-----
Iteration Number: 4366
Loss: 14.848160742957896
l2 norm of gradients: 0.16660264538302205
l2 norm of weights: 2.3824631587415035
-----
Iteration Number: 4367
Loss: 14.84525000330256
l2 norm of gradients: 0.16653840535337702
l2 norm of weights: 2.3822573199574864
-----
Iteration Number: 4368
Loss: 14.842341459820267
l2 norm of gradients: 0.16647420316404749
l2 norm of weights: 2.3820516583053406
-----
Iteration Number: 4369
Loss: 14.839435110711507
l2 norm of gradients: 0.16641003879072622
l2 norm of weights: 2.381846173675542
-----
Iteration Number: 4370
Loss: 14.836530954178277
l2 norm of gradients: 0.16634591220911626
l2 norm of weights: 2.3816408659586075
-----
Iteration Number: 4371
Loss: 14.833628988424078
l2 norm of gradients: 0.16628182339493067
l2 norm of weights: 2.3814357350450965
-----
Iteration Number: 4372
Loss: 14.83072921165392
l2 norm of gradients: 0.16621777232389276
l2 norm of weights: 2.3812307808256103
-----
Iteration Number: 4373
Loss: 14.827831622074306
l2 norm of gradients: 0.16615375897173618
l2 norm of weights: 2.381026003190793
-----
Iteration Number: 4374
Loss: 14.824936217893253
l2 norm of gradients: 0.16608978331420488
l2 norm of weights: 2.3808214020313287
-----
Iteration Number: 4375
Loss: 14.822042997320262

```

```

12 norm of gradients: 0.16602584532705297
12 norm of weights: 2.3806169772379455
-----
Iteration Number: 4376
Loss: 14.819151958566351
12 norm of gradients: 0.16596194498604522
12 norm of weights: 2.380412728701413
-----
Iteration Number: 4377
Loss: 14.816263099844027
12 norm of gradients: 0.16589808226695668
12 norm of weights: 2.380208656312543
-----
Iteration Number: 4378
Loss: 14.8133764193673
12 norm of gradients: 0.1658342571455729
12 norm of weights: 2.380004759962188
-----
Iteration Number: 4379
Loss: 14.810491915351674
12 norm of gradients: 0.16577046959769004
12 norm of weights: 2.3798010395412446
-----
Iteration Number: 4380
Loss: 14.807609586014145
12 norm of gradients: 0.1657067195991147
12 norm of weights: 2.3795974949406506
-----
Iteration Number: 4381
Loss: 14.804729429573213
12 norm of gradients: 0.16564300712566432
12 norm of weights: 2.3793941260513867
-----
Iteration Number: 4382
Loss: 14.801851444248861
12 norm of gradients: 0.16557933215316675
12 norm of weights: 2.379190932764475
-----
Iteration Number: 4383
Loss: 14.798975628262568
12 norm of gradients: 0.1655156946574607
12 norm of weights: 2.37898791497098
-----
Iteration Number: 4384
Loss: 14.79610197983731
12 norm of gradients: 0.16545209461439567
12 norm of weights: 2.3787850725620086
-----

```


Iteration Number: 4385
Loss: 14.793230497197548
l2 norm of gradients: 0.16538853199983186
l2 norm of weights: 2.3785824054287104

Iteration Number: 4386
Loss: 14.790361178569237
l2 norm of gradients: 0.1653250067896404
l2 norm of weights: 2.3783799134622763

Iteration Number: 4387
Loss: 14.78749402217981
l2 norm of gradients: 0.16526151895970323
l2 norm of weights: 2.378177596553941

Iteration Number: 4388
Loss: 14.784629026258196
l2 norm of gradients: 0.1651980684859133
l2 norm of weights: 2.37797545459498

Iteration Number: 4389
Loss: 14.781766189034812
l2 norm of gradients: 0.16513465534417449
l2 norm of weights: 2.3777734874767122

Iteration Number: 4390
Loss: 14.77890550874155
l2 norm of gradients: 0.1650712795104017
l2 norm of weights: 2.3775716950904995

Iteration Number: 4391
Loss: 14.776046983611797
l2 norm of gradients: 0.165007940960521
l2 norm of weights: 2.377370077327744

Iteration Number: 4392
Loss: 14.77319061188042
l2 norm of gradients: 0.1649446396704694
l2 norm of weights: 2.3771686340798928

Iteration Number: 4393
Loss: 14.77033639178376
l2 norm of gradients: 0.1648813756161952
l2 norm of weights: 2.376967365238434

Iteration Number: 4394
Loss: 14.76748432155965
l2 norm of gradients: 0.16481814877365786

```

12 norm of weights: 2.3767662706948984
-----
Iteration Number: 4395
Loss: 14.764634399447392
12 norm of gradients: 0.16475495911882806
12 norm of weights: 2.376565350340859
-----
Iteration Number: 4396
Loss: 14.761786623687781
12 norm of gradients: 0.16469180662768784
12 norm of weights: 2.3763646040679327
-----
Iteration Number: 4397
Loss: 14.758940992523076
12 norm of gradients: 0.16462869127623042
12 norm of weights: 2.376164031767778
-----
Iteration Number: 4398
Loss: 14.75609750419702
12 norm of gradients: 0.16456561304046058
12 norm of weights: 2.375963633332096
-----
Iteration Number: 4399
Loss: 14.753256156954825
12 norm of gradients: 0.16450257189639436
12 norm of weights: 2.3757634086526305
-----
Iteration Number: 4400
Loss: 14.75041694904319
12 norm of gradients: 0.1644395678200594
12 norm of weights: 2.3755633576211683
-----
Iteration Number: 4401
Loss: 14.747579878710273
12 norm of gradients: 0.1643766007874947
12 norm of weights: 2.375363480129538
-----
Iteration Number: 4402
Loss: 14.744744944205717
12 norm of gradients: 0.16431367077475095
12 norm of weights: 2.375163776069612
-----
Iteration Number: 4403
Loss: 14.741912143780622
12 norm of gradients: 0.16425077775789024
12 norm of weights: 2.3749642453333046
-----
Iteration Number: 4404

```

Loss: 14.739081475687575
l2 norm of gradients: 0.16418792171298655
l2 norm of weights: 2.374764887812573

Iteration Number: 4405
Loss: 14.736252938180614
l2 norm of gradients: 0.16412510261612526
l2 norm of weights: 2.3745657033994183

Iteration Number: 4406
Loss: 14.733426529515262
l2 norm of gradients: 0.16406232044340366
l2 norm of weights: 2.3743666919858826

Iteration Number: 4407
Loss: 14.7306022479485
l2 norm of gradients: 0.16399957517093075
l2 norm of weights: 2.3741678534640513

Iteration Number: 4408
Loss: 14.72778009173878
l2 norm of gradients: 0.1639368667748273
l2 norm of weights: 2.373969187726054

Iteration Number: 4409
Loss: 14.724960059146003
l2 norm of gradients: 0.16387419523122596
l2 norm of weights: 2.3737706946640604

Iteration Number: 4410
Loss: 14.72214214843156
l2 norm of gradients: 0.16381156051627116
l2 norm of weights: 2.3735723741702865

Iteration Number: 4411
Loss: 14.719326357858286
l2 norm of gradients: 0.16374896260611946
l2 norm of weights: 2.373374226136989

Iteration Number: 4412
Loss: 14.716512685690482
l2 norm of gradients: 0.1636864014769392
l2 norm of weights: 2.373176250456467

Iteration Number: 4413
Loss: 14.71370113019391
l2 norm of gradients: 0.16362387710491086
l2 norm of weights: 2.372978447021065

```

-----
Iteration Number: 4414
Loss: 14.71089168963579
l2 norm of gradients: 0.16356138946622684
l2 norm of weights: 2.372780815723168
-----
Iteration Number: 4415
Loss: 14.708084362284808
l2 norm of gradients: 0.16349893853709174
l2 norm of weights: 2.3725833564552046
-----
Iteration Number: 4416
Loss: 14.705279146411097
l2 norm of gradients: 0.1634365242937223
l2 norm of weights: 2.3723860691096483
-----
Iteration Number: 4417
Loss: 14.70247604028625
l2 norm of gradients: 0.1633741467123474
l2 norm of weights: 2.3721889535790126
-----
Iteration Number: 4418
Loss: 14.699675042183317
l2 norm of gradients: 0.16331180576920812
l2 norm of weights: 2.3719920097558567
-----
Iteration Number: 4419
Loss: 14.6968761503768
l2 norm of gradients: 0.16324950144055772
l2 norm of weights: 2.3717952375327807
-----
Iteration Number: 4420
Loss: 14.694079363142656
l2 norm of gradients: 0.16318723370266203
l2 norm of weights: 2.37159863680243
-----
Iteration Number: 4421
Loss: 14.69128467875829
l2 norm of gradients: 0.16312500253179898
l2 norm of weights: 2.3714022074574905
-----
Iteration Number: 4422
Loss: 14.688492095502566
l2 norm of gradients: 0.16306280790425898
l2 norm of weights: 2.3712059493906934
-----
Iteration Number: 4423
Loss: 14.685701611655787

```

```

12 norm of gradients: 0.16300064979634474
12 norm of weights: 2.371009862494813
-----
Iteration Number: 4424
Loss: 14.682913225499712
12 norm of gradients: 0.16293852818437157
12 norm of weights: 2.370813946662665
-----
Iteration Number: 4425
Loss: 14.680126935317544
12 norm of gradients: 0.16287644304466728
12 norm of weights: 2.37061820178711
-----
Iteration Number: 4426
Loss: 14.677342739393938
12 norm of gradients: 0.16281439435357217
12 norm of weights: 2.3704226277610507
-----
Iteration Number: 4427
Loss: 14.674560636014986
12 norm of gradients: 0.16275238208743908
12 norm of weights: 2.3702272244774347
-----
Iteration Number: 4428
Loss: 14.671780623468228
12 norm of gradients: 0.16269040622263356
12 norm of weights: 2.3700319918292507
-----
Iteration Number: 4429
Loss: 14.669002700042652
12 norm of gradients: 0.16262846673553377
12 norm of weights: 2.3698369297095314
-----
Iteration Number: 4430
Loss: 14.666226864028681
12 norm of gradients: 0.1625665636025306
12 norm of weights: 2.3696420380113543
-----
Iteration Number: 4431
Loss: 14.66345311371818
12 norm of gradients: 0.16250469680002771
12 norm of weights: 2.3694473166278383
-----
Iteration Number: 4432
Loss: 14.66068144740446
12 norm of gradients: 0.16244286630444146
12 norm of weights: 2.3692527654521465
-----

```

Iteration Number: 4433
Loss: 14.657911863382267
l2 norm of gradients: 0.1623810720922011
l2 norm of weights: 2.3690583843774853

Iteration Number: 4434
Loss: 14.655144359947784
l2 norm of gradients: 0.16231931413974865
l2 norm of weights: 2.368864173297104

Iteration Number: 4435
Loss: 14.652378935398627
l2 norm of gradients: 0.16225759242353918
l2 norm of weights: 2.3686701321042962

Iteration Number: 4436
Loss: 14.649615588033852
l2 norm of gradients: 0.16219590692004052
l2 norm of weights: 2.368476260692398

Iteration Number: 4437
Loss: 14.646854316153952
l2 norm of gradients: 0.1621342576057337
l2 norm of weights: 2.36828255895479

Iteration Number: 4438
Loss: 14.644095118060852
l2 norm of gradients: 0.16207264445711247
l2 norm of weights: 2.368089026784895

Iteration Number: 4439
Loss: 14.641337992057899
l2 norm of gradients: 0.16201106745068394
l2 norm of weights: 2.36789566407618

Iteration Number: 4440
Loss: 14.638582936449893
l2 norm of gradients: 0.1619495265629681
l2 norm of weights: 2.3677024707221563

Iteration Number: 4441
Loss: 14.635829949543039
l2 norm of gradients: 0.16188802177049816
l2 norm of weights: 2.367509446616377

Iteration Number: 4442
Loss: 14.633079029644986
l2 norm of gradients: 0.16182655304982052

```

12 norm of weights: 2.3673165916524397
-----
Iteration Number: 4443
Loss: 14.63033017506481
12 norm of gradients: 0.16176512037749463
12 norm of weights: 2.3671239057239855
-----
Iteration Number: 4444
Loss: 14.62758338411301
12 norm of gradients: 0.16170372373009348
12 norm of weights: 2.3669313887246988
-----
Iteration Number: 4445
Loss: 14.624838655101513
12 norm of gradients: 0.16164236308420302
12 norm of weights: 2.366739040548308
-----
Iteration Number: 4446
Loss: 14.622095986343666
12 norm of gradients: 0.16158103841642274
12 norm of weights: 2.3665468610885854
-----
Iteration Number: 4447
Loss: 14.619355376154244
12 norm of gradients: 0.1615197497033654
12 norm of weights: 2.3663548502393454
-----
Iteration Number: 4448
Loss: 14.616616822849446
12 norm of gradients: 0.1614584969216572
12 norm of weights: 2.366163007894449
-----
Iteration Number: 4449
Loss: 14.613880324746889
12 norm of gradients: 0.16139728004793769
12 norm of weights: 2.3659713339477966
-----
Iteration Number: 4450
Loss: 14.611145880165608
12 norm of gradients: 0.16133609905886
12 norm of weights: 2.3657798282933364
-----
Iteration Number: 4451
Loss: 14.608413487426063
12 norm of gradients: 0.16127495393109073
12 norm of weights: 2.365588490825058
-----
Iteration Number: 4452

```

```
Loss: 14.605683144850122
l2 norm of gradients: 0.16121384464130997
l2 norm of weights: 2.3653973214369954
-----
Iteration Number: 4453
Loss: 14.602954850761083
l2 norm of gradients: 0.1611527711662115
l2 norm of weights: 2.3652063200232263
-----
Iteration Number: 4454
Loss: 14.600228603483655
l2 norm of gradients: 0.16109173348250258
l2 norm of weights: 2.3650154864778723
-----
Iteration Number: 4455
Loss: 14.597504401343947
l2 norm of gradients: 0.1610307315669042
l2 norm of weights: 2.3648248206950986
-----
Iteration Number: 4456
Loss: 14.594782242669504
l2 norm of gradients: 0.16096976539615104
l2 norm of weights: 2.364634322569114
-----
Iteration Number: 4457
Loss: 14.592062125789273
l2 norm of gradients: 0.1609088349469916
l2 norm of weights: 2.3644439919941718
-----
Iteration Number: 4458
Loss: 14.58934404903361
l2 norm of gradients: 0.16084794019618798
l2 norm of weights: 2.3642538288645687
-----
Iteration Number: 4459
Loss: 14.586628010734286
l2 norm of gradients: 0.16078708112051612
l2 norm of weights: 2.3640638330746446
-----
Iteration Number: 4460
Loss: 14.583914009224474
l2 norm of gradients: 0.16072625769676585
l2 norm of weights: 2.363874004518784
-----
Iteration Number: 4461
Loss: 14.581202042838765
l2 norm of gradients: 0.16066546990174088
l2 norm of weights: 2.3636843430914163
```



```
-----  
Iteration Number: 4462  
Loss: 14.578492109913146  
l2 norm of gradients: 0.16060471771225876  
l2 norm of weights: 2.3634948486870133  
-----  
Iteration Number: 4463  
Loss: 14.575784208785022  
l2 norm of gradients: 0.16054400110515105  
l2 norm of weights: 2.36330552120009  
-----  
Iteration Number: 4464  
Loss: 14.57307833779319  
l2 norm of gradients: 0.1604833200572632  
l2 norm of weights: 2.363116360525208  
-----  
Iteration Number: 4465  
Loss: 14.570374495277859  
l2 norm of gradients: 0.1604226745454548  
l2 norm of weights: 2.3629273665569706  
-----  
Iteration Number: 4466  
Loss: 14.567672679580632  
l2 norm of gradients: 0.16036206454659938  
l2 norm of weights: 2.3627385391900266  
-----  
Iteration Number: 4467  
Loss: 14.56497288904453  
l2 norm of gradients: 0.1603014900375846  
l2 norm of weights: 2.3625498783190677  
-----  
Iteration Number: 4468  
Loss: 14.56227512201395  
l2 norm of gradients: 0.16024095099531221  
l2 norm of weights: 2.3623613838388295  
-----  
Iteration Number: 4469  
Loss: 14.55957937683471  
l2 norm of gradients: 0.16018044739669812  
l2 norm of weights: 2.3621730556440927  
-----  
Iteration Number: 4470  
Loss: 14.55688565185401  
l2 norm of gradients: 0.16011997921867246  
l2 norm of weights: 2.361984893629681  
-----  
Iteration Number: 4471  
Loss: 14.554193945420456
```

```
12 norm of gradients: 0.16005954643817955
12 norm of weights: 2.3617968976904637
-----
Iteration Number: 4472
Loss: 14.55150425588405
12 norm of gradients: 0.15999914903217805
12 norm of weights: 2.361609067721352
-----
Iteration Number: 4473
Loss: 14.54881658159618
12 norm of gradients: 0.15993878697764072
12 norm of weights: 2.361421403617303
-----
Iteration Number: 4474
Loss: 14.546130920909636
12 norm of gradients: 0.1598784602515548
12 norm of weights: 2.361233905273317
-----
Iteration Number: 4475
Loss: 14.543447272178595
12 norm of gradients: 0.15981816883092193
12 norm of weights: 2.361046572584439
-----
Iteration Number: 4476
Loss: 14.540765633758632
12 norm of gradients: 0.15975791269275796
12 norm of weights: 2.360859405445757
-----
Iteration Number: 4477
Loss: 14.538086004006702
12 norm of gradients: 0.15969769181409332
12 norm of weights: 2.3606724037524054
-----
Iteration Number: 4478
Loss: 14.53540838128116
12 norm of gradients: 0.15963750617197284
12 norm of weights: 2.36048556739956
-----
Iteration Number: 4479
Loss: 14.532732763941734
12 norm of gradients: 0.15957735574345577
12 norm of weights: 2.360298896282443
-----
Iteration Number: 4480
Loss: 14.530059150349555
12 norm of gradients: 0.15951724050561614
12 norm of weights: 2.36011239029632
-----
```

Iteration Number: 4481
Loss: 14.52738753886713
l2 norm of gradients: 0.15945716043554212
l2 norm of weights: 2.359926049336501

Iteration Number: 4482
Loss: 14.524717927858354
l2 norm of gradients: 0.1593971155103369
l2 norm of weights: 2.3597398732983397

Iteration Number: 4483
Loss: 14.522050315688501
l2 norm of gradients: 0.159337105707118
l2 norm of weights: 2.3595538620772345

Iteration Number: 4484
Loss: 14.519384700724233
l2 norm of gradients: 0.1592771310030177
l2 norm of weights: 2.3593680155686285

Iteration Number: 4485
Loss: 14.516721081333593
l2 norm of gradients: 0.1592171913751831
l2 norm of weights: 2.359182333668009

Iteration Number: 4486
Loss: 14.514059455885992
l2 norm of gradients: 0.15915728680077568
l2 norm of weights: 2.358996816270906

Iteration Number: 4487
Loss: 14.511399822752232
l2 norm of gradients: 0.15909741725697205
l2 norm of weights: 2.3588114632728963

Iteration Number: 4488
Loss: 14.508742180304491
l2 norm of gradients: 0.15903758272096333
l2 norm of weights: 2.358626274569599

Iteration Number: 4489
Loss: 14.506086526916324
l2 norm of gradients: 0.15897778316995564
l2 norm of weights: 2.35844125005668

Iteration Number: 4490
Loss: 14.503432860962661
l2 norm of gradients: 0.15891801858116986

```
l2 norm of weights: 2.358256389629846
-----
Iteration Number: 4491
Loss: 14.500781180819796
l2 norm of gradients: 0.1588582889318417
l2 norm of weights: 2.3580716931848515
-----
Iteration Number: 4492
Loss: 14.498131484865416
l2 norm of gradients: 0.15879859419922188
l2 norm of weights: 2.3578871606174934
-----
Iteration Number: 4493
Loss: 14.495483771478558
l2 norm of gradients: 0.15873893436057604
l2 norm of weights: 2.357702791823614
-----
Iteration Number: 4494
Loss: 14.492838039039649
l2 norm of gradients: 0.15867930939318475
l2 norm of weights: 2.3575185866990997
-----
Iteration Number: 4495
Loss: 14.490194285930471
l2 norm of gradients: 0.15861971927434362
l2 norm of weights: 2.3573345451398815
-----
Iteration Number: 4496
Loss: 14.487552510534188
l2 norm of gradients: 0.15856016398136324
l2 norm of weights: 2.357150667041934
-----
Iteration Number: 4497
Loss: 14.48491271123532
l2 norm of gradients: 0.15850064349156948
l2 norm of weights: 2.3569669523012777
-----
Iteration Number: 4498
Loss: 14.482274886419763
l2 norm of gradients: 0.15844115778230294
l2 norm of weights: 2.356783400813977
-----
Iteration Number: 4499
Loss: 14.479639034474765
l2 norm of gradients: 0.15838170683091968
l2 norm of weights: 2.3566000124761404
-----
Iteration Number: 4500
```

Loss: 14.477005153788955
l2 norm of gradients: 0.15832229061479064
l2 norm of weights: 2.3564167871839214

Iteration Number: 4501
Loss: 14.474373242752312
l2 norm of gradients: 0.15826290911130228
l2 norm of weights: 2.356233724833518

Iteration Number: 4502
Loss: 14.471743299756188
l2 norm of gradients: 0.15820356229785598
l2 norm of weights: 2.356050825321172

Iteration Number: 4503
Loss: 14.469115323193279
l2 norm of gradients: 0.15814425015186845
l2 norm of weights: 2.355868088543172

Iteration Number: 4504
Loss: 14.466489311457655
l2 norm of gradients: 0.1580849726507718
l2 norm of weights: 2.355685514395848

Iteration Number: 4505
Loss: 14.463865262944742
l2 norm of gradients: 0.1580257297720132
l2 norm of weights: 2.3555031027755775

Iteration Number: 4506
Loss: 14.461243176051322
l2 norm of gradients: 0.15796652149305546
l2 norm of weights: 2.355320853578781

Iteration Number: 4507
Loss: 14.45862304917553
l2 norm of gradients: 0.15790734779137647
l2 norm of weights: 2.3551387667019235

Iteration Number: 4508
Loss: 14.456004880716867
l2 norm of gradients: 0.15784820864446975
l2 norm of weights: 2.3549568420415157

Iteration Number: 4509
Loss: 14.453388669076167
l2 norm of gradients: 0.157789104029844
l2 norm of weights: 2.3547750794941122

```

-----
Iteration Number: 4510
Loss: 14.450774412655637
l2 norm of gradients: 0.15773003392502358
l2 norm of weights: 2.354593478956313
-----
Iteration Number: 4511
Loss: 14.448162109858826
l2 norm of gradients: 0.15767099830754822
l2 norm of weights: 2.3544120403247617
-----
Iteration Number: 4512
Loss: 14.445551759090641
l2 norm of gradients: 0.1576119971549732
l2 norm of weights: 2.354230763496148
-----
Iteration Number: 4513
Loss: 14.44294335875733
l2 norm of gradients: 0.15755303044486937
l2 norm of weights: 2.3540496483672047
-----
Iteration Number: 4514
Loss: 14.440336907266492
l2 norm of gradients: 0.15749409815482304
l2 norm of weights: 2.3538686948347105
-----
Iteration Number: 4515
Loss: 14.437732403027072
l2 norm of gradients: 0.15743520026243618
l2 norm of weights: 2.3536879027954885
-----
Iteration Number: 4516
Loss: 14.435129844449367
l2 norm of gradients: 0.1573763367453264
l2 norm of weights: 2.3535072721464068
-----
Iteration Number: 4517
Loss: 14.432529229945011
l2 norm of gradients: 0.15731750758112684
l2 norm of weights: 2.3533268027843777
-----
Iteration Number: 4518
Loss: 14.429930557926992
l2 norm of gradients: 0.1572587127474865
l2 norm of weights: 2.353146494606359
-----
Iteration Number: 4519
Loss: 14.427333826809623

```

```

12 norm of gradients: 0.15719995222206987
12 norm of weights: 2.3529663475093527
-----
Iteration Number: 4520
Loss: 14.42473903500858
12 norm of gradients: 0.15714122598255736
12 norm of weights: 2.352786361390406
-----
Iteration Number: 4521
Loss: 14.422146180940864
12 norm of gradients: 0.15708253400664504
12 norm of weights: 2.3526065361466113
-----
Iteration Number: 4522
Loss: 14.419555263024822
12 norm of gradients: 0.15702387627204475
12 norm of weights: 2.3524268716751044
-----
Iteration Number: 4523
Loss: 14.416966279680134
12 norm of gradients: 0.1569652527564842
12 norm of weights: 2.3522473678730678
-----
Iteration Number: 4524
Loss: 14.414379229327823
12 norm of gradients: 0.1569066634377069
12 norm of weights: 2.352068024637727
-----
Iteration Number: 4525
Loss: 14.411794110390247
12 norm of gradients: 0.15684810829347215
12 norm of weights: 2.3518888418663546
-----
Iteration Number: 4526
Loss: 14.40921092129109
12 norm of gradients: 0.15678958730155532
12 norm of weights: 2.3517098194562656
-----
Iteration Number: 4527
Loss: 14.406629660455382
12 norm of gradients: 0.1567311004397475
12 norm of weights: 2.351530957304822
-----
Iteration Number: 4528
Loss: 14.404050326309479
12 norm of gradients: 0.15667264768585581
12 norm of weights: 2.3513522553094295
-----

```

Iteration Number: 4529
Loss: 14.401472917281069
l2 norm of gradients: 0.15661422901770344
l2 norm of weights: 2.35117371336754

Iteration Number: 4530
Loss: 14.398897431799165
l2 norm of gradients: 0.15655584441312928
l2 norm of weights: 2.350995331376648

Iteration Number: 4531
Loss: 14.39632386829412
l2 norm of gradients: 0.1564974938499886
l2 norm of weights: 2.3508171092342955

Iteration Number: 4532
Loss: 14.393752225197606
l2 norm of gradients: 0.15643917730615237
l2 norm of weights: 2.3506390468380682

Iteration Number: 4533
Loss: 14.391182500942625
l2 norm of gradients: 0.1563808947595079
l2 norm of weights: 2.350461144085597

Iteration Number: 4534
Loss: 14.3886146939635
l2 norm of gradients: 0.15632264618795838
l2 norm of weights: 2.3502834008745577

Iteration Number: 4535
Loss: 14.386048802695893
l2 norm of gradients: 0.15626443156942327
l2 norm of weights: 2.3501058171026714

Iteration Number: 4536
Loss: 14.38348482557677
l2 norm of gradients: 0.15620625088183812
l2 norm of weights: 2.349928392667704

Iteration Number: 4537
Loss: 14.380922761044431
l2 norm of gradients: 0.15614810410315466
l2 norm of weights: 2.349751127467466

Iteration Number: 4538
Loss: 14.378362607538493
l2 norm of gradients: 0.15608999121134068


```
l2 norm of weights: 2.349574021399814
-----
Iteration Number: 4539
Loss: 14.375804363499896
l2 norm of gradients: 0.1560319121843804
l2 norm of weights: 2.3493970743626496
-----
Iteration Number: 4540
Loss: 14.373248027370895
l2 norm of gradients: 0.15597386700027407
l2 norm of weights: 2.3492202862539178
-----
Iteration Number: 4541
Loss: 14.370693597595066
l2 norm of gradients: 0.15591585563703841
l2 norm of weights: 2.34904365697161
-----
Iteration Number: 4542
Loss: 14.368141072617298
l2 norm of gradients: 0.15585787807270624
l2 norm of weights: 2.3488671864137634
-----
Iteration Number: 4543
Loss: 14.3655904508838
l2 norm of gradients: 0.15579993428532674
l2 norm of weights: 2.348690874478459
-----
Iteration Number: 4544
Loss: 14.363041730842092
l2 norm of gradients: 0.15574202425296557
l2 norm of weights: 2.348514721063823
-----
Iteration Number: 4545
Loss: 14.360494910941005
l2 norm of gradients: 0.15568414795370447
l2 norm of weights: 2.348338726068028
-----
Iteration Number: 4546
Loss: 14.357949989630685
l2 norm of gradients: 0.15562630536564176
l2 norm of weights: 2.34816288938929
-----
Iteration Number: 4547
Loss: 14.355406965362592
l2 norm of gradients: 0.15556849646689205
l2 norm of weights: 2.3479872109258717
-----
Iteration Number: 4548
```

```

Loss: 14.35286583658949
l2 norm of gradients: 0.1555107212355865
l2 norm of weights: 2.34781169057608
-----
Iteration Number: 4549
Loss: 14.350326601765452
l2 norm of gradients: 0.1554529796498726
l2 norm of weights: 2.347636328238267
-----
Iteration Number: 4550
Loss: 14.347789259345861
l2 norm of gradients: 0.15539527168791437
l2 norm of weights: 2.347461123810831
-----
Iteration Number: 4551
Loss: 14.345253807787406
l2 norm of gradients: 0.15533759732789226
l2 norm of weights: 2.3472860771922135
-----
Iteration Number: 4552
Loss: 14.34272024554808
l2 norm of gradients: 0.15527995654800328
l2 norm of weights: 2.347111188280904
-----
Iteration Number: 4553
Loss: 14.340188571087184
l2 norm of gradients: 0.15522234932646095
l2 norm of weights: 2.3469364569754347
-----
Iteration Number: 4554
Loss: 14.337658782865308
l2 norm of gradients: 0.15516477564149536
l2 norm of weights: 2.346761883174384
-----
Iteration Number: 4555
Loss: 14.335130879344367
l2 norm of gradients: 0.1551072354713532
l2 norm of weights: 2.346587466776377
-----
Iteration Number: 4556
Loss: 14.332604858987555
l2 norm of gradients: 0.15504972879429763
l2 norm of weights: 2.346413207680081
-----
Iteration Number: 4557
Loss: 14.330080720259373
l2 norm of gradients: 0.1549922555886086
l2 norm of weights: 2.3462391057842114

```

```

-----
Iteration Number: 4558
Loss: 14.327558461625628
l2 norm of gradients: 0.15493481583258256
l2 norm of weights: 2.3460651609875267
-----
Iteration Number: 4559
Loss: 14.325038081553412
l2 norm of gradients: 0.15487740950453274
l2 norm of weights: 2.3458913731888322
-----
Iteration Number: 4560
Loss: 14.322519578511129
l2 norm of gradients: 0.154820036582789
l2 norm of weights: 2.345717742286978
-----
Iteration Number: 4561
Loss: 14.320002950968451
l2 norm of gradients: 0.15476269704569784
l2 norm of weights: 2.3455442681808596
-----
Iteration Number: 4562
Loss: 14.317488197396372
l2 norm of gradients: 0.15470539087162252
l2 norm of weights: 2.3453709507694183
-----
Iteration Number: 4563
Loss: 14.314975316267162
l2 norm of gradients: 0.15464811803894302
l2 norm of weights: 2.345197789951639
-----
Iteration Number: 4564
Loss: 14.312464306054391
l2 norm of gradients: 0.15459087852605619
l2 norm of weights: 2.345024785626554
-----
Iteration Number: 4565
Loss: 14.309955165232914
l2 norm of gradients: 0.15453367231137552
l2 norm of weights: 2.34485193769324
-----
Iteration Number: 4566
Loss: 14.307447892278876
l2 norm of gradients: 0.1544764993733314
l2 norm of weights: 2.344679246050819
-----
Iteration Number: 4567
Loss: 14.304942485669711

```

```

12 norm of gradients: 0.15441935969037096
12 norm of weights: 2.3445067105984587
-----
Iteration Number: 4568
Loss: 14.302438943884143
12 norm of gradients: 0.15436225324095826
12 norm of weights: 2.344334331235372
-----
Iteration Number: 4569
Loss: 14.299937265402175
12 norm of gradients: 0.15430518000357407
12 norm of weights: 2.344162107860817
-----
Iteration Number: 4570
Loss: 14.297437448705102
12 norm of gradients: 0.15424813995671618
12 norm of weights: 2.343990040374098
-----
Iteration Number: 4571
Loss: 14.294939492275498
12 norm of gradients: 0.15419113307889923
12 norm of weights: 2.343818128674564
-----
Iteration Number: 4572
Loss: 14.292443394597221
12 norm of gradients: 0.1541341593486548
12 norm of weights: 2.3436463726616092
-----
Iteration Number: 4573
Loss: 14.289949154155405
12 norm of gradients: 0.1540772187445312
12 norm of weights: 2.343474772234674
-----
Iteration Number: 4574
Loss: 14.287456769436474
12 norm of gradients: 0.15402031124509402
12 norm of weights: 2.3433033272932438
-----
Iteration Number: 4575
Loss: 14.284966238928131
12 norm of gradients: 0.15396343682892563
12 norm of weights: 2.34313203773685
-----
Iteration Number: 4576
Loss: 14.282477561119347
12 norm of gradients: 0.15390659547462535
12 norm of weights: 2.3429609034650687
-----

```

Iteration Number: 4577
Loss: 14.279990734500366
l2 norm of gradients: 0.15384978716080958
l2 norm of weights: 2.3427899243775214

Iteration Number: 4578
Loss: 14.277505757562729
l2 norm of gradients: 0.15379301186611175
l2 norm of weights: 2.342619100373876

Iteration Number: 4579
Loss: 14.275022628799235
l2 norm of gradients: 0.15373626956918227
l2 norm of weights: 2.342448431353845

Iteration Number: 4580
Loss: 14.27254134670396
l2 norm of gradients: 0.1536795602486886
l2 norm of weights: 2.342277917217188

Iteration Number: 4581
Loss: 14.270061909772254
l2 norm of gradients: 0.15362288388331535
l2 norm of weights: 2.342107557863708

Iteration Number: 4582
Loss: 14.26758431650073
l2 norm of gradients: 0.15356624045176412
l2 norm of weights: 2.341937353193254

Iteration Number: 4583
Loss: 14.265108565387285
l2 norm of gradients: 0.15350962993275372
l2 norm of weights: 2.3417673031057222

Iteration Number: 4584
Loss: 14.262634654931073
l2 norm of gradients: 0.15345305230502
l2 norm of weights: 2.3415974075010526

Iteration Number: 4585
Loss: 14.260162583632521
l2 norm of gradients: 0.15339650754731599
l2 norm of weights: 2.3414276662792313

Iteration Number: 4586
Loss: 14.257692349993325
l2 norm of gradients: 0.1533399956384118

```

12 norm of weights: 2.3412580793402897
-----
Iteration Number: 4587
Loss: 14.255223952516442
12 norm of gradients: 0.15328351655709477
12 norm of weights: 2.3410886465843057
-----
Iteration Number: 4588
Loss: 14.25275738970609
12 norm of gradients: 0.15322707028216948
12 norm of weights: 2.3409193679114018
-----
Iteration Number: 4589
Loss: 14.250292660067759
12 norm of gradients: 0.15317065679245767
12 norm of weights: 2.340750243221747
-----
Iteration Number: 4590
Loss: 14.247829762108195
12 norm of gradients: 0.15311427606679817
12 norm of weights: 2.3405812724155544
-----
Iteration Number: 4591
Loss: 14.245368694335406
12 norm of gradients: 0.15305792808404722
12 norm of weights: 2.340412455393085
-----
Iteration Number: 4592
Loss: 14.24290945525866
12 norm of gradients: 0.1530016128230782
12 norm of weights: 2.340243792054643
-----
Iteration Number: 4593
Loss: 14.240452043388485
12 norm of gradients: 0.15294533026278181
12 norm of weights: 2.3400752823005795
-----
Iteration Number: 4594
Loss: 14.237996457236665
12 norm of gradients: 0.15288908038206594
12 norm of weights: 2.3399069260312912
-----
Iteration Number: 4595
Loss: 14.235542695316246
12 norm of gradients: 0.15283286315985595
12 norm of weights: 2.3397387231472213
-----
Iteration Number: 4596

```

Loss: 14.233090756141518
l2 norm of gradients: 0.15277667857509422
l2 norm of weights: 2.3395706735488564

Iteration Number: 4597
Loss: 14.230640638228035
l2 norm of gradients: 0.15272052660674074
l2 norm of weights: 2.3394027771367303

Iteration Number: 4598
Loss: 14.2281923400926
l2 norm of gradients: 0.15266440723377261
l2 norm of weights: 2.339235033811423

Iteration Number: 4599
Loss: 14.225745860253262
l2 norm of gradients: 0.1526083204351844
l2 norm of weights: 2.3390674434735583

Iteration Number: 4600
Loss: 14.223301197229336
l2 norm of gradients: 0.15255226618998807
l2 norm of weights: 2.338900006023808

Iteration Number: 4601
Loss: 14.220858349541373
l2 norm of gradients: 0.15249624447721277
l2 norm of weights: 2.3387327213628875

Iteration Number: 4602
Loss: 14.218417315711182
l2 norm of gradients: 0.15244025527590527
l2 norm of weights: 2.3385655893915596

Iteration Number: 4603
Loss: 14.215978094261805
l2 norm of gradients: 0.15238429856512947
l2 norm of weights: 2.338398610010632

Iteration Number: 4604
Loss: 14.21354068371755
l2 norm of gradients: 0.15232837432396698
l2 norm of weights: 2.3382317831209574

Iteration Number: 4605
Loss: 14.211105082603956
l2 norm of gradients: 0.15227248253151665
l2 norm of weights: 2.3380651086234354

```

-----
Iteration Number: 4606
Loss: 14.208671289447814
l2 norm of gradients: 0.15221662316689485
l2 norm of weights: 2.3378985864190116
-----
Iteration Number: 4607
Loss: 14.20623930277715
l2 norm of gradients: 0.15216079620923526
l2 norm of weights: 2.3377322164086767
-----
Iteration Number: 4608
Loss: 14.20380912112123
l2 norm of gradients: 0.15210500163768917
l2 norm of weights: 2.3375659984934662
-----
Iteration Number: 4609
Loss: 14.201380743010581
l2 norm of gradients: 0.15204923943142523
l2 norm of weights: 2.3373999325744634
-----
Iteration Number: 4610
Loss: 14.198954166976941
l2 norm of gradients: 0.1519935095696297
l2 norm of weights: 2.3372340185527958
-----
Iteration Number: 4611
Loss: 14.196529391553312
l2 norm of gradients: 0.15193781203150628
l2 norm of weights: 2.3370682563296374
-----
Iteration Number: 4612
Loss: 14.194106415273911
l2 norm of gradients: 0.1518821467962761
l2 norm of weights: 2.336902645806208
-----
Iteration Number: 4613
Loss: 14.191685236674212
l2 norm of gradients: 0.15182651384317797
l2 norm of weights: 2.3367371868837727
-----
Iteration Number: 4614
Loss: 14.189265854290909
l2 norm of gradients: 0.15177091315146807
l2 norm of weights: 2.3365718794636434
-----
Iteration Number: 4615
Loss: 14.186848266661933

```



```
l2 norm of gradients: 0.15171534470042022
l2 norm of weights: 2.3364067234471766
-----
Iteration Number: 4616
Loss: 14.184432472326455
l2 norm of gradients: 0.15165980846932578
l2 norm of weights: 2.336241718735775
-----
Iteration Number: 4617
Loss: 14.182018469824865
l2 norm of gradients: 0.15160430443749362
l2 norm of weights: 2.336076865230888
-----
Iteration Number: 4618
Loss: 14.179606257698797
l2 norm of gradients: 0.15154883258425025
l2 norm of weights: 2.33591216283401
-----
Iteration Number: 4619
Loss: 14.17719583449111
l2 norm of gradients: 0.15149339288893968
l2 norm of weights: 2.335747611446681
-----
Iteration Number: 4620
Loss: 14.174787198745882
l2 norm of gradients: 0.15143798533092365
l2 norm of weights: 2.3355832109704875
-----
Iteration Number: 4621
Loss: 14.172380349008433
l2 norm of gradients: 0.15138260988958135
l2 norm of weights: 2.3354189613070617
-----
Iteration Number: 4622
Loss: 14.169975283825298
l2 norm of gradients: 0.1513272665443097
l2 norm of weights: 2.335254862358082
-----
Iteration Number: 4623
Loss: 14.167572001744244
l2 norm of gradients: 0.1512719552745231
l2 norm of weights: 2.335090914025272
-----
Iteration Number: 4624
Loss: 14.165170501314257
l2 norm of gradients: 0.15121667605965378
l2 norm of weights: 2.3349271162104013
-----
```

Iteration Number: 4625
Loss: 14.162770781085552
l2 norm of gradients: 0.15116142887915146
l2 norm of weights: 2.334763468815286

Iteration Number: 4626
Loss: 14.160372839609554
l2 norm of gradients: 0.15110621371248353
l2 norm of weights: 2.3345999717417873

Iteration Number: 4627
Loss: 14.157976675438917
l2 norm of gradients: 0.15105103053913513
l2 norm of weights: 2.334436624891813

Iteration Number: 4628
Loss: 14.15558228712752
l2 norm of gradients: 0.15099587933860895
l2 norm of weights: 2.3342734281673168

Iteration Number: 4629
Loss: 14.153189673230447
l2 norm of gradients: 0.1509407600904254
l2 norm of weights: 2.334110381470297

Iteration Number: 4630
Loss: 14.150798832304009
l2 norm of gradients: 0.15088567277412257
l2 norm of weights: 2.3339474847028

Iteration Number: 4631
Loss: 14.148409762905727
l2 norm of gradients: 0.15083061736925624
l2 norm of weights: 2.3337847377669165

Iteration Number: 4632
Loss: 14.146022463594342
l2 norm of gradients: 0.1507755938554
l2 norm of weights: 2.3336221405647843

Iteration Number: 4633
Loss: 14.143636932929807
l2 norm of gradients: 0.15072060221214492
l2 norm of weights: 2.3334596929985856

Iteration Number: 4634
Loss: 14.141253169473286
l2 norm of gradients: 0.15066564241909994

```
l2 norm of weights: 2.3332973949705496
-----
Iteration Number: 4635
Loss: 14.138871171787157
l2 norm of gradients: 0.15061071445589178
l2 norm of weights: 2.3331352463829518
-----
Iteration Number: 4636
Loss: 14.136490938435008
l2 norm of gradients: 0.15055581830216466
l2 norm of weights: 2.332973247138113
-----
Iteration Number: 4637
Loss: 14.134112467981632
l2 norm of gradients: 0.15050095393758076
l2 norm of weights: 2.332811397138401
-----
Iteration Number: 4638
Loss: 14.131735758993042
l2 norm of gradients: 0.15044612134181987
l2 norm of weights: 2.3326496962862278
-----
Iteration Number: 4639
Loss: 14.12936081003644
l2 norm of gradients: 0.15039132049457962
l2 norm of weights: 2.332488144484052
-----
Iteration Number: 4640
Loss: 14.12698761968025
l2 norm of gradients: 0.15033655137557533
l2 norm of weights: 2.33232674163438
-----
Iteration Number: 4641
Loss: 14.1246161864941
l2 norm of gradients: 0.15028181396454013
l2 norm of weights: 2.3321654876397617
-----
Iteration Number: 4642
Loss: 14.122246509048807
l2 norm of gradients: 0.15022710824122482
l2 norm of weights: 2.3320043824027947
-----
Iteration Number: 4643
Loss: 14.119878585916412
l2 norm of gradients: 0.15017243418539808
l2 norm of weights: 2.331843425826122
-----
Iteration Number: 4644
```

Loss: 14.117512415670143
l2 norm of gradients: 0.1501177917768464
l2 norm of weights: 2.3316826178124317

Iteration Number: 4645
Loss: 14.115147996884426
l2 norm of gradients: 0.15006318099537386
l2 norm of weights: 2.3315219582644606

Iteration Number: 4646
Loss: 14.1127853281349
l2 norm of gradients: 0.15000860182080256
l2 norm of weights: 2.3313614470849884

Iteration Number: 4647
Loss: 14.110424407998394
l2 norm of gradients: 0.14995405423297223
l2 norm of weights: 2.331201084176843

Iteration Number: 4648
Loss: 14.10806523505293
l2 norm of gradients: 0.14989953821174046
l2 norm of weights: 2.3310408694428975

Iteration Number: 4649
Loss: 14.105707807877739
l2 norm of gradients: 0.14984505373698268
l2 norm of weights: 2.3308808027860715

Iteration Number: 4650
Loss: 14.103352125053236
l2 norm of gradients: 0.1497906007885921
l2 norm of weights: 2.330720884109329

Iteration Number: 4651
Loss: 14.100998185161036
l2 norm of gradients: 0.14973617934647973
l2 norm of weights: 2.3305611133156834

Iteration Number: 4652
Loss: 14.09864598678394
l2 norm of gradients: 0.1496817893905744
l2 norm of weights: 2.330401490308191

Iteration Number: 4653
Loss: 14.096295528505948
l2 norm of gradients: 0.14962743090082278
l2 norm of weights: 2.330242014989956

```

-----
Iteration Number: 4654
Loss: 14.093946808912245
l2 norm of gradients: 0.14957310385718933
l2 norm of weights: 2.3300826872641274
-----
Iteration Number: 4655
Loss: 14.091599826589215
l2 norm of gradients: 0.14951880823965644
l2 norm of weights: 2.3299235070339015
-----
Iteration Number: 4656
Loss: 14.089254580124415
l2 norm of gradients: 0.14946454402822415
l2 norm of weights: 2.32976447420252
-----
Iteration Number: 4657
Loss: 14.086911068106607
l2 norm of gradients: 0.14941031120291054
l2 norm of weights: 2.329605588673271
-----
Iteration Number: 4658
Loss: 14.084569289125724
l2 norm of gradients: 0.1493561097437514
l2 norm of weights: 2.329446850349488
-----
Iteration Number: 4659
Loss: 14.082229241772891
l2 norm of gradients: 0.14930193963080038
l2 norm of weights: 2.3292882591345525
-----
Iteration Number: 4660
Loss: 14.07989092464042
l2 norm of gradients: 0.14924780084412903
l2 norm of weights: 2.32912981493189
-----
Iteration Number: 4661
Loss: 14.077554336321803
l2 norm of gradients: 0.1491936933638267
l2 norm of weights: 2.328971517644973
-----
Iteration Number: 4662
Loss: 14.07521947541171
l2 norm of gradients: 0.14913961717000054
l2 norm of weights: 2.3288133671773203
-----
Iteration Number: 4663
Loss: 14.072886340505994

```

```

12 norm of gradients: 0.14908557224277574
12 norm of weights: 2.3286553634324965
-----
Iteration Number: 4664
Loss: 14.070554930201695
12 norm of gradients: 0.14903155856229514
12 norm of weights: 2.328497506314113
-----
Iteration Number: 4665
Loss: 14.068225243097022
12 norm of gradients: 0.14897757610871953
12 norm of weights: 2.3283397957258263
-----
Iteration Number: 4666
Loss: 14.065897277791365
12 norm of gradients: 0.14892362486222752
12 norm of weights: 2.3281822315713403
-----
Iteration Number: 4667
Loss: 14.06357103288529
12 norm of gradients: 0.14886970480301567
12 norm of weights: 2.3280248137544035
-----
Iteration Number: 4668
Loss: 14.061246506980547
12 norm of gradients: 0.14881581591129828
12 norm of weights: 2.3278675421788124
-----
Iteration Number: 4669
Loss: 14.05892369868004
12 norm of gradients: 0.1487619581673076
12 norm of weights: 2.327710416748408
-----
Iteration Number: 4670
Loss: 14.056602606587868
12 norm of gradients: 0.14870813155129375
12 norm of weights: 2.3275534373670785
-----
Iteration Number: 4671
Loss: 14.054283229309283
12 norm of gradients: 0.1486543360435246
12 norm of weights: 2.327396603938758
-----
Iteration Number: 4672
Loss: 14.051965565450729
12 norm of gradients: 0.148600571624286
12 norm of weights: 2.327239916367427
-----

```

Iteration Number: 4673
Loss: 14.0496496136198
l2 norm of gradients: 0.14854683827388168
l2 norm of weights: 2.3270833745571116

Iteration Number: 4674
Loss: 14.047335372425268
l2 norm of gradients: 0.14849313597263314
l2 norm of weights: 2.3269269784118847

Iteration Number: 4675
Loss: 14.045022840477076
l2 norm of gradients: 0.14843946470087985
l2 norm of weights: 2.3267707278358656

Iteration Number: 4676
Loss: 14.042712016386329
l2 norm of gradients: 0.148385824438979
l2 norm of weights: 2.326614622733219

Iteration Number: 4677
Loss: 14.040402898765299
l2 norm of gradients: 0.14833221516730571
l2 norm of weights: 2.326458663008156

Iteration Number: 4678
Loss: 14.038095486227418
l2 norm of gradients: 0.14827863686625317
l2 norm of weights: 2.3263028485649353

Iteration Number: 4679
Loss: 14.035789777387285
l2 norm of gradients: 0.14822508951623206
l2 norm of weights: 2.3261471793078594

Iteration Number: 4680
Loss: 14.033485770860665
l2 norm of gradients: 0.1481715730976713
l2 norm of weights: 2.3259916551412787

Iteration Number: 4681
Loss: 14.031183465264478
l2 norm of gradients: 0.14811808759101738
l2 norm of weights: 2.32583627596959

Iteration Number: 4682
Loss: 14.02888285921681
l2 norm of gradients: 0.1480646329767348

```
l2 norm of weights: 2.3256810416972353
-----
Iteration Number: 4683
Loss: 14.026583951336905
l2 norm of gradients: 0.1480112092353059
l2 norm of weights: 2.3255259522287033
-----
Iteration Number: 4684
Loss: 14.024286740245156
l2 norm of gradients: 0.14795781634723082
l2 norm of weights: 2.3253710074685294
-----
Iteration Number: 4685
Loss: 14.021991224563125
l2 norm of gradients: 0.14790445429302765
l2 norm of weights: 2.3252162073212945
-----
Iteration Number: 4686
Loss: 14.019697402913526
l2 norm of gradients: 0.1478511230532323
l2 norm of weights: 2.325061551691627
-----
Iteration Number: 4687
Loss: 14.017405273920225
l2 norm of gradients: 0.14779782260839852
l2 norm of weights: 2.324907040484199
-----
Iteration Number: 4688
Loss: 14.015114836208244
l2 norm of gradients: 0.14774455293909788
l2 norm of weights: 2.3247526736037325
-----
Iteration Number: 4689
Loss: 14.012826088403758
l2 norm of gradients: 0.14769131402591992
l2 norm of weights: 2.324598450954993
-----
Iteration Number: 4690
Loss: 14.010539029134089
l2 norm of gradients: 0.14763810584947185
l2 norm of weights: 2.3244443724427923
-----
Iteration Number: 4691
Loss: 14.008253657027723
l2 norm of gradients: 0.14758492839037887
l2 norm of weights: 2.3242904379719906
-----
Iteration Number: 4692
```



```

Loss: 14.005969970714274
12 norm of gradients: 0.14753178162928401
12 norm of weights: 2.3241366474474927
-----
Iteration Number: 4693
Loss: 14.003687968824524
12 norm of gradients: 0.14747866554684802
12 norm of weights: 2.3239830007742497
-----
Iteration Number: 4694
Loss: 14.001407649990393
12 norm of gradients: 0.14742558012374965
12 norm of weights: 2.32382949785726
-----
Iteration Number: 4695
Loss: 13.999129012844946
12 norm of gradients: 0.14737252534068535
12 norm of weights: 2.323676138601567
-----
Iteration Number: 4696
Loss: 13.996852056022403
12 norm of gradients: 0.1473195011783695
12 norm of weights: 2.323522922912262
-----
Iteration Number: 4697
Loss: 13.994576778158118
12 norm of gradients: 0.14726650761753424
12 norm of weights: 2.323369850694481
-----
Iteration Number: 4698
Loss: 13.992303177888587
12 norm of gradients: 0.14721354463892944
12 norm of weights: 2.3232169218534073
-----
Iteration Number: 4699
Loss: 13.990031253851459
12 norm of gradients: 0.14716061222332302
12 norm of weights: 2.3230641362942706
-----
Iteration Number: 4700
Loss: 13.987761004685511
12 norm of gradients: 0.14710771035150053
12 norm of weights: 2.3229114939223465
-----
Iteration Number: 4701
Loss: 13.985492429030673
12 norm of gradients: 0.1470548390042654
12 norm of weights: 2.322758994642956

```

```

-----
Iteration Number: 4702
Loss: 13.983225525528004
l2 norm of gradients: 0.14700199816243883
l2 norm of weights: 2.3226066383614685
-----
Iteration Number: 4703
Loss: 13.980960292819702
l2 norm of gradients: 0.14694918780685984
l2 norm of weights: 2.322454424983299
-----
Iteration Number: 4704
Loss: 13.978696729549107
l2 norm of gradients: 0.14689640791838526
l2 norm of weights: 2.322302354413907
-----
Iteration Number: 4705
Loss: 13.976434834360685
l2 norm of gradients: 0.14684365847788966
l2 norm of weights: 2.3221504265588013
-----
Iteration Number: 4706
Loss: 13.97417460590005
l2 norm of gradients: 0.14679093946626545
l2 norm of weights: 2.321998641323535
-----
Iteration Number: 4707
Loss: 13.971916042813941
l2 norm of gradients: 0.1467382508644228
l2 norm of weights: 2.3218469986137085
-----
Iteration Number: 4708
Loss: 13.969659143750228
l2 norm of gradients: 0.1466855926532896
l2 norm of weights: 2.321695498334968
-----
Iteration Number: 4709
Loss: 13.967403907357914
l2 norm of gradients: 0.14663296481381155
l2 norm of weights: 2.3215441403930064
-----
Iteration Number: 4710
Loss: 13.96515033228714
l2 norm of gradients: 0.14658036732695218
l2 norm of weights: 2.3213929246935625
-----
Iteration Number: 4711
Loss: 13.962898417189159

```

```

12 norm of gradients: 0.14652780017369266
12 norm of weights: 2.321241851142422
-----
Iteration Number: 4712
Loss: 13.96064816071638
12 norm of gradients: 0.14647526333503197
12 norm of weights: 2.3210909196454175
-----
Iteration Number: 4713
Loss: 13.958399561522306
12 norm of gradients: 0.14642275679198677
12 norm of weights: 2.3209401301084265
-----
Iteration Number: 4714
Loss: 13.95615261826159
12 norm of gradients: 0.14637028052559162
12 norm of weights: 2.3207894824373736
-----
Iteration Number: 4715
Loss: 13.953907329590008
12 norm of gradients: 0.14631783451689853
12 norm of weights: 2.32063897653823
-----
Iteration Number: 4716
Loss: 13.951663694164443
12 norm of gradients: 0.14626541874697754
12 norm of weights: 2.3204886123170136
-----
Iteration Number: 4717
Loss: 13.949421710642921
12 norm of gradients: 0.1462130331969162
12 norm of weights: 2.320338389679788
-----
Iteration Number: 4718
Loss: 13.947181377684585
12 norm of gradients: 0.1461606778478198
12 norm of weights: 2.320188308532663
-----
Iteration Number: 4719
Loss: 13.944942693949681
12 norm of gradients: 0.1461083526808114
12 norm of weights: 2.3200383687817956
-----
Iteration Number: 4720
Loss: 13.942705658099605
12 norm of gradients: 0.14605605767703167
12 norm of weights: 2.3198885703333882
-----

```

Iteration Number: 4721
Loss: 13.940470268796851
l2 norm of gradients: 0.1460037928176391
l2 norm of weights: 2.3197389130936914

Iteration Number: 4722
Loss: 13.938236524705038
l2 norm of gradients: 0.1459515580838096
l2 norm of weights: 2.319589396969

Iteration Number: 4723
Loss: 13.936004424488894
l2 norm of gradients: 0.14589935345673705
l2 norm of weights: 2.3194400218656566

Iteration Number: 4724
Loss: 13.93377396681428
l2 norm of gradients: 0.14584717891763282
l2 norm of weights: 2.3192907876900497

Iteration Number: 4725
Loss: 13.931545150348152
l2 norm of gradients: 0.14579503444772587
l2 norm of weights: 2.3191416943486147

Iteration Number: 4726
Loss: 13.92931797375859
l2 norm of gradients: 0.14574292002826303
l2 norm of weights: 2.3189927417478327

Iteration Number: 4727
Loss: 13.927092435714787
l2 norm of gradients: 0.1456908356405085
l2 norm of weights: 2.318843929794232

Iteration Number: 4728
Loss: 13.924868534887043
l2 norm of gradients: 0.1456387812657443
l2 norm of weights: 2.3186952583943867

Iteration Number: 4729
Loss: 13.922646269946775
l2 norm of gradients: 0.14558675688527004
l2 norm of weights: 2.3185467274549176

Iteration Number: 4730
Loss: 13.9204256395665
l2 norm of gradients: 0.1455347624804028

```

12 norm of weights: 2.318398336882492
-----
Iteration Number: 4731
Loss: 13.918206642419856
12 norm of gradients: 0.14548279803247743
12 norm of weights: 2.318250086583823
-----
Iteration Number: 4732
Loss: 13.915989277181575
12 norm of gradients: 0.1454308635228463
12 norm of weights: 2.318101976465672
-----
Iteration Number: 4733
Loss: 13.913773542527505
12 norm of gradients: 0.14537895893287928
12 norm of weights: 2.317954006434844
-----
Iteration Number: 4734
Loss: 13.911559437134603
12 norm of gradients: 0.14532708424396396
12 norm of weights: 2.317806176398193
-----
Iteration Number: 4735
Loss: 13.909346959680914
12 norm of gradients: 0.14527523943750537
12 norm of weights: 2.317658486262618
-----
Iteration Number: 4736
Loss: 13.907136108845599
12 norm of gradients: 0.14522342449492612
12 norm of weights: 2.317510935935065
-----
Iteration Number: 4737
Loss: 13.90492688330892
12 norm of gradients: 0.1451716393976663
12 norm of weights: 2.317363525322526
-----
Iteration Number: 4738
Loss: 13.902719281752239
12 norm of gradients: 0.14511988412718377
12 norm of weights: 2.3172162543320405
-----
Iteration Number: 4739
Loss: 13.90051330285802
12 norm of gradients: 0.14506815866495357
12 norm of weights: 2.3170691228706937
-----
Iteration Number: 4740

```

Loss: 13.898308945309816
l2 norm of gradients: 0.14501646299246843
l2 norm of weights: 2.3169221308456165

Iteration Number: 4741
Loss: 13.8961062077923
l2 norm of gradients: 0.14496479709123858
l2 norm of weights: 2.3167752781639877

Iteration Number: 4742
Loss: 13.893905088991215
l2 norm of gradients: 0.1449131609427917
l2 norm of weights: 2.3166285647330316

Iteration Number: 4743
Loss: 13.891705587593425
l2 norm of gradients: 0.1448615545286729
l2 norm of weights: 2.31648199046002

Iteration Number: 4744
Loss: 13.889507702286874
l2 norm of gradients: 0.14480997783044486
l2 norm of weights: 2.3163355552522695

Iteration Number: 4745
Loss: 13.887311431760605
l2 norm of gradients: 0.14475843082968753
l2 norm of weights: 2.316189259017145

Iteration Number: 4746
Loss: 13.88511677470475
l2 norm of gradients: 0.1447069135079985
l2 norm of weights: 2.316043101662057

Iteration Number: 4747
Loss: 13.88292372981054
l2 norm of gradients: 0.1446554258469927
l2 norm of weights: 2.3158970830944625

Iteration Number: 4748
Loss: 13.880732295770297
l2 norm of gradients: 0.1446039678283024
l2 norm of weights: 2.3157512032218643

Iteration Number: 4749
Loss: 13.878542471277424
l2 norm of gradients: 0.14455253943357732
l2 norm of weights: 2.3156054619518134

```

-----
Iteration Number: 4750
Loss: 13.87635425502642
l2 norm of gradients: 0.14450114064448472
l2 norm of weights: 2.315459859191906
-----
Iteration Number: 4751
Loss: 13.874167645712877
l2 norm of gradients: 0.14444977144270893
l2 norm of weights: 2.3153143948497847
-----
Iteration Number: 4752
Loss: 13.871982642033462
l2 norm of gradients: 0.14439843180995193
l2 norm of weights: 2.315169068833139
-----
Iteration Number: 4753
Loss: 13.869799242685932
l2 norm of gradients: 0.14434712172793282
l2 norm of weights: 2.315023881049706
-----
Iteration Number: 4754
Loss: 13.867617446369142
l2 norm of gradients: 0.14429584117838826
l2 norm of weights: 2.314878831407267
-----
Iteration Number: 4755
Loss: 13.865437251783009
l2 norm of gradients: 0.14424459014307203
l2 norm of weights: 2.314733919813651
-----
Iteration Number: 4756
Loss: 13.863258657628549
l2 norm of gradients: 0.14419336860375534
l2 norm of weights: 2.314589146176734
-----
Iteration Number: 4757
Loss: 13.861081662607855
l2 norm of gradients: 0.14414217654222666
l2 norm of weights: 2.314444510404438
-----
Iteration Number: 4758
Loss: 13.858906265424107
l2 norm of gradients: 0.1440910139402918
l2 norm of weights: 2.3143000124047313
-----
Iteration Number: 4759
Loss: 13.856732464781551

```

```

12 norm of gradients: 0.1440398807797737
12 norm of weights: 2.3141556520856295
-----
Iteration Number: 4760
Loss: 13.854560259385524
12 norm of gradients: 0.14398877704251264
12 norm of weights: 2.314011429355193
-----
Iteration Number: 4761
Loss: 13.852389647942438
12 norm of gradients: 0.14393770271036624
12 norm of weights: 2.3138673441215305
-----
Iteration Number: 4762
Loss: 13.850220629159784
12 norm of gradients: 0.14388665776520912
12 norm of weights: 2.313723396292797
-----
Iteration Number: 4763
Loss: 13.848053201746119
12 norm of gradients: 0.14383564218893335
12 norm of weights: 2.313579585777193
-----
Iteration Number: 4764
Loss: 13.845887364411093
12 norm of gradients: 0.143784655963448
12 norm of weights: 2.3134359124829658
-----
Iteration Number: 4765
Loss: 13.84372311586541
12 norm of gradients: 0.14373369907067945
12 norm of weights: 2.3132923763184103
-----
Iteration Number: 4766
Loss: 13.841560454820861
12 norm of gradients: 0.14368277149257122
12 norm of weights: 2.3131489771918665
-----
Iteration Number: 4767
Loss: 13.839399379990306
12 norm of gradients: 0.14363187321108392
12 norm of weights: 2.3130057150117223
-----
Iteration Number: 4768
Loss: 13.837239890087675
12 norm of gradients: 0.1435810042081954
12 norm of weights: 2.312862589686411
-----

```


Iteration Number: 4769
Loss: 13.835081983827967
l2 norm of gradients: 0.14353016446590053
l2 norm of weights: 2.312719601124413

Iteration Number: 4770
Loss: 13.83292565992725
l2 norm of gradients: 0.14347935396621125
l2 norm of weights: 2.3125767492342546

Iteration Number: 4771
Loss: 13.830770917102662
l2 norm of gradients: 0.1434285726911569
l2 norm of weights: 2.3124340339245095

Iteration Number: 4772
Loss: 13.828617754072408
l2 norm of gradients: 0.14337782062278348
l2 norm of weights: 2.3122914551037974

Iteration Number: 4773
Loss: 13.826466169555758
l2 norm of gradients: 0.1433270977431543
l2 norm of weights: 2.3121490126807855

Iteration Number: 4774
Loss: 13.824316162273048
l2 norm of gradients: 0.14327640403434963
l2 norm of weights: 2.3120067065641856

Iteration Number: 4775
Loss: 13.822167730945676
l2 norm of gradients: 0.1432257394784668
l2 norm of weights: 2.3118645366627573

Iteration Number: 4776
Loss: 13.820020874296107
l2 norm of gradients: 0.14317510405762016
l2 norm of weights: 2.3117225028853072

Iteration Number: 4777
Loss: 13.817875591047866
l2 norm of gradients: 0.143124497753941
l2 norm of weights: 2.3115806051406875

Iteration Number: 4778
Loss: 13.81573187992554
l2 norm of gradients: 0.14307392054957763

```
12 norm of weights: 2.3114388433377977
-----
Iteration Number: 4779
Loss: 13.813589739654772
12 norm of gradients: 0.14302337242669527
12 norm of weights: 2.3112972173855826
-----
Iteration Number: 4780
Loss: 13.81144916896227
12 norm of gradients: 0.1429728533674762
12 norm of weights: 2.3111557271930354
-----
Iteration Number: 4781
Loss: 13.8093101665758
12 norm of gradients: 0.14292236335411942
12 norm of weights: 2.3110143726691943
-----
Iteration Number: 4782
Loss: 13.80717273122418
12 norm of gradients: 0.14287190236884115
12 norm of weights: 2.3108731537231444
-----
Iteration Number: 4783
Loss: 13.80503686163729
12 norm of gradients: 0.14282147039387413
12 norm of weights: 2.3107320702640184
-----
Iteration Number: 4784
Loss: 13.80290255654606
12 norm of gradients: 0.14277106741146828
12 norm of weights: 2.3105911222009934
-----
Iteration Number: 4785
Loss: 13.800769814682486
12 norm of gradients: 0.14272069340389024
12 norm of weights: 2.3104503094432958
-----
Iteration Number: 4786
Loss: 13.7986386347796
12 norm of gradients: 0.1426703483534235
12 norm of weights: 2.3103096319001963
-----
Iteration Number: 4787
Loss: 13.796509015571498
12 norm of gradients: 0.1426200322423683
12 norm of weights: 2.310169089481013
-----
Iteration Number: 4788
```

Loss: 13.794380955793327
l2 norm of gradients: 0.1425697450530419
l2 norm of weights: 2.3100286820951106

Iteration Number: 4789
Loss: 13.792254454181277
l2 norm of gradients: 0.14251948676777812
l2 norm of weights: 2.309888409651901

Iteration Number: 4790
Loss: 13.790129509472598
l2 norm of gradients: 0.14246925736892765
l2 norm of weights: 2.309748272060841

Iteration Number: 4791
Loss: 13.788006120405582
l2 norm of gradients: 0.14241905683885794
l2 norm of weights: 2.3096082692314353

Iteration Number: 4792
Loss: 13.785884285719572
l2 norm of gradients: 0.14236888515995313
l2 norm of weights: 2.309468401073235

Iteration Number: 4793
Loss: 13.783764004154952
l2 norm of gradients: 0.14231874231461403
l2 norm of weights: 2.3093286674958375

Iteration Number: 4794
Loss: 13.78164527445316
l2 norm of gradients: 0.14226862828525827
l2 norm of weights: 2.3091890684088865

Iteration Number: 4795
Loss: 13.779528095356666
l2 norm of gradients: 0.14221854305432002
l2 norm of weights: 2.309049603722073

Iteration Number: 4796
Loss: 13.777412465609002
l2 norm of gradients: 0.14216848660425013
l2 norm of weights: 2.3089102733451345

Iteration Number: 4797
Loss: 13.775298383954725
l2 norm of gradients: 0.1421184589175162
l2 norm of weights: 2.308771077187854

```

-----
Iteration Number: 4798
Loss: 13.773185849139448
l2 norm of gradients: 0.1420684599766023
l2 norm of weights: 2.308632015160062
-----
Iteration Number: 4799
Loss: 13.771074859909817
l2 norm of gradients: 0.14201848976400916
l2 norm of weights: 2.3084930871716356
-----
Iteration Number: 4800
Loss: 13.768965415013515
l2 norm of gradients: 0.14196854826225408
l2 norm of weights: 2.3083542931324983
-----
Iteration Number: 4801
Loss: 13.766857513199273
l2 norm of gradients: 0.14191863545387093
l2 norm of weights: 2.30821563295262
-----
Iteration Number: 4802
Loss: 13.764751153216855
l2 norm of gradients: 0.14186875132141008
l2 norm of weights: 2.3080771065420174
-----
Iteration Number: 4803
Loss: 13.76264633381706
l2 norm of gradients: 0.14181889584743845
l2 norm of weights: 2.3079387138107537
-----
Iteration Number: 4804
Loss: 13.76054305375173
l2 norm of gradients: 0.14176906901453945
l2 norm of weights: 2.307800454668939
-----
Iteration Number: 4805
Loss: 13.758441311773732
l2 norm of gradients: 0.14171927080531294
l2 norm of weights: 2.3076623290267286
-----
Iteration Number: 4806
Loss: 13.75634110663698
l2 norm of gradients: 0.14166950120237531
l2 norm of weights: 2.307524336794327
-----
Iteration Number: 4807
Loss: 13.754242437096412

```

```

12 norm of gradients: 0.1416197601883592
12 norm of weights: 2.307386477881982
-----
Iteration Number: 4808
Loss: 13.752145301908
12 norm of gradients: 0.14157004774591386
12 norm of weights: 2.307248752199991
-----
Iteration Number: 4809
Loss: 13.750049699828747
12 norm of gradients: 0.14152036385770492
12 norm of weights: 2.3071111596586964
-----
Iteration Number: 4810
Loss: 13.74795562961669
12 norm of gradients: 0.14147070850641424
12 norm of weights: 2.3069737001684874
-----
Iteration Number: 4811
Loss: 13.745863090030893
12 norm of gradients: 0.14142108167474016
12 norm of weights: 2.306836373639799
-----
Iteration Number: 4812
Loss: 13.743772079831452
12 norm of gradients: 0.14137148334539731
12 norm of weights: 2.306699179983115
-----
Iteration Number: 4813
Loss: 13.741682597779482
12 norm of gradients: 0.14132191350111664
12 norm of weights: 2.306562119108964
-----
Iteration Number: 4814
Loss: 13.739594642637137
12 norm of gradients: 0.1412723721246453
12 norm of weights: 2.306425190927921
-----
Iteration Number: 4815
Loss: 13.737508213167585
12 norm of gradients: 0.14122285919874686
12 norm of weights: 2.3062883953506086
-----
Iteration Number: 4816
Loss: 13.735423308135024
12 norm of gradients: 0.14117337470620103
12 norm of weights: 2.3061517322876957
-----

```

Iteration Number: 4817
Loss: 13.733339926304678
l2 norm of gradients: 0.14112391862980367
l2 norm of weights: 2.3060152016498967

Iteration Number: 4818
Loss: 13.731258066442795
l2 norm of gradients: 0.14107449095236707
l2 norm of weights: 2.3058788033479747

Iteration Number: 4819
Loss: 13.729177727316635
l2 norm of gradients: 0.14102509165671948
l2 norm of weights: 2.3057425372927383

Iteration Number: 4820
Loss: 13.72709890769449
l2 norm of gradients: 0.14097572072570538
l2 norm of weights: 2.3056064033950414

Iteration Number: 4821
Loss: 13.725021606345672
l2 norm of gradients: 0.14092637814218548
l2 norm of weights: 2.305470401565787

Iteration Number: 4822
Loss: 13.722945822040503
l2 norm of gradients: 0.14087706388903645
l2 norm of weights: 2.305334531715922

Iteration Number: 4823
Loss: 13.72087155355033
l2 norm of gradients: 0.14082777794915102
l2 norm of weights: 2.305198793756443

Iteration Number: 4824
Loss: 13.718798799647518
l2 norm of gradients: 0.14077852030543825
l2 norm of weights: 2.3050631875983902

Iteration Number: 4825
Loss: 13.71672755910545
l2 norm of gradients: 0.14072929094082293
l2 norm of weights: 2.304927713152852

Iteration Number: 4826
Loss: 13.714657830698513
l2 norm of gradients: 0.14068008983824615

```

12 norm of weights: 2.3047923703309627
-----
Iteration Number: 4827
Loss: 13.712589613202127
12 norm of gradients: 0.14063091698066468
12 norm of weights: 2.304657159043904
-----
Iteration Number: 4828
Loss: 13.71052290539271
12 norm of gradients: 0.1405817723510516
12 norm of weights: 2.3045220792029038
-----
Iteration Number: 4829
Loss: 13.708457706047698
12 norm of gradients: 0.14053265593239572
12 norm of weights: 2.3043871307192356
-----
Iteration Number: 4830
Loss: 13.706394013945543
12 norm of gradients: 0.14048356770770182
12 norm of weights: 2.3042523135042217
-----
Iteration Number: 4831
Loss: 13.704331827865703
12 norm of gradients: 0.1404345076599906
12 norm of weights: 2.3041176274692283
-----
Iteration Number: 4832
Loss: 13.702271146588652
12 norm of gradients: 0.1403854757722987
12 norm of weights: 2.303983072525671
-----
Iteration Number: 4833
Loss: 13.700211968895864
12 norm of gradients: 0.14033647202767852
12 norm of weights: 2.303848648585009
-----
Iteration Number: 4834
Loss: 13.698154293569825
12 norm of gradients: 0.14028749640919833
12 norm of weights: 2.303714355558751
-----
Iteration Number: 4835
Loss: 13.696098119394037
12 norm of gradients: 0.14023854889994225
12 norm of weights: 2.3035801933584503
-----
Iteration Number: 4836

```

```

Loss: 13.694043445152994
l2 norm of gradients: 0.14018962948301017
l2 norm of weights: 2.303446161895707
-----
Iteration Number: 4837
Loss: 13.691990269632205
l2 norm of gradients: 0.14014073814151765
l2 norm of weights: 2.3033122610821692
-----
Iteration Number: 4838
Loss: 13.689938591618185
l2 norm of gradients: 0.14009187485859612
l2 norm of weights: 2.3031784908295303
-----
Iteration Number: 4839
Loss: 13.687888409898441
l2 norm of gradients: 0.14004303961739262
l2 norm of weights: 2.3030448510495303
-----
Iteration Number: 4840
Loss: 13.685839723261498
l2 norm of gradients: 0.13999423240107
l2 norm of weights: 2.3029113416539557
-----
Iteration Number: 4841
Loss: 13.68379253049687
l2 norm of gradients: 0.1399454531928066
l2 norm of weights: 2.3027779625546407
-----
Iteration Number: 4842
Loss: 13.681746830395083
l2 norm of gradients: 0.1398967019757966
l2 norm of weights: 2.302644713663465
-----
Iteration Number: 4843
Loss: 13.679702621747651
l2 norm of gradients: 0.13984797873324956
l2 norm of weights: 2.3025115948923554
-----
Iteration Number: 4844
Loss: 13.677659903347106
l2 norm of gradients: 0.13979928344839082
l2 norm of weights: 2.302378606153285
-----
Iteration Number: 4845
Loss: 13.675618673986955
l2 norm of gradients: 0.13975061610446127
l2 norm of weights: 2.3022457473582736

```



```

-----
Iteration Number: 4846
Loss: 13.673578932461716
l2 norm of gradients: 0.1397019766847173
l2 norm of weights: 2.302113018419387
-----
Iteration Number: 4847
Loss: 13.671540677566904
l2 norm of gradients: 0.13965336517243068
l2 norm of weights: 2.3019804192487396
-----
Iteration Number: 4848
Loss: 13.669503908099026
l2 norm of gradients: 0.13960478155088896
l2 norm of weights: 2.30184794975849
-----
Iteration Number: 4849
Loss: 13.667468622855584
l2 norm of gradients: 0.1395562258033949
l2 norm of weights: 2.3017156098608447
-----
Iteration Number: 4850
Loss: 13.665434820635078
l2 norm of gradients: 0.13950769791326684
l2 norm of weights: 2.301583399468056
-----
Iteration Number: 4851
Loss: 13.66340250023699
l2 norm of gradients: 0.13945919786383845
l2 norm of weights: 2.3014513184924232
-----
Iteration Number: 4852
Loss: 13.661371660461809
l2 norm of gradients: 0.13941072563845883
l2 norm of weights: 2.301319366846293
-----
Iteration Number: 4853
Loss: 13.659342300111003
l2 norm of gradients: 0.13936228122049246
l2 norm of weights: 2.301187544442057
-----
Iteration Number: 4854
Loss: 13.657314417987038
l2 norm of gradients: 0.1393138645933191
l2 norm of weights: 2.3010558511921544
-----
Iteration Number: 4855
Loss: 13.655288012893365

```

```

12 norm of gradients: 0.13926547574033388
12 norm of weights: 2.3009242870090714
-----
Iteration Number: 4856
Loss: 13.653263083634426
12 norm of gradients: 0.13921711464494718
12 norm of weights: 2.3007928518053395
-----
Iteration Number: 4857
Loss: 13.651239629015649
12 norm of gradients: 0.1391687812905846
12 norm of weights: 2.300661545493538
-----
Iteration Number: 4858
Loss: 13.64921764784345
12 norm of gradients: 0.13912047566068708
12 norm of weights: 2.3005303679862923
-----
Iteration Number: 4859
Loss: 13.647197138925229
12 norm of gradients: 0.1390721977387107
12 norm of weights: 2.3003993191962744
-----
Iteration Number: 4860
Loss: 13.645178101069373
12 norm of gradients: 0.1390239475081267
12 norm of weights: 2.3002683990362023
-----
Iteration Number: 4861
Loss: 13.643160533085252
12 norm of gradients: 0.1389757249524215
12 norm of weights: 2.3001376074188413
-----
Iteration Number: 4862
Loss: 13.641144433783223
12 norm of gradients: 0.1389275300550966
12 norm of weights: 2.3000069442570035
-----
Iteration Number: 4863
Loss: 13.63912980197462
12 norm of gradients: 0.13887936279966864
12 norm of weights: 2.299876409463547
-----
Iteration Number: 4864
Loss: 13.63711663647176
12 norm of gradients: 0.13883122316966942
12 norm of weights: 2.2997460029513763
-----

```

Iteration Number: 4865
Loss: 13.63510493608794
l2 norm of gradients: 0.13878311114864567
l2 norm of weights: 2.299615724633443

Iteration Number: 4866
Loss: 13.633094699637441
l2 norm of gradients: 0.13873502672015908
l2 norm of weights: 2.2994855744227456

Iteration Number: 4867
Loss: 13.631085925935519
l2 norm of gradients: 0.13868696986778653
l2 norm of weights: 2.2993555522323277

Iteration Number: 4868
Loss: 13.629078613798406
l2 norm of gradients: 0.13863894057511966
l2 norm of weights: 2.2992256579752812

Iteration Number: 4869
Loss: 13.627072762043317
l2 norm of gradients: 0.13859093882576523
l2 norm of weights: 2.2990958915647437

Iteration Number: 4870
Loss: 13.625068369488446
l2 norm of gradients: 0.13854296460334475
l2 norm of weights: 2.2989662529138988

Iteration Number: 4871
Loss: 13.623065434952943
l2 norm of gradients: 0.13849501789149468
l2 norm of weights: 2.298836741935978

Iteration Number: 4872
Loss: 13.62106395725696
l2 norm of gradients: 0.13844709867386637
l2 norm of weights: 2.2987073585442586

Iteration Number: 4873
Loss: 13.6190639352216
l2 norm of gradients: 0.13839920693412597
l2 norm of weights: 2.2985781026520646

Iteration Number: 4874
Loss: 13.617065367668955
l2 norm of gradients: 0.13835134265595447

```

12 norm of weights: 2.2984489741727665
-----
Iteration Number: 4875
Loss: 13.615068253422077
12 norm of gradients: 0.13830350582304748
12 norm of weights: 2.2983199730197814
-----
Iteration Number: 4876
Loss: 13.613072591304995
12 norm of gradients: 0.13825569641911556
12 norm of weights: 2.2981910991065724
-----
Iteration Number: 4877
Loss: 13.611078380142713
12 norm of gradients: 0.13820791442788385
12 norm of weights: 2.2980623523466503
-----
Iteration Number: 4878
Loss: 13.609085618761192
12 norm of gradients: 0.13816015983309227
12 norm of weights: 2.2979337326535716
-----
Iteration Number: 4879
Loss: 13.607094305987367
12 norm of gradients: 0.1381124326184953
12 norm of weights: 2.29780523994094
-----
Iteration Number: 4880
Loss: 13.605104440649153
12 norm of gradients: 0.13806473276786213
12 norm of weights: 2.2976768741224047
-----
Iteration Number: 4881
Loss: 13.60311602157541
12 norm of gradients: 0.13801706026497648
12 norm of weights: 2.297548635111663
-----
Iteration Number: 4882
Loss: 13.60112904759598
12 norm of gradients: 0.13796941509363678
12 norm of weights: 2.2974205228224576
-----
Iteration Number: 4883
Loss: 13.59914351754166
12 norm of gradients: 0.13792179723765588
12 norm of weights: 2.2972925371685773
-----
Iteration Number: 4884

```

Loss: 13.597159430244226
l2 norm of gradients: 0.1378742066808612
l2 norm of weights: 2.2971646780638593

Iteration Number: 4885
Loss: 13.595176784536402
l2 norm of gradients: 0.1378266434070946
l2 norm of weights: 2.2970369454221853

Iteration Number: 4886
Loss: 13.59319557925188
l2 norm of gradients: 0.13777910740021257
l2 norm of weights: 2.2969093391574846

Iteration Number: 4887
Loss: 13.591215813225322
l2 norm of gradients: 0.13773159864408585
l2 norm of weights: 2.296781859183734

Iteration Number: 4888
Loss: 13.58923748529233
l2 norm of gradients: 0.13768411712259956
l2 norm of weights: 2.2966545054149545

Iteration Number: 4889
Loss: 13.587260594289491
l2 norm of gradients: 0.13763666281965342
l2 norm of weights: 2.2965272777652155

Iteration Number: 4890
Loss: 13.585285139054335
l2 norm of gradients: 0.13758923571916135
l2 norm of weights: 2.296400176148632

Iteration Number: 4891
Loss: 13.583311118425357
l2 norm of gradients: 0.13754183580505153
l2 norm of weights: 2.2962732004793662

Iteration Number: 4892
Loss: 13.581338531242007
l2 norm of gradients: 0.13749446306126648
l2 norm of weights: 2.2961463506716266

Iteration Number: 4893
Loss: 13.579367376344692
l2 norm of gradients: 0.1374471174717631
l2 norm of weights: 2.2960196266396684

```

-----
Iteration Number: 4894
Loss: 13.577397652574774
l2 norm of gradients: 0.1373997990205124
l2 norm of weights: 2.2958930282977925
-----
Iteration Number: 4895
Loss: 13.575429358774574
l2 norm of gradients: 0.13735250769149956
l2 norm of weights: 2.2957665555603466
-----
Iteration Number: 4896
Loss: 13.573462493787366
l2 norm of gradients: 0.13730524346872405
l2 norm of weights: 2.2956402083417267
-----
Iteration Number: 4897
Loss: 13.571497056457375
l2 norm of gradients: 0.1372580063361994
l2 norm of weights: 2.2955139865563723
-----
Iteration Number: 4898
Loss: 13.569533045629777
l2 norm of gradients: 0.1372107962779533
l2 norm of weights: 2.295387890118772
-----
Iteration Number: 4899
Loss: 13.567570460150705
l2 norm of gradients: 0.13716361327802742
l2 norm of weights: 2.2952619189434595
-----
Iteration Number: 4900
Loss: 13.565609298867244
l2 norm of gradients: 0.13711645732047772
l2 norm of weights: 2.2951360729450156
-----
Iteration Number: 4901
Loss: 13.563649560627418
l2 norm of gradients: 0.1370693283893739
l2 norm of weights: 2.295010352038068
-----
Iteration Number: 4902
Loss: 13.561691244280214
l2 norm of gradients: 0.1370222264687999
l2 norm of weights: 2.29488475613729
-----
Iteration Number: 4903
Loss: 13.55973434867556

```

```

12 norm of gradients: 0.13697515154285342
12 norm of weights: 2.2947592851574012
-----
Iteration Number: 4904
Loss: 13.557778872664331
12 norm of gradients: 0.13692810359564625
12 norm of weights: 2.294633939013169
-----
Iteration Number: 4905
Loss: 13.555824815098356
12 norm of gradients: 0.13688108261130402
12 norm of weights: 2.2945087176194066
-----
Iteration Number: 4906
Loss: 13.5538721748304
12 norm of gradients: 0.13683408857396626
12 norm of weights: 2.294383620890974
-----
Iteration Number: 4907
Loss: 13.55192095071418
12 norm of gradients: 0.13678712146778635
12 norm of weights: 2.2942586487427765
-----
Iteration Number: 4908
Loss: 13.549971141604356
12 norm of gradients: 0.13674018127693144
12 norm of weights: 2.294133801089768
-----
Iteration Number: 4909
Loss: 13.548022746356528
12 norm of gradients: 0.13669326798558254
12 norm of weights: 2.294009077846947
-----
Iteration Number: 4910
Loss: 13.546075763827245
12 norm of gradients: 0.13664638157793435
12 norm of weights: 2.2938844789293595
-----
Iteration Number: 4911
Loss: 13.544130192873997
12 norm of gradients: 0.1365995220381953
12 norm of weights: 2.293760004252098
-----
Iteration Number: 4912
Loss: 13.54218603235521
12 norm of gradients: 0.13655268935058767
12 norm of weights: 2.293635653730301
-----

```

Iteration Number: 4913
Loss: 13.540243281130248
l2 norm of gradients: 0.13650588349934714
l2 norm of weights: 2.293511427279154

Iteration Number: 4914
Loss: 13.538301938059426
l2 norm of gradients: 0.13645910446872322
l2 norm of weights: 2.2933873248138887

Iteration Number: 4915
Loss: 13.53636200200399
l2 norm of gradients: 0.13641235224297896
l2 norm of weights: 2.293263346249783

Iteration Number: 4916
Loss: 13.534423471826129
l2 norm of gradients: 0.13636562680639097
l2 norm of weights: 2.293139491502162

Iteration Number: 4917
Loss: 13.532486346388959
l2 norm of gradients: 0.13631892814324953
l2 norm of weights: 2.293015760486397

Iteration Number: 4918
Loss: 13.53055062455654
l2 norm of gradients: 0.13627225623785824
l2 norm of weights: 2.2928921531179056

Iteration Number: 4919
Loss: 13.52861630519387
l2 norm of gradients: 0.13622561107453432
l2 norm of weights: 2.2927686693121516

Iteration Number: 4920
Loss: 13.526683387166875
l2 norm of gradients: 0.13617899263760838
l2 norm of weights: 2.292645308984646

Iteration Number: 4921
Loss: 13.524751869342415
l2 norm of gradients: 0.13613240091142453
l2 norm of weights: 2.292522072050946

Iteration Number: 4922
Loss: 13.52282175058829
l2 norm of gradients: 0.13608583588034015


```
l2 norm of weights: 2.2923989584266553
-----
Iteration Number: 4923
Loss: 13.520893029773227
l2 norm of gradients: 0.13603929752872604
l2 norm of weights: 2.2922759680274236
-----
Iteration Number: 4924
Loss: 13.518965705766888
l2 norm of gradients: 0.1359927858409664
l2 norm of weights: 2.2921531007689477
-----
Iteration Number: 4925
Loss: 13.517039777439859
l2 norm of gradients: 0.13594630080145864
l2 norm of weights: 2.2920303565669706
-----
Iteration Number: 4926
Loss: 13.515115243663658
l2 norm of gradients: 0.1358998423946135
l2 norm of weights: 2.2919077353372814
-----
Iteration Number: 4927
Loss: 13.51319210331074
l2 norm of gradients: 0.13585341060485484
l2 norm of weights: 2.291785236995717
-----
Iteration Number: 4928
Loss: 13.511270355254478
l2 norm of gradients: 0.13580700541661989
l2 norm of weights: 2.2916628614581582
-----
Iteration Number: 4929
Loss: 13.509349998369181
l2 norm of gradients: 0.13576062681435888
l2 norm of weights: 2.2915406086405357
-----
Iteration Number: 4930
Loss: 13.507431031530084
l2 norm of gradients: 0.13571427478253542
l2 norm of weights: 2.2914184784588234
-----
Iteration Number: 4931
Loss: 13.505513453613336
l2 norm of gradients: 0.13566794930562592
l2 norm of weights: 2.2912964708290438
-----
Iteration Number: 4932
```

Loss: 13.503597263496024
l2 norm of gradients: 0.13562165036812018
l2 norm of weights: 2.2911745856672647

Iteration Number: 4933
Loss: 13.50168246005615
l2 norm of gradients: 0.13557537795452082
l2 norm of weights: 2.2910528228896005

Iteration Number: 4934
Loss: 13.499769042172659
l2 norm of gradients: 0.13552913204934366
l2 norm of weights: 2.290931182412213

Iteration Number: 4935
Loss: 13.49785700872539
l2 norm of gradients: 0.13548291263711734
l2 norm of weights: 2.290809664151309

Iteration Number: 4936
Loss: 13.495946358595125
l2 norm of gradients: 0.13543671970238355
l2 norm of weights: 2.2906882680231426

Iteration Number: 4937
Loss: 13.49403709066356
l2 norm of gradients: 0.13539055322969695
l2 norm of weights: 2.2905669939440148

Iteration Number: 4938
Loss: 13.492129203813313
l2 norm of gradients: 0.13534441320362497
l2 norm of weights: 2.290445841830272

Iteration Number: 4939
Loss: 13.490222696927923
l2 norm of gradients: 0.13529829960874803
l2 norm of weights: 2.2903248115983064

Iteration Number: 4940
Loss: 13.48831756889184
l2 norm of gradients: 0.13525221242965924
l2 norm of weights: 2.290203903164559

Iteration Number: 4941
Loss: 13.486413818590446
l2 norm of gradients: 0.13520615165096467
l2 norm of weights: 2.2900831164455155

```

-----
Iteration Number: 4942
Loss: 13.484511444910025
l2 norm of gradients: 0.13516011725728302
l2 norm of weights: 2.289962451357708
-----
Iteration Number: 4943
Loss: 13.482610446737791
l2 norm of gradients: 0.13511410923324585
l2 norm of weights: 2.289841907817716
-----
Iteration Number: 4944
Loss: 13.480710822961864
l2 norm of gradients: 0.1350681275634973
l2 norm of weights: 2.289721485742164
-----
Iteration Number: 4945
Loss: 13.478812572471286
l2 norm of gradients: 0.1350221722326942
l2 norm of weights: 2.2896011850477245
-----
Iteration Number: 4946
Loss: 13.476915694156009
l2 norm of gradients: 0.13497624322550614
l2 norm of weights: 2.289481005651115
-----
Iteration Number: 4947
Loss: 13.4750201869069
l2 norm of gradients: 0.13493034052661523
l2 norm of weights: 2.2893609474691003
-----
Iteration Number: 4948
Loss: 13.473126049615738
l2 norm of gradients: 0.13488446412071617
l2 norm of weights: 2.289241010418491
-----
Iteration Number: 4949
Loss: 13.471233281175214
l2 norm of gradients: 0.1348386139925162
l2 norm of weights: 2.2891211944161447
-----
Iteration Number: 4950
Loss: 13.469341880478934
l2 norm of gradients: 0.1347927901267351
l2 norm of weights: 2.2890014993789647
-----
Iteration Number: 4951
Loss: 13.46745184642141

```

```

12 norm of gradients: 0.13474699250810504
12 norm of weights: 2.2888819252239014
-----
Iteration Number: 4952
Loss: 13.465563177898062
12 norm of gradients: 0.1347012211213708
12 norm of weights: 2.288762471867951
-----
Iteration Number: 4953
Loss: 13.463675873805226
12 norm of gradients: 0.1346554759512895
12 norm of weights: 2.288643139228157
-----
Iteration Number: 4954
Loss: 13.461789933040139
12 norm of gradients: 0.13460975698263056
12 norm of weights: 2.2885239272216076
-----
Iteration Number: 4955
Loss: 13.459905354500947
12 norm of gradients: 0.13456406420017594
12 norm of weights: 2.288404835765439
-----
Iteration Number: 4956
Loss: 13.458022137086713
12 norm of gradients: 0.13451839758871972
12 norm of weights: 2.2882858647768325
-----
Iteration Number: 4957
Loss: 13.456140279697388
12 norm of gradients: 0.13447275713306842
12 norm of weights: 2.2881670141730166
-----
Iteration Number: 4958
Loss: 13.454259781233837
12 norm of gradients: 0.1344271428180408
12 norm of weights: 2.288048283871267
-----
Iteration Number: 4959
Loss: 13.45238064059784
12 norm of gradients: 0.1343815546284678
12 norm of weights: 2.287929673788903
-----
Iteration Number: 4960
Loss: 13.450502856692061
12 norm of gradients: 0.13433599254919254
12 norm of weights: 2.2878111838432926
-----

```

Iteration Number: 4961
Loss: 13.448626428420077
l2 norm of gradients: 0.13429045656507035
l2 norm of weights: 2.28769281395185

Iteration Number: 4962
Loss: 13.446751354686368
l2 norm of gradients: 0.13424494666096873
l2 norm of weights: 2.287574564032035

Iteration Number: 4963
Loss: 13.444877634396313
l2 norm of gradients: 0.13419946282176717
l2 norm of weights: 2.287456434001354

Iteration Number: 4964
Loss: 13.443005266456199
l2 norm of gradients: 0.13415400503235733
l2 norm of weights: 2.2873384237773595

Iteration Number: 4965
Loss: 13.441134249773194
l2 norm of gradients: 0.1341085732776428
l2 norm of weights: 2.28722053327765

Iteration Number: 4966
Loss: 13.439264583255387
l2 norm of gradients: 0.1340631675425393
l2 norm of weights: 2.287102762419872

Iteration Number: 4967
Loss: 13.437396265811753
l2 norm of gradients: 0.13401778781197443
l2 norm of weights: 2.286985111121717

Iteration Number: 4968
Loss: 13.435529296352165
l2 norm of gradients: 0.13397243407088777
l2 norm of weights: 2.2868675793009223

Iteration Number: 4969
Loss: 13.433663673787402
l2 norm of gradients: 0.13392710630423077
l2 norm of weights: 2.286750166875273

Iteration Number: 4970
Loss: 13.431799397029131
l2 norm of gradients: 0.13388180449696677

```

12 norm of weights: 2.286632873762599
-----
Iteration Number: 4971
Loss: 13.429936464989911
12 norm of gradients: 0.13383652863407097
12 norm of weights: 2.2865156998807783
-----
Iteration Number: 4972
Loss: 13.428074876583208
12 norm of gradients: 0.1337912787005304
12 norm of weights: 2.2863986451477336
-----
Iteration Number: 4973
Loss: 13.42621463072337
12 norm of gradients: 0.13374605468134376
12 norm of weights: 2.286281709481434
-----
Iteration Number: 4974
Loss: 13.424355726325643
12 norm of gradients: 0.13370085656152164
12 norm of weights: 2.2861648927998957
-----
Iteration Number: 4975
Loss: 13.422498162306171
12 norm of gradients: 0.13365568432608624
12 norm of weights: 2.2860481950211815
-----
Iteration Number: 4976
Loss: 13.42064193758198
12 norm of gradients: 0.13361053796007152
12 norm of weights: 2.285931616063399
-----
Iteration Number: 4977
Loss: 13.418787051070987
12 norm of gradients: 0.13356541744852296
12 norm of weights: 2.2858151558447033
-----
Iteration Number: 4978
Loss: 13.416933501692018
12 norm of gradients: 0.1335203227764978
12 norm of weights: 2.2856988142832955
-----
Iteration Number: 4979
Loss: 13.415081288364764
12 norm of gradients: 0.13347525392906484
12 norm of weights: 2.285582591297423
-----
Iteration Number: 4980

```

Loss: 13.413230410009819
l2 norm of gradients: 0.13343021089130436
l2 norm of weights: 2.2854664868053787

Iteration Number: 4981
Loss: 13.411380865548654
l2 norm of gradients: 0.1333851936483082
l2 norm of weights: 2.2853505007255026

Iteration Number: 4982
Loss: 13.409532653903646
l2 norm of gradients: 0.13334020218517967
l2 norm of weights: 2.2852346329761812

Iteration Number: 4983
Loss: 13.407685773998045
l2 norm of gradients: 0.13329523648703356
l2 norm of weights: 2.285118883475847

Iteration Number: 4984
Loss: 13.405840224755986
l2 norm of gradients: 0.13325029653899612
l2 norm of weights: 2.2850032521429777

Iteration Number: 4985
Loss: 13.403996005102499
l2 norm of gradients: 0.1332053823262049
l2 norm of weights: 2.284887738896099

Iteration Number: 4986
Loss: 13.402153113963484
l2 norm of gradients: 0.1331604938338089
l2 norm of weights: 2.2847723436537817

Iteration Number: 4987
Loss: 13.400311550265736
l2 norm of gradients: 0.1331156310469683
l2 norm of weights: 2.284657066334643

Iteration Number: 4988
Loss: 13.398471312936934
l2 norm of gradients: 0.13307079395085483
l2 norm of weights: 2.284541906857346

Iteration Number: 4989
Loss: 13.396632400905633
l2 norm of gradients: 0.13302598253065118
l2 norm of weights: 2.2844268651406017

```

-----
Iteration Number: 4990
Loss: 13.394794813101274
l2 norm of gradients: 0.13298119677155149
l2 norm of weights: 2.284311941103165
-----
Iteration Number: 4991
Loss: 13.392958548454178
l2 norm of gradients: 0.13293643665876106
l2 norm of weights: 2.284197134663839
-----
Iteration Number: 4992
Loss: 13.391123605895544
l2 norm of gradients: 0.13289170217749624
l2 norm of weights: 2.2840824457414715
-----
Iteration Number: 4993
Loss: 13.38928998435745
l2 norm of gradients: 0.1328469933129846
l2 norm of weights: 2.283967874254957
-----
Iteration Number: 4994
Loss: 13.38745768277286
l2 norm of gradients: 0.1328023100504649
l2 norm of weights: 2.283853420123237
-----
Iteration Number: 4995
Loss: 13.385626700075607
l2 norm of gradients: 0.1327576523751868
l2 norm of weights: 2.283739083265299
-----
Iteration Number: 4996
Loss: 13.383797035200407
l2 norm of gradients: 0.13271302027241105
l2 norm of weights: 2.2836248636001746
-----
Iteration Number: 4997
Loss: 13.381968687082848
l2 norm of gradients: 0.13266841372740945
l2 norm of weights: 2.2835107610469443
-----
Iteration Number: 4998
Loss: 13.380141654659404
l2 norm of gradients: 0.13262383272546477
l2 norm of weights: 2.283396775524734
-----
Iteration Number: 4999
Loss: 13.378315936867411

```



```

12 norm of gradients: 0.13257927725187063
12 norm of weights: 2.283282906952715
-----
Iteration Number: 5000
Loss: 13.376491532645089
12 norm of gradients: 0.13253474729193163
12 norm of weights: 2.2831691552501057
-----
Iteration Number: 5001
Loss: 13.374668440931528
12 norm of gradients: 0.13249024283096328
12 norm of weights: 2.2830555203361698
-----
Iteration Number: 5002
Loss: 13.372846660666692
12 norm of gradients: 0.13244576385429185
12 norm of weights: 2.282942002130218
-----
Iteration Number: 5003
Loss: 13.371026190791415
12 norm of gradients: 0.1324013103472545
12 norm of weights: 2.2828286005516065
-----
Iteration Number: 5004
Loss: 13.36920703024741
12 norm of gradients: 0.13235688229519907
12 norm of weights: 2.282715315519738
-----
Iteration Number: 5005
Loss: 13.367389177977252
12 norm of gradients: 0.1323124796834842
12 norm of weights: 2.2826021469540616
-----
Iteration Number: 5006
Loss: 13.365572632924392
12 norm of gradients: 0.1322681024974793
12 norm of weights: 2.282489094774072
-----
Iteration Number: 5007
Loss: 13.363757394033147
12 norm of gradients: 0.13222375072256431
12 norm of weights: 2.2823761588993103
-----
Iteration Number: 5008
Loss: 13.361943460248714
12 norm of gradients: 0.13217942434413005
12 norm of weights: 2.282263339249364
-----

```

Iteration Number: 5009
Loss: 13.360130830517136
l2 norm of gradients: 0.1321351233475777
l2 norm of weights: 2.282150635743866

Iteration Number: 5010
Loss: 13.358319503785347
l2 norm of gradients: 0.1320908477183193
l2 norm of weights: 2.282038048302496

Iteration Number: 5011
Loss: 13.356509479001131
l2 norm of gradients: 0.1320465974417771
l2 norm of weights: 2.2819255768449795

Iteration Number: 5012
Loss: 13.35470075511315
l2 norm of gradients: 0.13200237250338426
l2 norm of weights: 2.2818132212910887

Iteration Number: 5013
Loss: 13.352893331070925
l2 norm of gradients: 0.13195817288858405
l2 norm of weights: 2.2817009815606406

Iteration Number: 5014
Loss: 13.351087205824848
l2 norm of gradients: 0.13191399858283054
l2 norm of weights: 2.2815888575735

Iteration Number: 5015
Loss: 13.349282378326164
l2 norm of gradients: 0.13186984957158798
l2 norm of weights: 2.2814768492495756

Iteration Number: 5016
Loss: 13.34747884752699
l2 norm of gradients: 0.1318257258403311
l2 norm of weights: 2.281364956508825

Iteration Number: 5017
Loss: 13.3456766123803
l2 norm of gradients: 0.13178162737454496
l2 norm of weights: 2.28125317927125

Iteration Number: 5018
Loss: 13.343875671839944
l2 norm of gradients: 0.1317375541597251

```

12 norm of weights: 2.281141517456898
-----
Iteration Number: 5019
Loss: 13.342076024860612
12 norm of gradients: 0.13169350618137715
12 norm of weights: 2.281029970985864
-----
Iteration Number: 5020
Loss: 13.340277670397873
12 norm of gradients: 0.1316494834250171
12 norm of weights: 2.2809185397782885
-----
Iteration Number: 5021
Loss: 13.338480607408144
12 norm of gradients: 0.1316054858761711
12 norm of weights: 2.280807223754358
-----
Iteration Number: 5022
Loss: 13.336684834848711
12 norm of gradients: 0.1315615135203757
12 norm of weights: 2.280696022834305
-----
Iteration Number: 5023
Loss: 13.334890351677707
12 norm of gradients: 0.13151756634317738
12 norm of weights: 2.2805849369384075
-----
Iteration Number: 5024
Loss: 13.333097156854137
12 norm of gradients: 0.13147364433013295
12 norm of weights: 2.2804739659869906
-----
Iteration Number: 5025
Loss: 13.331305249337852
12 norm of gradients: 0.1314297474668092
12 norm of weights: 2.2803631099004256
-----
Iteration Number: 5026
Loss: 13.329514628089571
12 norm of gradients: 0.13138587573878296
12 norm of weights: 2.280252368599128
-----
Iteration Number: 5027
Loss: 13.32772529207085
12 norm of gradients: 0.1313420291316413
12 norm of weights: 2.2801417420035617
-----
Iteration Number: 5028

```

Loss: 13.325937240244116
l2 norm of gradients: 0.13129820763098113
l2 norm of weights: 2.2800312300342345

Iteration Number: 5029
Loss: 13.324150471572654
l2 norm of gradients: 0.13125441122240936
l2 norm of weights: 2.2799208326117015

Iteration Number: 5030
Loss: 13.322364985020586
l2 norm of gradients: 0.13121063989154289
l2 norm of weights: 2.2798105496565637

Iteration Number: 5031
Loss: 13.3205807795529
l2 norm of gradients: 0.13116689362400855
l2 norm of weights: 2.279700381089468

Iteration Number: 5032
Loss: 13.318797854135436
l2 norm of gradients: 0.13112317240544294
l2 norm of weights: 2.279590326831107

Iteration Number: 5033
Loss: 13.317016207734882
l2 norm of gradients: 0.1310794762214927
l2 norm of weights: 2.279480386802219

Iteration Number: 5034
Loss: 13.315235839318774
l2 norm of gradients: 0.13103580505781412
l2 norm of weights: 2.2793705609235895

Iteration Number: 5035
Loss: 13.313456747855511
l2 norm of gradients: 0.1309921589000734
l2 norm of weights: 2.279260849116049

Iteration Number: 5036
Loss: 13.311678932314326
l2 norm of gradients: 0.13094853773394638
l2 norm of weights: 2.2791512513004735

Iteration Number: 5037
Loss: 13.309902391665313
l2 norm of gradients: 0.1309049415451187
l2 norm of weights: 2.2790417673977865

```

-----
Iteration Number: 5038
Loss: 13.308127124879414
l2 norm of gradients: 0.1308613703192857
l2 norm of weights: 2.278932397328957
-----
Iteration Number: 5039
Loss: 13.30635313092841
l2 norm of gradients: 0.13081782404215245
l2 norm of weights: 2.278823141014999
-----
Iteration Number: 5040
Loss: 13.304580408784938
l2 norm of gradients: 0.13077430269943344
l2 norm of weights: 2.278713998376973
-----
Iteration Number: 5041
Loss: 13.302808957422481
l2 norm of gradients: 0.130730806276853
l2 norm of weights: 2.2786049693359853
-----
Iteration Number: 5042
Loss: 13.30103877581536
l2 norm of gradients: 0.13068733476014482
l2 norm of weights: 2.2784960538131886
-----
Iteration Number: 5043
Loss: 13.299269862938749
l2 norm of gradients: 0.1306438881350523
l2 norm of weights: 2.278387251729782
-----
Iteration Number: 5044
Loss: 13.29750221776867
l2 norm of gradients: 0.13060046638732822
l2 norm of weights: 2.2782785630070084
-----
Iteration Number: 5045
Loss: 13.295735839281974
l2 norm of gradients: 0.13055706950273493
l2 norm of weights: 2.2781699875661587
-----
Iteration Number: 5046
Loss: 13.293970726456372
l2 norm of gradients: 0.13051369746704422
l2 norm of weights: 2.278061525328569
-----
Iteration Number: 5047
Loss: 13.292206878270408

```

```

12 norm of gradients: 0.13047035026603723
12 norm of weights: 2.2779531762156218
-----
Iteration Number: 5048
Loss: 13.290444293703468
12 norm of gradients: 0.13042702788550445
12 norm of weights: 2.277844940148744
-----
Iteration Number: 5049
Loss: 13.288682971735788
12 norm of gradients: 0.13038373031124592
12 norm of weights: 2.27773681704941
-----
Iteration Number: 5050
Loss: 13.286922911348434
12 norm of gradients: 0.13034045752907072
12 norm of weights: 2.277628806839139
-----
Iteration Number: 5051
Loss: 13.285164111523319
12 norm of gradients: 0.13029720952479748
12 norm of weights: 2.277520909439497
-----
Iteration Number: 5052
Loss: 13.283406571243187
12 norm of gradients: 0.13025398628425391
12 norm of weights: 2.277413124772096
-----
Iteration Number: 5053
Loss: 13.281650289491633
12 norm of gradients: 0.130210787793277
12 norm of weights: 2.2773054527585916
-----
Iteration Number: 5054
Loss: 13.279895265253082
12 norm of gradients: 0.13016761403771296
12 norm of weights: 2.2771978933206882
-----
Iteration Number: 5055
Loss: 13.278141497512799
12 norm of gradients: 0.1301244650034172
12 norm of weights: 2.2770904463801345
-----
Iteration Number: 5056
Loss: 13.276388985256883
12 norm of gradients: 0.13008134067625415
12 norm of weights: 2.2769831118587254
-----

```

Iteration Number: 5057
Loss: 13.274637727472271
l2 norm of gradients: 0.13003824104209746
l2 norm of weights: 2.2768758896783012

Iteration Number: 5058
Loss: 13.27288772314674
l2 norm of gradients: 0.1299951660868298
l2 norm of weights: 2.2767687797607485

Iteration Number: 5059
Loss: 13.271138971268893
l2 norm of gradients: 0.12995211579634283
l2 norm of weights: 2.276661782028

Iteration Number: 5060
Loss: 13.269391470828177
l2 norm of gradients: 0.12990909015653734
l2 norm of weights: 2.2765548964020335

Iteration Number: 5061
Loss: 13.267645220814861
l2 norm of gradients: 0.12986608915332307
l2 norm of weights: 2.2764481228048727

Iteration Number: 5062
Loss: 13.26590022022006
l2 norm of gradients: 0.12982311277261857
l2 norm of weights: 2.276341461158588

Iteration Number: 5063
Loss: 13.264156468035706
l2 norm of gradients: 0.12978016100035153
l2 norm of weights: 2.276234911385295

Iteration Number: 5064
Loss: 13.26241396325458
l2 norm of gradients: 0.12973723382245841
l2 norm of weights: 2.2761284734071534

Iteration Number: 5065
Loss: 13.260672704870284
l2 norm of gradients: 0.12969433122488455
l2 norm of weights: 2.2760221471463717

Iteration Number: 5066
Loss: 13.25893269187725
l2 norm of gradients: 0.12965145319358404

```

12 norm of weights: 2.2759159325252027
-----
Iteration Number: 5067
Loss: 13.25719392327074
12 norm of gradients: 0.12960859971451996
12 norm of weights: 2.275809829465945
-----
Iteration Number: 5068
Loss: 13.255456398046851
12 norm of gradients: 0.12956577077366396
12 norm of weights: 2.275703837890942
-----
Iteration Number: 5069
Loss: 13.253720115202501
12 norm of gradients: 0.12952296635699653
12 norm of weights: 2.275597957722585
-----
Iteration Number: 5070
Loss: 13.251985073735442
12 norm of gradients: 0.12948018645050693
12 norm of weights: 2.2754921888833093
-----
Iteration Number: 5071
Loss: 13.250251272644249
12 norm of gradients: 0.1294374310401929
12 norm of weights: 2.2753865312955965
-----
Iteration Number: 5072
Loss: 13.248518710928321
12 norm of gradients: 0.129394700112061
12 norm of weights: 2.2752809848819737
-----
Iteration Number: 5073
Loss: 13.246787387587895
12 norm of gradients: 0.1293519936521264
12 norm of weights: 2.2751755495650143
-----
Iteration Number: 5074
Loss: 13.245057301624017
12 norm of gradients: 0.1293093116464128
12 norm of weights: 2.275070225267337
-----
Iteration Number: 5075
Loss: 13.243328452038574
12 norm of gradients: 0.12926665408095245
12 norm of weights: 2.2749650119116063
-----
Iteration Number: 5076

```


Loss: 13.241600837834264
l2 norm of gradients: 0.12922402094178623
l2 norm of weights: 2.274859909420532

Iteration Number: 5077
Loss: 13.239874458014613
l2 norm of gradients: 0.12918141221496335
l2 norm of weights: 2.27475491771687

Iteration Number: 5078
Loss: 13.238149311583978
l2 norm of gradients: 0.12913882788654166
l2 norm of weights: 2.2746500367234224

Iteration Number: 5079
Loss: 13.23642539754752
l2 norm of gradients: 0.12909626794258736
l2 norm of weights: 2.274545266363036

Iteration Number: 5080
Loss: 13.23470271491124
l2 norm of gradients: 0.12905373236917506
l2 norm of weights: 2.274440606558603

Iteration Number: 5081
Loss: 13.23298126268195
l2 norm of gradients: 0.1290112211523878
l2 norm of weights: 2.274336057233063

Iteration Number: 5082
Loss: 13.231261039867283
l2 norm of gradients: 0.12896873427831695
l2 norm of weights: 2.274231618309399

Iteration Number: 5083
Loss: 13.229542045475696
l2 norm of gradients: 0.1289262717330622
l2 norm of weights: 2.274127289710642

Iteration Number: 5084
Loss: 13.227824278516461
l2 norm of gradients: 0.1288838335027315
l2 norm of weights: 2.2740230713598666

Iteration Number: 5085
Loss: 13.22610773799967
l2 norm of gradients: 0.12884141957344108
l2 norm of weights: 2.273918963180194

```

-----
Iteration Number: 5086
Loss: 13.224392422936234
l2 norm of gradients: 0.12879902993131542
l2 norm of weights: 2.273814965094791
-----
Iteration Number: 5087
Loss: 13.22267833233788
l2 norm of gradients: 0.12875666456248727
l2 norm of weights: 2.27371107702687
-----
Iteration Number: 5088
Loss: 13.220965465217148
l2 norm of gradients: 0.1287143234530974
l2 norm of weights: 2.2736072988996883
-----
Iteration Number: 5089
Loss: 13.219253820587404
l2 norm of gradients: 0.1286720065892949
l2 norm of weights: 2.27350363063655
-----
Iteration Number: 5090
Loss: 13.217543397462817
l2 norm of gradients: 0.1286297139572368
l2 norm of weights: 2.2734000721608036
-----
Iteration Number: 5091
Loss: 13.215834194858378
l2 norm of gradients: 0.12858744554308837
l2 norm of weights: 2.273296623395844
-----
Iteration Number: 5092
Loss: 13.214126211789898
l2 norm of gradients: 0.12854520133302286
l2 norm of weights: 2.273193284265112
-----
Iteration Number: 5093
Loss: 13.212419447273987
l2 norm of gradients: 0.12850298131322158
l2 norm of weights: 2.2730900546920925
-----
Iteration Number: 5094
Loss: 13.21071390032808
l2 norm of gradients: 0.12846078546987386
l2 norm of weights: 2.2729869346003166
-----
Iteration Number: 5095
Loss: 13.209009569970418

```

```

12 norm of gradients: 0.12841861378917696
12 norm of weights: 2.272883923913362
-----
Iteration Number: 5096
Loss: 13.207306455220056
12 norm of gradients: 0.1283764662573361
12 norm of weights: 2.272781022554851
-----
Iteration Number: 5097
Loss: 13.20560455509686
12 norm of gradients: 0.12833434286056444
12 norm of weights: 2.272678230448451
-----
Iteration Number: 5098
Loss: 13.20390386862151
12 norm of gradients: 0.12829224358508295
12 norm of weights: 2.2725755475178757
-----
Iteration Number: 5099
Loss: 13.202204394815485
12 norm of gradients: 0.1282501684171206
12 norm of weights: 2.2724729736868836
-----
Iteration Number: 5100
Loss: 13.200506132701088
12 norm of gradients: 0.12820811734291412
12 norm of weights: 2.27237050887928
-----
Iteration Number: 5101
Loss: 13.198809081301418
12 norm of gradients: 0.128166090348708
12 norm of weights: 2.2722681530189144
-----
Iteration Number: 5102
Loss: 13.197113239640391
12 norm of gradients: 0.12812408742075454
12 norm of weights: 2.2721659060296817
-----
Iteration Number: 5103
Loss: 13.195418606742725
12 norm of gradients: 0.12808210854531368
12 norm of weights: 2.2720637678355238
-----
Iteration Number: 5104
Loss: 13.193725181633951
12 norm of gradients: 0.12804015370865343
12 norm of weights: 2.2719617383604263
-----

```

Iteration Number: 5105
Loss: 13.192032963340395
l2 norm of gradients: 0.127998222897049
l2 norm of weights: 2.271859817528421

Iteration Number: 5106
Loss: 13.190341950889204
l2 norm of gradients: 0.12795631609678368
l2 norm of weights: 2.2717580052635857

Iteration Number: 5107
Loss: 13.188652143308314
l2 norm of gradients: 0.12791443329414812
l2 norm of weights: 2.271656301490043

Iteration Number: 5108
Loss: 13.186963539626479
l2 norm of gradients: 0.12787257447544065
l2 norm of weights: 2.27155470613196

Iteration Number: 5109
Loss: 13.185276138873252
l2 norm of gradients: 0.12783073962696728
l2 norm of weights: 2.271453219113552

Iteration Number: 5110
Loss: 13.183589940078985
l2 norm of gradients: 0.12778892873504147
l2 norm of weights: 2.2713518403590767

Iteration Number: 5111
Loss: 13.181904942274837
l2 norm of gradients: 0.12774714178598426
l2 norm of weights: 2.2712505697928393

Iteration Number: 5112
Loss: 13.180221144492771
l2 norm of gradients: 0.1277053787661241
l2 norm of weights: 2.2711494073391885

Iteration Number: 5113
Loss: 13.17853854576555
l2 norm of gradients: 0.127663639661797
l2 norm of weights: 2.2710483529225205

Iteration Number: 5114
Loss: 13.176857145126728
l2 norm of gradients: 0.12762192445934636

```

12 norm of weights: 2.2709474064672754
-----
Iteration Number: 5115
Loss: 13.175176941610683
12 norm of gradients: 0.12758023314512298
12 norm of weights: 2.2708465678979395
-----
Iteration Number: 5116
Loss: 13.173497934252568
12 norm of gradients: 0.1275385657054852
12 norm of weights: 2.270745837139043
-----
Iteration Number: 5117
Loss: 13.171820122088345
12 norm of gradients: 0.12749692212679845
12 norm of weights: 2.270645214115164
-----
Iteration Number: 5118
Loss: 13.170143504154776
12 norm of gradients: 0.12745530239543562
12 norm of weights: 2.2705446987509235
-----
Iteration Number: 5119
Loss: 13.168468079489426
12 norm of gradients: 0.12741370649777706
12 norm of weights: 2.270444290970989
-----
Iteration Number: 5120
Loss: 13.166793847130645
12 norm of gradients: 0.12737213442021014
12 norm of weights: 2.2703439907000726
-----
Iteration Number: 5121
Loss: 13.165120806117582
12 norm of gradients: 0.12733058614912968
12 norm of weights: 2.2702437978629333
-----
Iteration Number: 5122
Loss: 13.163448955490196
12 norm of gradients: 0.12728906167093754
12 norm of weights: 2.2701437123843733
-----
Iteration Number: 5123
Loss: 13.16177829428922
12 norm of gradients: 0.12724756097204293
12 norm of weights: 2.270043734189242
-----
Iteration Number: 5124

```

Loss: 13.160108821556207
l2 norm of gradients: 0.1272060840388622
l2 norm of weights: 2.269943863202433

Iteration Number: 5125
Loss: 13.15844053633348
l2 norm of gradients: 0.12716463085781873
l2 norm of weights: 2.2698440993488855

Iteration Number: 5126
Loss: 13.156773437664173
l2 norm of gradients: 0.1271232014153432
l2 norm of weights: 2.2697444425535833

Iteration Number: 5127
Loss: 13.155107524592205
l2 norm of gradients: 0.12708179569787323
l2 norm of weights: 2.2696448927415562

Iteration Number: 5128
Loss: 13.15344279616229
l2 norm of gradients: 0.1270404136918536
l2 norm of weights: 2.2695454498378793

Iteration Number: 5129
Loss: 13.151779251419935
l2 norm of gradients: 0.1269990553837361
l2 norm of weights: 2.269446113767672

Iteration Number: 5130
Loss: 13.150116889411438
l2 norm of gradients: 0.1269577207599794
l2 norm of weights: 2.269346884456101

Iteration Number: 5131
Loss: 13.148455709183883
l2 norm of gradients: 0.12691640980704946
l2 norm of weights: 2.269247761828376

Iteration Number: 5132
Loss: 13.146795709785152
l2 norm of gradients: 0.12687512251141886
l2 norm of weights: 2.2691487458097526

Iteration Number: 5133
Loss: 13.14513689026392
l2 norm of gradients: 0.12683385885956733
l2 norm of weights: 2.269049836325532

```

-----
Iteration Number: 5134
Loss: 13.143479249669635
l2 norm of gradients: 0.12679261883798146
l2 norm of weights: 2.2689510333010605
-----
Iteration Number: 5135
Loss: 13.141822787052549
l2 norm of gradients: 0.12675140243315464
l2 norm of weights: 2.268852336661729
-----
Iteration Number: 5136
Loss: 13.14016750146369
l2 norm of gradients: 0.1267102096315872
l2 norm of weights: 2.2687537463329743
-----
Iteration Number: 5137
Loss: 13.138513391954888
l2 norm of gradients: 0.1266690404197863
l2 norm of weights: 2.2686552622402782
-----
Iteration Number: 5138
Loss: 13.136860457578747
l2 norm of gradients: 0.12662789478426592
l2 norm of weights: 2.2685568843091675
-----
Iteration Number: 5139
Loss: 13.135208697388663
l2 norm of gradients: 0.12658677271154664
l2 norm of weights: 2.268458612465214
-----
Iteration Number: 5140
Loss: 13.133558110438816
l2 norm of gradients: 0.12654567418815604
l2 norm of weights: 2.2683604466340346
-----
Iteration Number: 5141
Loss: 13.13190869578418
l2 norm of gradients: 0.12650459920062831
l2 norm of weights: 2.268262386741292
-----
Iteration Number: 5142
Loss: 13.130260452480492
l2 norm of gradients: 0.12646354773550433
l2 norm of weights: 2.268164432712693
-----
Iteration Number: 5143
Loss: 13.128613379584294

```

```
12 norm of gradients: 0.12642251977933167
12 norm of weights: 2.2680665844739902
-----
Iteration Number: 5144
Loss: 13.126967476152906
12 norm of gradients: 0.12638151531866457
12 norm of weights: 2.267968841950981
-----
Iteration Number: 5145
Loss: 13.125322741244428
12 norm of gradients: 0.12634053434006387
12 norm of weights: 2.2678712050695085
-----
Iteration Number: 5146
Loss: 13.123679173917742
12 norm of gradients: 0.12629957683009704
12 norm of weights: 2.26777367375546
-----
Iteration Number: 5147
Loss: 13.122036773232509
12 norm of gradients: 0.1262586427753381
12 norm of weights: 2.2676762479347676
-----
Iteration Number: 5148
Loss: 13.12039553824918
12 norm of gradients: 0.12621773216236762
12 norm of weights: 2.2675789275334104
-----
Iteration Number: 5149
Loss: 13.118755468028985
12 norm of gradients: 0.12617684497777276
12 norm of weights: 2.26748171247741
-----
Iteration Number: 5150
Loss: 13.117116561633921
12 norm of gradients: 0.12613598120814704
12 norm of weights: 2.2673846026928346
-----
Iteration Number: 5151
Loss: 13.115478818126782
12 norm of gradients: 0.12609514084009057
12 norm of weights: 2.2672875981057974
-----
Iteration Number: 5152
Loss: 13.113842236571134
12 norm of gradients: 0.12605432386020998
12 norm of weights: 2.2671906986424557
-----
```


Iteration Number: 5153
Loss: 13.112206816031314
l2 norm of gradients: 0.12601353025511813
l2 norm of weights: 2.267093904229013

Iteration Number: 5154
Loss: 13.110572555572448
l2 norm of gradients: 0.1259727600114344
l2 norm of weights: 2.2669972147917163

Iteration Number: 5155
Loss: 13.108939454260437
l2 norm of gradients: 0.1259320131157846
l2 norm of weights: 2.2669006302568597

Iteration Number: 5156
Loss: 13.107307511161949
l2 norm of gradients: 0.1258912895548008
l2 norm of weights: 2.2668041505507794

Iteration Number: 5157
Loss: 13.105676725344438
l2 norm of gradients: 0.12585058931512141
l2 norm of weights: 2.2667077755998597

Iteration Number: 5158
Loss: 13.104047095876133
l2 norm of gradients: 0.1258099123833912
l2 norm of weights: 2.266611505330527

Iteration Number: 5159
Loss: 13.102418621826033
l2 norm of gradients: 0.12576925874626124
l2 norm of weights: 2.2665153396692546

Iteration Number: 5160
Loss: 13.100791302263914
l2 norm of gradients: 0.1257286283903887
l2 norm of weights: 2.2664192785425596

Iteration Number: 5161
Loss: 13.099165136260332
l2 norm of gradients: 0.1256880213024372
l2 norm of weights: 2.266323321877005

Iteration Number: 5162
Loss: 13.097540122886599
l2 norm of gradients: 0.12564743746907647

```

12 norm of weights: 2.2662274695991975
-----
Iteration Number: 5163
Loss: 13.095916261214823
12 norm of gradients: 0.12560687687698233
12 norm of weights: 2.26613172163579
-----
Iteration Number: 5164
Loss: 13.094293550317863
12 norm of gradients: 0.12556633951283697
12 norm of weights: 2.2660360779134794
-----
Iteration Number: 5165
Loss: 13.09267198926936
12 norm of gradients: 0.12552582536332854
12 norm of weights: 2.265940538359007
-----
Iteration Number: 5166
Loss: 13.091051577143729
12 norm of gradients: 0.12548533441515144
12 norm of weights: 2.2658451028991613
-----
Iteration Number: 5167
Loss: 13.089432313016147
12 norm of gradients: 0.12544486665500607
12 norm of weights: 2.2657497714607726
-----
Iteration Number: 5168
Loss: 13.087814195962567
12 norm of gradients: 0.12540442206959898
12 norm of weights: 2.2656545439707174
-----
Iteration Number: 5169
Loss: 13.086197225059706
12 norm of gradients: 0.12536400064564263
12 norm of weights: 2.2655594203559177
-----
Iteration Number: 5170
Loss: 13.08458139938506
12 norm of gradients: 0.12532360236985568
12 norm of weights: 2.2654644005433395
-----
Iteration Number: 5171
Loss: 13.08296671801688
12 norm of gradients: 0.12528322722896273
12 norm of weights: 2.2653694844599936
-----
Iteration Number: 5172

```

```

Loss: 13.0813531800342
12 norm of gradients: 0.1252428752096943
12 norm of weights: 2.265274672032936
-----
Iteration Number: 5173
Loss: 13.079740784516803
12 norm of gradients: 0.12520254629878688
12 norm of weights: 2.2651799631892673
-----
Iteration Number: 5174
Loss: 13.07812953054525
12 norm of gradients: 0.125162240482983
12 norm of weights: 2.2650853578561323
-----
Iteration Number: 5175
Loss: 13.076519417200865
12 norm of gradients: 0.12512195774903093
12 norm of weights: 2.2649908559607215
-----
Iteration Number: 5176
Loss: 13.074910443565749
12 norm of gradients: 0.125081698083685
12 norm of weights: 2.26489645743027
-----
Iteration Number: 5177
Loss: 13.073302608722743
12 norm of gradients: 0.12504146147370526
12 norm of weights: 2.264802162192056
-----
Iteration Number: 5178
Loss: 13.071695911755475
12 norm of gradients: 0.12500124790585773
12 norm of weights: 2.2647079701734056
-----
Iteration Number: 5179
Loss: 13.07009035174833
12 norm of gradients: 0.1249610573669142
12 norm of weights: 2.2646138813016865
-----
Iteration Number: 5180
Loss: 13.068485927786448
12 norm of gradients: 0.12492088984365217
12 norm of weights: 2.264519895504313
-----
Iteration Number: 5181
Loss: 13.066882638955745
12 norm of gradients: 0.12488074532285504
12 norm of weights: 2.264426012708743

```

```
-----  
Iteration Number: 5182  
Loss: 13.06528048434289  
l2 norm of gradients: 0.12484062379131197  
l2 norm of weights: 2.2643322328424804  
-----  
Iteration Number: 5183  
Loss: 13.063679463035323  
l2 norm of gradients: 0.12480052523581774  
l2 norm of weights: 2.2642385558330718  
-----  
Iteration Number: 5184  
Loss: 13.062079574121231  
l2 norm of gradients: 0.12476044964317298  
l2 norm of weights: 2.2641449816081107  
-----  
Iteration Number: 5185  
Loss: 13.060480816689571  
l2 norm of gradients: 0.1247203970001839  
l2 norm of weights: 2.264051510095233  
-----  
Iteration Number: 5186  
Loss: 13.058883189830063  
l2 norm of gradients: 0.12468036729366241  
l2 norm of weights: 2.2639581412221212  
-----  
Iteration Number: 5187  
Loss: 13.057286692633175  
l2 norm of gradients: 0.12464036051042608  
l2 norm of weights: 2.2638648749165013  
-----  
Iteration Number: 5188  
Loss: 13.055691324190146  
l2 norm of gradients: 0.12460037663729814  
l2 norm of weights: 2.263771711106144  
-----  
Iteration Number: 5189  
Loss: 13.054097083592968  
l2 norm of gradients: 0.12456041566110736  
l2 norm of weights: 2.263678649718865  
-----  
Iteration Number: 5190  
Loss: 13.05250396993439  
l2 norm of gradients: 0.12452047756868811  
l2 norm of weights: 2.2635856906825245  
-----  
Iteration Number: 5191  
Loss: 13.05091198230792
```

```

12 norm of gradients: 0.12448056234688029
12 norm of weights: 2.2634928339250266
-----
Iteration Number: 5192
Loss: 13.049321119807816
12 norm of gradients: 0.12444066998252945
12 norm of weights: 2.2634000793743208
-----
Iteration Number: 5193
Loss: 13.047731381529111
12 norm of gradients: 0.12440080046248665
12 norm of weights: 2.263307426958401
-----
Iteration Number: 5194
Loss: 13.04614276656757
12 norm of gradients: 0.12436095377360824
12 norm of weights: 2.2632148766053053
-----
Iteration Number: 5195
Loss: 13.04455527401973
12 norm of gradients: 0.12432112990275625
12 norm of weights: 2.263122428243116
-----
Iteration Number: 5196
Loss: 13.042968902982878
12 norm of gradients: 0.12428132883679817
12 norm of weights: 2.263030081799961
-----
Iteration Number: 5197
Loss: 13.04138365255505
12 norm of gradients: 0.12424155056260683
12 norm of weights: 2.2629378372040128
-----
Iteration Number: 5198
Loss: 13.03979952183504
12 norm of gradients: 0.12420179506706054
12 norm of weights: 2.262845694383486
-----
Iteration Number: 5199
Loss: 13.038216509922394
12 norm of gradients: 0.12416206233704294
12 norm of weights: 2.2627536532666426
-----
Iteration Number: 5200
Loss: 13.036634615917414
12 norm of gradients: 0.12412235235944313
12 norm of weights: 2.2626617137817875
-----

```

Iteration Number: 5201
Loss: 13.035053838921149
l2 norm of gradients: 0.1240826651211555
l2 norm of weights: 2.2625698758572703

Iteration Number: 5202
Loss: 13.033474178035405
l2 norm of gradients: 0.12404300060907984
l2 norm of weights: 2.262478139421485

Iteration Number: 5203
Loss: 13.03189563236273
l2 norm of gradients: 0.1240033588101212
l2 norm of weights: 2.2623865044028704

Iteration Number: 5204
Loss: 13.030318201006427
l2 norm of gradients: 0.12396373971118993
l2 norm of weights: 2.2622949707299096

Iteration Number: 5205
Loss: 13.028741883070554
l2 norm of gradients: 0.12392414329920165
l2 norm of weights: 2.2622035383311294

Iteration Number: 5206
Loss: 13.02716667765991
l2 norm of gradients: 0.12388456956107727
l2 norm of weights: 2.262112207135102

Iteration Number: 5207
Loss: 13.025592583880048
l2 norm of gradients: 0.1238450184837429
l2 norm of weights: 2.262020977070444

Iteration Number: 5208
Loss: 13.024019600837267
l2 norm of gradients: 0.1238054900541299
l2 norm of weights: 2.2619298480658148

Iteration Number: 5209
Loss: 13.022447727638614
l2 norm of gradients: 0.12376598425917484
l2 norm of weights: 2.2618388200499195

Iteration Number: 5210
Loss: 13.020876963391881
l2 norm of gradients: 0.12372650108581935

```
l2 norm of weights: 2.261747892951508
-----
Iteration Number: 5211
Loss: 13.019307307205613
l2 norm of gradients: 0.12368704052101044
l2 norm of weights: 2.2616570666993736
-----
Iteration Number: 5212
Loss: 13.017738758189093
l2 norm of gradients: 0.12364760255169999
l2 norm of weights: 2.2615663412223537
-----
Iteration Number: 5213
Loss: 13.016171315452352
l2 norm of gradients: 0.12360818716484516
l2 norm of weights: 2.2614757164493304
-----
Iteration Number: 5214
Loss: 13.014604978106174
l2 norm of gradients: 0.12356879434740826
l2 norm of weights: 2.2613851923092305
-----
Iteration Number: 5215
Loss: 13.013039745262077
l2 norm of gradients: 0.12352942408635656
l2 norm of weights: 2.2612947687310245
-----
Iteration Number: 5216
Loss: 13.011475616032326
l2 norm of gradients: 0.12349007636866241
l2 norm of weights: 2.2612044456437275
-----
Iteration Number: 5217
Loss: 13.00991258952993
l2 norm of gradients: 0.12345075118130323
l2 norm of weights: 2.2611142229763983
-----
Iteration Number: 5218
Loss: 13.00835066486864
l2 norm of gradients: 0.12341144851126153
l2 norm of weights: 2.261024100658141
-----
Iteration Number: 5219
Loss: 13.006789841162957
l2 norm of gradients: 0.12337216834552471
l2 norm of weights: 2.2609340786181025
-----
Iteration Number: 5220
```

```
Loss: 13.005230117528109
12 norm of gradients: 0.12333291067108523
12 norm of weights: 2.2608441567854753
-----
Iteration Number: 5221
Loss: 13.003671493080079
12 norm of gradients: 0.12329367547494051
12 norm of weights: 2.260754335089495
-----
Iteration Number: 5222
Loss: 13.002113966935585
12 norm of gradients: 0.12325446274409288
12 norm of weights: 2.260664613459442
-----
Iteration Number: 5223
Loss: 13.000557538212078
12 norm of gradients: 0.12321527246554968
12 norm of weights: 2.260574991824641
-----
Iteration Number: 5224
Loss: 12.999002206027766
12 norm of gradients: 0.1231761046263231
12 norm of weights: 2.2604854701144608
-----
Iteration Number: 5225
Loss: 12.997447969501582
12 norm of gradients: 0.12313695921343032
12 norm of weights: 2.260396048258313
-----
Iteration Number: 5226
Loss: 12.995894827753208
12 norm of gradients: 0.12309783621389322
12 norm of weights: 2.260306726185656
-----
Iteration Number: 5227
Loss: 12.994342779903045
12 norm of gradients: 0.1230587356147387
12 norm of weights: 2.2602175038259893
-----
Iteration Number: 5228
Loss: 12.99279182507226
12 norm of gradients: 0.12301965740299844
12 norm of weights: 2.260128381108859
-----
Iteration Number: 5229
Loss: 12.991241962382729
12 norm of gradients: 0.12298060156570899
12 norm of weights: 2.260039357963854
```



```

-----
Iteration Number: 5230
Loss: 12.989693190957084
l2 norm of gradients: 0.12294156808991169
l2 norm of weights: 2.259950434320608
-----
Iteration Number: 5231
Loss: 12.988145509918692
l2 norm of gradients: 0.12290255696265262
l2 norm of weights: 2.259861610108797
-----
Iteration Number: 5232
Loss: 12.986598918391639
l2 norm of gradients: 0.12286356817098269
l2 norm of weights: 2.2597728852581436
-----
Iteration Number: 5233
Loss: 12.985053415500769
l2 norm of gradients: 0.12282460170195757
l2 norm of weights: 2.2596842596984135
-----
Iteration Number: 5234
Loss: 12.98350900037164
l2 norm of gradients: 0.12278565754263761
l2 norm of weights: 2.2595957333594145
-----
Iteration Number: 5235
Loss: 12.981965672130553
l2 norm of gradients: 0.12274673568008797
l2 norm of weights: 2.259507306171002
-----
Iteration Number: 5236
Loss: 12.980423429904548
l2 norm of gradients: 0.12270783610137846
l2 norm of weights: 2.2594189780630725
-----
Iteration Number: 5237
Loss: 12.978882272821389
l2 norm of gradients: 0.1226689587935836
l2 norm of weights: 2.2593307489655667
-----
Iteration Number: 5238
Loss: 12.977342200009572
l2 norm of gradients: 0.12263010374378251
l2 norm of weights: 2.2592426188084715
-----
Iteration Number: 5239
Loss: 12.975803210598336

```

```

12 norm of gradients: 0.1225912709390591
12 norm of weights: 2.2591545875218157
-----
Iteration Number: 5240
Loss: 12.974265303717637
12 norm of gradients: 0.12255246036650183
12 norm of weights: 2.2590666550356726
-----
Iteration Number: 5241
Loss: 12.972728478498174
12 norm of gradients: 0.12251367201320376
12 norm of weights: 2.258978821280159
-----
Iteration Number: 5242
Loss: 12.97119273407137
12 norm of gradients: 0.12247490586626264
12 norm of weights: 2.2588910861854368
-----
Iteration Number: 5243
Loss: 12.969658069569377
12 norm of gradients: 0.12243616191278074
12 norm of weights: 2.2588034496817104
-----
Iteration Number: 5244
Loss: 12.968124484125077
12 norm of gradients: 0.12239744013986488
12 norm of weights: 2.25871591169923
-----
Iteration Number: 5245
Loss: 12.966591976872087
12 norm of gradients: 0.12235874053462663
12 norm of weights: 2.2586284721682865
-----
Iteration Number: 5246
Loss: 12.965060546944741
12 norm of gradients: 0.1223200630841818
12 norm of weights: 2.2585411310192183
-----
Iteration Number: 5247
Loss: 12.963530193478112
12 norm of gradients: 0.1222814077756509
12 norm of weights: 2.2584538881824057
-----
Iteration Number: 5248
Loss: 12.962000915607996
12 norm of gradients: 0.12224277459615902
12 norm of weights: 2.2583667435882724
-----

```

Iteration Number: 5249
Loss: 12.960472712470912
l2 norm of gradients: 0.12220416353283559
l2 norm of weights: 2.2582796971672865

Iteration Number: 5250
Loss: 12.958945583204109
l2 norm of gradients: 0.12216557457281461
l2 norm of weights: 2.2581927488499605

Iteration Number: 5251
Loss: 12.957419526945563
l2 norm of gradients: 0.12212700770323444
l2 norm of weights: 2.2581058985668507

Iteration Number: 5252
Loss: 12.955894542833974
l2 norm of gradients: 0.12208846291123807
l2 norm of weights: 2.2580191462485555

Iteration Number: 5253
Loss: 12.95437063000876
l2 norm of gradients: 0.12204994018397275
l2 norm of weights: 2.2579324918257195

Iteration Number: 5254
Loss: 12.952847787610075
l2 norm of gradients: 0.12201143950859022
l2 norm of weights: 2.2578459352290285

Iteration Number: 5255
Loss: 12.951326014778793
l2 norm of gradients: 0.12197296087224664
l2 norm of weights: 2.257759476389215

Iteration Number: 5256
Loss: 12.949805310656501
l2 norm of gradients: 0.12193450426210248
l2 norm of weights: 2.257673115237052

Iteration Number: 5257
Loss: 12.948285674385524
l2 norm of gradients: 0.12189606966532265
l2 norm of weights: 2.257586851703358

Iteration Number: 5258
Loss: 12.946767105108895
l2 norm of gradients: 0.1218576570690764

```

12 norm of weights: 2.2575006857189956
-----
Iteration Number: 5259
Loss: 12.945249601970383
12 norm of gradients: 0.12181926646053734
12 norm of weights: 2.25741461721487
-----
Iteration Number: 5260
Loss: 12.943733164114466
12 norm of gradients: 0.12178089782688331
12 norm of weights: 2.257328646121931
-----
Iteration Number: 5261
Loss: 12.942217790686348
12 norm of gradients: 0.12174255115529664
12 norm of weights: 2.2572427723711708
-----
Iteration Number: 5262
Loss: 12.940703480831953
12 norm of gradients: 0.1217042264329638
12 norm of weights: 2.257156995893626
-----
Iteration Number: 5263
Loss: 12.939190233697923
12 norm of gradients: 0.12166592364707561
12 norm of weights: 2.2570713166203773
-----
Iteration Number: 5264
Loss: 12.937678048431621
12 norm of gradients: 0.12162764278482716
12 norm of weights: 2.2569857344825484
-----
Iteration Number: 5265
Loss: 12.936166924181126
12 norm of gradients: 0.12158938383341768
12 norm of weights: 2.256900249411307
-----
Iteration Number: 5266
Loss: 12.934656860095238
12 norm of gradients: 0.12155114678005093
12 norm of weights: 2.256814861337863
-----
Iteration Number: 5267
Loss: 12.933147855323474
12 norm of gradients: 0.12151293161193459
12 norm of weights: 2.2567295701934724
-----
Iteration Number: 5268

```

Loss: 12.931639909016065
l2 norm of gradients: 0.12147473831628072
l2 norm of weights: 2.256644375909432

Iteration Number: 5269
Loss: 12.930133020323964
l2 norm of gradients: 0.12143656688030546
l2 norm of weights: 2.2565592784170843

Iteration Number: 5270
Loss: 12.928627188398835
l2 norm of gradients: 0.12139841729122933
l2 norm of weights: 2.256474277647814

Iteration Number: 5271
Loss: 12.927122412393059
l2 norm of gradients: 0.12136028953627681
l2 norm of weights: 2.25638937353305

Iteration Number: 5272
Loss: 12.925618691459732
l2 norm of gradients: 0.12132218360267667
l2 norm of weights: 2.256304566004264

Iteration Number: 5273
Loss: 12.924116024752665
l2 norm of gradients: 0.12128409947766176
l2 norm of weights: 2.2562198549929717

Iteration Number: 5274
Loss: 12.922614411426386
l2 norm of gradients: 0.1212460371484691
l2 norm of weights: 2.2561352404307327

Iteration Number: 5275
Loss: 12.921113850636129
l2 norm of gradients: 0.12120799660233984
l2 norm of weights: 2.2560507222491486

Iteration Number: 5276
Loss: 12.919614341537848
l2 norm of gradients: 0.12116997782651916
l2 norm of weights: 2.2559663003798662

Iteration Number: 5277
Loss: 12.918115883288207
l2 norm of gradients: 0.12113198080825646
l2 norm of weights: 2.255881974754575

```

-----
Iteration Number: 5278
Loss: 12.916618475044585
l2 norm of gradients: 0.12109400553480511
l2 norm of weights: 2.2557977453050064
-----
Iteration Number: 5279
Loss: 12.915122115965064
l2 norm of gradients: 0.12105605199342254
l2 norm of weights: 2.2557136119629377
-----
Iteration Number: 5280
Loss: 12.913626805208443
l2 norm of gradients: 0.12101812017137024
l2 norm of weights: 2.255629574660188
-----
Iteration Number: 5281
Loss: 12.912132541934238
l2 norm of gradients: 0.1209802100559139
l2 norm of weights: 2.2555456333286203
-----
Iteration Number: 5282
Loss: 12.910639325302661
l2 norm of gradients: 0.12094232163432303
l2 norm of weights: 2.255461787900141
-----
Iteration Number: 5283
Loss: 12.909147154474642
l2 norm of gradients: 0.12090445489387121
l2 norm of weights: 2.2553780383066995
-----
Iteration Number: 5284
Loss: 12.907656028611822
l2 norm of gradients: 0.12086660982183611
l2 norm of weights: 2.255294384480288
-----
Iteration Number: 5285
Loss: 12.906165946876543
l2 norm of gradients: 0.12082878640549932
l2 norm of weights: 2.2552108263529433
-----
Iteration Number: 5286
Loss: 12.904676908431862
l2 norm of gradients: 0.12079098463214641
l2 norm of weights: 2.2551273638567446
-----
Iteration Number: 5287
Loss: 12.903188912441536

```

```
12 norm of gradients: 0.1207532044890669
12 norm of weights: 2.2550439969238147
-----
Iteration Number: 5288
Loss: 12.901701958070039
12 norm of gradients: 0.12071544596355432
12 norm of weights: 2.254960725486319
-----
Iteration Number: 5289
Loss: 12.900216044482546
12 norm of gradients: 0.12067770904290613
12 norm of weights: 2.2548775494764675
-----
Iteration Number: 5290
Loss: 12.89873117084493
12 norm of gradients: 0.12063999371442372
12 norm of weights: 2.254794468826512
-----
Iteration Number: 5291
Loss: 12.89724733632379
12 norm of gradients: 0.12060229996541229
12 norm of weights: 2.254711483468748
-----
Iteration Number: 5292
Loss: 12.89576454008641
12 norm of gradients: 0.12056462778318114
12 norm of weights: 2.2546285933355144
-----
Iteration Number: 5293
Loss: 12.894282781300783
12 norm of gradients: 0.12052697715504329
12 norm of weights: 2.254545798359193
-----
Iteration Number: 5294
Loss: 12.892802059135617
12 norm of gradients: 0.12048934806831578
12 norm of weights: 2.254463098472209
-----
Iteration Number: 5295
Loss: 12.891322372760312
12 norm of gradients: 0.12045174051031944
12 norm of weights: 2.2543804936070306
-----
Iteration Number: 5296
Loss: 12.889843721344974
12 norm of gradients: 0.12041415446837897
12 norm of weights: 2.2542979836961696
-----
```

Iteration Number: 5297
Loss: 12.888366104060411
l2 norm of gradients: 0.12037658992982297
l2 norm of weights: 2.2542155686721794

Iteration Number: 5298
Loss: 12.886889520078137
l2 norm of gradients: 0.12033904688198382
l2 norm of weights: 2.254133248467658

Iteration Number: 5299
Loss: 12.885413968570363
l2 norm of gradients: 0.12030152531219777
l2 norm of weights: 2.2540510230152466

Iteration Number: 5300
Loss: 12.883939448710006
l2 norm of gradients: 0.12026402520780485
l2 norm of weights: 2.253968892247628

Iteration Number: 5301
Loss: 12.882465959670673
l2 norm of gradients: 0.12022654655614895
l2 norm of weights: 2.25388685609753

Iteration Number: 5302
Loss: 12.880993500626687
l2 norm of gradients: 0.12018908934457774
l2 norm of weights: 2.253804914497721

Iteration Number: 5303
Loss: 12.879522070753051
l2 norm of gradients: 0.12015165356044256
l2 norm of weights: 2.2537230673810145

Iteration Number: 5304
Loss: 12.87805166922549
l2 norm of gradients: 0.12011423919109872
l2 norm of weights: 2.253641314680266

Iteration Number: 5305
Loss: 12.876582295220405
l2 norm of gradients: 0.12007684622390513
l2 norm of weights: 2.253559656328375

Iteration Number: 5306
Loss: 12.875113947914915
l2 norm of gradients: 0.12003947464622458


```

12 norm of weights: 2.2534780922582813
-----
Iteration Number: 5307
Loss: 12.873646626486817
12 norm of gradients: 0.12000212444542355
12 norm of weights: 2.2533966224029722
-----
Iteration Number: 5308
Loss: 12.872180330114624
12 norm of gradients: 0.11996479560887213
12 norm of weights: 2.253315246695473
-----
Iteration Number: 5309
Loss: 12.87071505797753
12 norm of gradients: 0.11992748812394441
12 norm of weights: 2.2532339650688558
-----
Iteration Number: 5310
Loss: 12.869250809255437
12 norm of gradients: 0.11989020197801796
12 norm of weights: 2.253152777456233
-----
Iteration Number: 5311
Loss: 12.86778758312893
12 norm of gradients: 0.1198529371584741
12 norm of weights: 2.2530716837907616
-----
Iteration Number: 5312
Loss: 12.866325378779305
12 norm of gradients: 0.11981569365269791
12 norm of weights: 2.25299068400564
-----
Iteration Number: 5313
Loss: 12.86486419538854
12 norm of gradients: 0.11977847144807811
12 norm of weights: 2.252909778034111
-----
Iteration Number: 5314
Loss: 12.863404032139309
12 norm of gradients: 0.11974127053200709
12 norm of weights: 2.252828965809459
-----
Iteration Number: 5315
Loss: 12.861944888214984
12 norm of gradients: 0.11970409089188092
12 norm of weights: 2.252748247265012
-----
Iteration Number: 5316

```

Loss: 12.860486762799628
l2 norm of gradients: 0.11966693251509931
l2 norm of weights: 2.25266762233414

Iteration Number: 5317
Loss: 12.859029655077993
l2 norm of gradients: 0.11962979538906562
l2 norm of weights: 2.2525870909502568

Iteration Number: 5318
Loss: 12.857573564235533
l2 norm of gradients: 0.11959267950118686
l2 norm of weights: 2.252506653046818

Iteration Number: 5319
Loss: 12.85611848945838
l2 norm of gradients: 0.11955558483887367
l2 norm of weights: 2.252426308557322

Iteration Number: 5320
Loss: 12.85466442993337
l2 norm of gradients: 0.11951851138954026
l2 norm of weights: 2.252346057415312

Iteration Number: 5321
Loss: 12.853211384848022
l2 norm of gradients: 0.11948145914060451
l2 norm of weights: 2.2522658995543705

Iteration Number: 5322
Loss: 12.851759353390545
l2 norm of gradients: 0.11944442807948784
l2 norm of weights: 2.252185834908125

Iteration Number: 5323
Loss: 12.850308334749844
l2 norm of gradients: 0.11940741819361532
l2 norm of weights: 2.2521058634102458

Iteration Number: 5324
Loss: 12.848858328115506
l2 norm of gradients: 0.11937042947041562
l2 norm of weights: 2.2520259849944444

Iteration Number: 5325
Loss: 12.847409332677813
l2 norm of gradients: 0.11933346189732087
l2 norm of weights: 2.2519461995944763

```
-----
Iteration Number: 5326
Loss: 12.84596134762773
l2 norm of gradients: 0.11929651546176688
l2 norm of weights: 2.251866507144139
-----
Iteration Number: 5327
Loss: 12.844514372156908
l2 norm of gradients: 0.11925959015119297
l2 norm of weights: 2.2517869075772725
-----
Iteration Number: 5328
Loss: 12.843068405457696
l2 norm of gradients: 0.11922268595304195
l2 norm of weights: 2.2517074008277596
-----
Iteration Number: 5329
Loss: 12.84162344672312
l2 norm of gradients: 0.11918580285476033
l2 norm of weights: 2.251627986829526
-----
Iteration Number: 5330
Loss: 12.840179495146891
l2 norm of gradients: 0.119148940843798
l2 norm of weights: 2.2515486655165406
-----
Iteration Number: 5331
Loss: 12.838736549923418
l2 norm of gradients: 0.11911209990760842
l2 norm of weights: 2.2514694368228123
-----
Iteration Number: 5332
Loss: 12.83729461024778
l2 norm of gradients: 0.1190752800336486
l2 norm of weights: 2.2513903006823957
-----
Iteration Number: 5333
Loss: 12.835853675315752
l2 norm of gradients: 0.11903848120937899
l2 norm of weights: 2.251311257029385
-----
Iteration Number: 5334
Loss: 12.834413744323786
l2 norm of gradients: 0.11900170342226365
l2 norm of weights: 2.2512323057979198
-----
Iteration Number: 5335
Loss: 12.832974816469024
```

```

12 norm of gradients: 0.11896494665976996
12 norm of weights: 2.25115344692218
-----
Iteration Number: 5336
Loss: 12.831536890949284
12 norm of gradients: 0.11892821090936895
12 norm of weights: 2.251074680336389
-----
Iteration Number: 5337
Loss: 12.830099966963076
12 norm of gradients: 0.1188914961585351
12 norm of weights: 2.2509960059748115
-----
Iteration Number: 5338
Loss: 12.828664043709583
12 norm of gradients: 0.11885480239474626
12 norm of weights: 2.250917423771757
-----
Iteration Number: 5339
Loss: 12.827229120388683
12 norm of gradients: 0.1188181296054838
12 norm of weights: 2.250838933661575
-----
Iteration Number: 5340
Loss: 12.825795196200914
12 norm of gradients: 0.11878147777823257
12 norm of weights: 2.2507605355786584
-----
Iteration Number: 5341
Loss: 12.824362270347521
12 norm of gradients: 0.1187448469004809
12 norm of weights: 2.2506822294574422
-----
Iteration Number: 5342
Loss: 12.822930342030403
12 norm of gradients: 0.11870823695972038
12 norm of weights: 2.2506040152324047
-----
Iteration Number: 5343
Loss: 12.821499410452166
12 norm of gradients: 0.11867164794344623
12 norm of weights: 2.250525892838066
-----
Iteration Number: 5344
Loss: 12.820069474816071
12 norm of gradients: 0.11863507983915701
12 norm of weights: 2.250447862208987
-----

```

Iteration Number: 5345
Loss: 12.818640534326072
l2 norm of gradients: 0.11859853263435474
l2 norm of weights: 2.2503699232797736

Iteration Number: 5346
Loss: 12.817212588186798
l2 norm of gradients: 0.1185620063165448
l2 norm of weights: 2.250292075985073

Iteration Number: 5347
Loss: 12.815785635603557
l2 norm of gradients: 0.118525500873236
l2 norm of weights: 2.250214320259573

Iteration Number: 5348
Loss: 12.814359675782331
l2 norm of gradients: 0.11848901629194057
l2 norm of weights: 2.2501366560380056

Iteration Number: 5349
Loss: 12.812934707929786
l2 norm of gradients: 0.11845255256017405
l2 norm of weights: 2.2500590832551453

Iteration Number: 5350
Loss: 12.811510731253257
l2 norm of gradients: 0.11841610966545552
l2 norm of weights: 2.2499816018458074

Iteration Number: 5351
Loss: 12.810087744960757
l2 norm of gradients: 0.11837968759530722
l2 norm of weights: 2.2499042117448504

Iteration Number: 5352
Loss: 12.80866574826098
l2 norm of gradients: 0.11834328633725497
l2 norm of weights: 2.249826912887174

Iteration Number: 5353
Loss: 12.807244740363284
l2 norm of gradients: 0.1183069058788279
l2 norm of weights: 2.2497497052077216

Iteration Number: 5354
Loss: 12.805824720477716
l2 norm of gradients: 0.11827054620755838

```

12 norm of weights: 2.2496725886414777
-----
Iteration Number: 5355
Loss: 12.804405687814986
12 norm of gradients: 0.11823420731098229
12 norm of weights: 2.2495955631234685
-----
Iteration Number: 5356
Loss: 12.802987641586482
12 norm of gradients: 0.11819788917663879
12 norm of weights: 2.249518628588764
-----
Iteration Number: 5357
Loss: 12.80157058100426
12 norm of gradients: 0.11816159179207036
12 norm of weights: 2.2494417849724746
-----
Iteration Number: 5358
Loss: 12.800154505281057
12 norm of gradients: 0.11812531514482287
12 norm of weights: 2.249365032209754
-----
Iteration Number: 5359
Loss: 12.79873941363028
12 norm of gradients: 0.11808905922244547
12 norm of weights: 2.249288370235797
-----
Iteration Number: 5360
Loss: 12.797325305266007
12 norm of gradients: 0.11805282401249066
12 norm of weights: 2.249211798985842
-----
Iteration Number: 5361
Loss: 12.795912179402979
12 norm of gradients: 0.1180166095025143
12 norm of weights: 2.2491353183951666
-----
Iteration Number: 5362
Loss: 12.794500035256622
12 norm of gradients: 0.1179804156800755
12 norm of weights: 2.249058928399094
-----
Iteration Number: 5363
Loss: 12.793088872043022
12 norm of gradients: 0.11794424253273667
12 norm of weights: 2.2489826289329864
-----
Iteration Number: 5364

```

Loss: 12.791678688978939
l2 norm of gradients: 0.11790809004806359
l2 norm of weights: 2.2489064199322497

Iteration Number: 5365
Loss: 12.7902694852818
l2 norm of gradients: 0.11787195821362531
l2 norm of weights: 2.248830301332331

Iteration Number: 5366
Loss: 12.788861260169705
l2 norm of gradients: 0.1178358470169941
l2 norm of weights: 2.2487542730687204

Iteration Number: 5367
Loss: 12.787454012861415
l2 norm of gradients: 0.11779975644574564
l2 norm of weights: 2.2486783350769484

Iteration Number: 5368
Loss: 12.786047742576368
l2 norm of gradients: 0.11776368648745882
l2 norm of weights: 2.2486024872925885

Iteration Number: 5369
Loss: 12.784642448534665
l2 norm of gradients: 0.11772763712971582
l2 norm of weights: 2.2485267296512554

Iteration Number: 5370
Loss: 12.78323812995707
l2 norm of gradients: 0.11769160836010209
l2 norm of weights: 2.248451062088607

Iteration Number: 5371
Loss: 12.78183478606502
l2 norm of gradients: 0.11765560016620637
l2 norm of weights: 2.248375484540341

Iteration Number: 5372
Loss: 12.780432416080613
l2 norm of gradients: 0.11761961253562063
l2 norm of weights: 2.248299996942199

Iteration Number: 5373
Loss: 12.779031019226617
l2 norm of gradients: 0.1175836454559401
l2 norm of weights: 2.248224599229963

```

-----
Iteration Number: 5374
Loss: 12.777630594726459
l2 norm of gradients: 0.11754769891476326
l2 norm of weights: 2.2481492913394576
-----
Iteration Number: 5375
Loss: 12.776231141804235
l2 norm of gradients: 0.11751177289969192
l2 norm of weights: 2.2480740732065483
-----
Iteration Number: 5376
Loss: 12.774832659684709
l2 norm of gradients: 0.117475867398331
l2 norm of weights: 2.247998944767144
-----
Iteration Number: 5377
Loss: 12.773435147593299
l2 norm of gradients: 0.11743998239828879
l2 norm of weights: 2.247923905957194
-----
Iteration Number: 5378
Loss: 12.77203860475609
l2 norm of gradients: 0.1174041178871767
l2 norm of weights: 2.247848956712689
-----
Iteration Number: 5379
Loss: 12.770643030399835
l2 norm of gradients: 0.11736827385260952
l2 norm of weights: 2.247774096969663
-----
Iteration Number: 5380
Loss: 12.769248423751938
l2 norm of gradients: 0.1173324502822051
l2 norm of weights: 2.2476993266641907
-----
Iteration Number: 5381
Loss: 12.767854784040475
l2 norm of gradients: 0.11729664716358458
l2 norm of weights: 2.247624645732388
-----
Iteration Number: 5382
Loss: 12.766462110494183
l2 norm of gradients: 0.11726086448437238
l2 norm of weights: 2.2475500541104134
-----
Iteration Number: 5383
Loss: 12.765070402342449

```



```

12 norm of gradients: 0.11722510223219612
12 norm of weights: 2.2474755517344667
-----
Iteration Number: 5384
Loss: 12.763679658815331
12 norm of gradients: 0.11718936039468655
12 norm of weights: 2.2474011385407895
-----
Iteration Number: 5385
Loss: 12.762289879143541
12 norm of gradients: 0.11715363895947768
12 norm of weights: 2.247326814465665
-----
Iteration Number: 5386
Loss: 12.76090106255845
12 norm of gradients: 0.11711793791420676
12 norm of weights: 2.247252579445417
-----
Iteration Number: 5387
Loss: 12.759513208292098
12 norm of gradients: 0.1170822572465142
12 norm of weights: 2.247178433416413
-----
Iteration Number: 5388
Loss: 12.758126315577169
12 norm of gradients: 0.11704659694404358
12 norm of weights: 2.2471043763150593
-----
Iteration Number: 5389
Loss: 12.756740383647013
12 norm of gradients: 0.11701095699444172
12 norm of weights: 2.2470304080778067
-----
Iteration Number: 5390
Loss: 12.755355411735632
12 norm of gradients: 0.11697533738535867
12 norm of weights: 2.246956528641145
-----
Iteration Number: 5391
Loss: 12.753971399077692
12 norm of gradients: 0.11693973810444756
12 norm of weights: 2.246882737941607
-----
Iteration Number: 5392
Loss: 12.752588344908508
12 norm of gradients: 0.11690415913936483
12 norm of weights: 2.2468090359157666
-----

```

Iteration Number: 5393
Loss: 12.751206248464058
l2 norm of gradients: 0.11686860047777003
l2 norm of weights: 2.2467354225002385

Iteration Number: 5394
Loss: 12.749825108980973
l2 norm of gradients: 0.11683306210732584
l2 norm of weights: 2.24666189763168

Iteration Number: 5395
Loss: 12.748444925696536
l2 norm of gradients: 0.1167975440156982
l2 norm of weights: 2.2465884612467892

Iteration Number: 5396
Loss: 12.747065697848685
l2 norm of gradients: 0.11676204619055625
l2 norm of weights: 2.246515113282306

Iteration Number: 5397
Loss: 12.745687424676012
l2 norm of gradients: 0.11672656861957216
l2 norm of weights: 2.24644185367501

Iteration Number: 5398
Loss: 12.744310105417771
l2 norm of gradients: 0.11669111129042142
l2 norm of weights: 2.246368682361725

Iteration Number: 5399
Loss: 12.742933739313855
l2 norm of gradients: 0.11665567419078261
l2 norm of weights: 2.246295599279314

Iteration Number: 5400
Loss: 12.741558325604823
l2 norm of gradients: 0.11662025730833746
l2 norm of weights: 2.2462226043646822

Iteration Number: 5401
Loss: 12.740183863531875
l2 norm of gradients: 0.1165848606307709
l2 norm of weights: 2.246149697554776

Iteration Number: 5402
Loss: 12.738810352336866
l2 norm of gradients: 0.11654948414577103

```

12 norm of weights: 2.246076878786583
-----
Iteration Number: 5403
Loss: 12.737437791262305
12 norm of gradients: 0.116514127841029
12 norm of weights: 2.246004147997132
-----
Iteration Number: 5404
Loss: 12.736066179551354
12 norm of gradients: 0.11647879170423925
12 norm of weights: 2.2459315051234934
-----
Iteration Number: 5405
Loss: 12.73469551644782
12 norm of gradients: 0.11644347572309931
12 norm of weights: 2.2458589501027784
-----
Iteration Number: 5406
Loss: 12.733325801196159
12 norm of gradients: 0.11640817988530983
12 norm of weights: 2.2457864828721394
-----
Iteration Number: 5407
Loss: 12.731957033041478
12 norm of gradients: 0.11637290417857461
12 norm of weights: 2.2457141033687713
-----
Iteration Number: 5408
Loss: 12.730589211229534
12 norm of gradients: 0.11633764859060071
12 norm of weights: 2.2456418115299077
-----
Iteration Number: 5409
Loss: 12.729222335006737
12 norm of gradients: 0.11630241310909811
12 norm of weights: 2.2455696072928255
-----
Iteration Number: 5410
Loss: 12.72785640362013
12 norm of gradients: 0.11626719772178015
12 norm of weights: 2.2454974905948424
-----
Iteration Number: 5411
Loss: 12.726491416317415
12 norm of gradients: 0.11623200241636325
12 norm of weights: 2.245425461373316
-----
Iteration Number: 5412

```

```

Loss: 12.725127372346947
12 norm of gradients: 0.11619682718056688
12 norm of weights: 2.2453535195656467
-----
Iteration Number: 5413
Loss: 12.723764270957709
12 norm of gradients: 0.1161616720021137
12 norm of weights: 2.245281665109275
-----
Iteration Number: 5414
Loss: 12.722402111399346
12 norm of gradients: 0.11612653686872952
12 norm of weights: 2.245209897941682
-----
Iteration Number: 5415
Loss: 12.721040892922138
12 norm of gradients: 0.11609142176814331
12 norm of weights: 2.2451382180003914
-----
Iteration Number: 5416
Loss: 12.719680614777019
12 norm of gradients: 0.11605632668808703
12 norm of weights: 2.2450666252229663
-----
Iteration Number: 5417
Loss: 12.718321276215562
12 norm of gradients: 0.11602125161629596
12 norm of weights: 2.2449951195470117
-----
Iteration Number: 5418
Loss: 12.716962876489982
12 norm of gradients: 0.11598619654050837
12 norm of weights: 2.244923700910174
-----
Iteration Number: 5419
Loss: 12.71560541485314
12 norm of gradients: 0.11595116144846576
12 norm of weights: 2.244852369250139
-----
Iteration Number: 5420
Loss: 12.714248890558546
12 norm of gradients: 0.11591614632791264
12 norm of weights: 2.244781124504635
-----
Iteration Number: 5421
Loss: 12.712893302860346
12 norm of gradients: 0.11588115116659674
12 norm of weights: 2.244709966611431

```

```
-----  
Iteration Number: 5422  
Loss: 12.711538651013328  
l2 norm of gradients: 0.11584617595226882  
l2 norm of weights: 2.244638895508337  
-----  
Iteration Number: 5423  
Loss: 12.710184934272926  
l2 norm of gradients: 0.11581122067268292  
l2 norm of weights: 2.244567911133202  
-----  
Iteration Number: 5424  
Loss: 12.708832151895201  
l2 norm of gradients: 0.11577628531559601  
l2 norm of weights: 2.244497013423919  
-----  
Iteration Number: 5425  
Loss: 12.707480303136885  
l2 norm of gradients: 0.11574136986876832  
l2 norm of weights: 2.2444262023184187  
-----  
Iteration Number: 5426  
Loss: 12.706129387255316  
l2 norm of gradients: 0.11570647431996313  
l2 norm of weights: 2.2443554777546755  
-----  
Iteration Number: 5427  
Loss: 12.704779403508496  
l2 norm of gradients: 0.11567159865694689  
l2 norm of weights: 2.244284839670703  
-----  
Iteration Number: 5428  
Loss: 12.703430351155054  
l2 norm of gradients: 0.11563674286748912  
l2 norm of weights: 2.2442142880045557  
-----  
Iteration Number: 5429  
Loss: 12.702082229454264  
l2 norm of gradients: 0.11560190693936252  
l2 norm of weights: 2.2441438226943293  
-----  
Iteration Number: 5430  
Loss: 12.700735037666032  
l2 norm of gradients: 0.1155670908603428  
l2 norm of weights: 2.2440734436781598  
-----  
Iteration Number: 5431  
Loss: 12.699388775050908
```

```

12 norm of gradients: 0.11553229461820887
12 norm of weights: 2.244003150894225
-----
Iteration Number: 5432
Loss: 12.69804344087008
12 norm of gradients: 0.11549751820074279
12 norm of weights: 2.2439329442807416
-----
Iteration Number: 5433
Loss: 12.696699034385361
12 norm of gradients: 0.11546276159572964
12 norm of weights: 2.2438628237759684
-----
Iteration Number: 5434
Loss: 12.695355554859221
12 norm of gradients: 0.11542802479095766
12 norm of weights: 2.243792789318205
-----
Iteration Number: 5435
Loss: 12.69401300155475
12 norm of gradients: 0.11539330777421826
12 norm of weights: 2.243722840845791
-----
Iteration Number: 5436
Loss: 12.692671373735676
12 norm of gradients: 0.11535861053330586
12 norm of weights: 2.243652978297107
-----
Iteration Number: 5437
Loss: 12.691330670666368
12 norm of gradients: 0.1153239330560181
12 norm of weights: 2.2435832016105732
-----
Iteration Number: 5438
Loss: 12.689990891611824
12 norm of gradients: 0.11528927533015561
12 norm of weights: 2.243513510724652
-----
Iteration Number: 5439
Loss: 12.688652035837679
12 norm of gradients: 0.1152546373435223
12 norm of weights: 2.243443905577846
-----
Iteration Number: 5440
Loss: 12.687314102610202
12 norm of gradients: 0.11522001908392505
12 norm of weights: 2.2433743861086977
-----

```

Iteration Number: 5441
Loss: 12.685977091196294
l2 norm of gradients: 0.11518542053917392
l2 norm of weights: 2.24330495225579

Iteration Number: 5442
Loss: 12.684641000863486
l2 norm of gradients: 0.11515084169708212
l2 norm of weights: 2.243235603957748

Iteration Number: 5443
Loss: 12.683305830879945
l2 norm of gradients: 0.11511628254546591
l2 norm of weights: 2.243166341153235

Iteration Number: 5444
Loss: 12.681971580514473
l2 norm of gradients: 0.11508174307214465
l2 norm of weights: 2.2430971637809565

Iteration Number: 5445
Loss: 12.680638249036495
l2 norm of gradients: 0.11504722326494093
l2 norm of weights: 2.2430280717796585

Iteration Number: 5446
Loss: 12.679305835716072
l2 norm of gradients: 0.11501272311168032
l2 norm of weights: 2.2429590650881255

Iteration Number: 5447
Loss: 12.677974339823892
l2 norm of gradients: 0.11497824260019165
l2 norm of weights: 2.2428901436451847

Iteration Number: 5448
Loss: 12.676643760631283
l2 norm of gradients: 0.11494378171830678
l2 norm of weights: 2.2428213073897023

Iteration Number: 5449
Loss: 12.675314097410185
l2 norm of gradients: 0.11490934045386064
l2 norm of weights: 2.242752556260586

Iteration Number: 5450
Loss: 12.673985349433186
l2 norm of gradients: 0.11487491879469142

```
12 norm of weights: 2.242683890196783
-----
Iteration Number: 5451
Loss: 12.672657515973489
12 norm of gradients: 0.11484051672864032
12 norm of weights: 2.242615309137281
-----
Iteration Number: 5452
Loss: 12.671330596304928
12 norm of gradients: 0.1148061342435517
12 norm of weights: 2.242546813021108
-----
Iteration Number: 5453
Loss: 12.670004589701973
12 norm of gradients: 0.11477177132727309
12 norm of weights: 2.2424784017873334
-----
Iteration Number: 5454
Loss: 12.668679495439706
12 norm of gradients: 0.11473742796765499
12 norm of weights: 2.2424100753750644
-----
Iteration Number: 5455
Loss: 12.667355312793845
12 norm of gradients: 0.11470310415255124
12 norm of weights: 2.2423418337234517
-----
Iteration Number: 5456
Loss: 12.666032041040737
12 norm of gradients: 0.11466879986981861
12 norm of weights: 2.2422736767716835
-----
Iteration Number: 5457
Loss: 12.66470967945735
12 norm of gradients: 0.11463451510731715
12 norm of weights: 2.2422056044589898
-----
Iteration Number: 5458
Loss: 12.663388227321272
12 norm of gradients: 0.11460024985290992
12 norm of weights: 2.24213761672464
-----
Iteration Number: 5459
Loss: 12.662067683910731
12 norm of gradients: 0.11456600409446324
12 norm of weights: 2.242069713507944
-----
Iteration Number: 5460
```


Loss: 12.66074804850456
l2 norm of gradients: 0.11453177781984637
l2 norm of weights: 2.2420018947482525

Iteration Number: 5461
Loss: 12.659429320382237
l2 norm of gradients: 0.11449757101693193
l2 norm of weights: 2.2419341603849556

Iteration Number: 5462
Loss: 12.658111498823843
l2 norm of gradients: 0.11446338367359549
l2 norm of weights: 2.2418665103574833

Iteration Number: 5463
Loss: 12.656794583110095
l2 norm of gradients: 0.11442921577771578
l2 norm of weights: 2.2417989446053066

Iteration Number: 5464
Loss: 12.655478572522327
l2 norm of gradients: 0.11439506731717476
l2 norm of weights: 2.2417314630679352

Iteration Number: 5465
Loss: 12.654163466342496
l2 norm of gradients: 0.11436093827985751
l2 norm of weights: 2.2416640656849207

Iteration Number: 5466
Loss: 12.652849263853186
l2 norm of gradients: 0.11432682865365214
l2 norm of weights: 2.241596752395854

Iteration Number: 5467
Loss: 12.651535964337592
l2 norm of gradients: 0.11429273842645003
l2 norm of weights: 2.2415295231403647

Iteration Number: 5468
Loss: 12.650223567079536
l2 norm of gradients: 0.11425866758614568
l2 norm of weights: 2.2414623778581246

Iteration Number: 5469
Loss: 12.648912071363458
l2 norm of gradients: 0.11422461612063663
l2 norm of weights: 2.2413953164888443

```
-----  
Iteration Number: 5470  
Loss: 12.647601476474419  
l2 norm of gradients: 0.11419058401782367  
l2 norm of weights: 2.241328338972274  
-----  
Iteration Number: 5471  
Loss: 12.646291781698098  
l2 norm of gradients: 0.11415657126561068  
l2 norm of weights: 2.241261445248205  
-----  
Iteration Number: 5472  
Loss: 12.644982986320795  
l2 norm of gradients: 0.11412257785190481  
l2 norm of weights: 2.241194635256468  
-----  
Iteration Number: 5473  
Loss: 12.643675089629419  
l2 norm of gradients: 0.11408860376461619  
l2 norm of weights: 2.241127908936933  
-----  
Iteration Number: 5474  
Loss: 12.642368090911516  
l2 norm of gradients: 0.1140546489916582  
l2 norm of weights: 2.241061266229511  
-----  
Iteration Number: 5475  
Loss: 12.641061989455228  
l2 norm of gradients: 0.11402071352094739  
l2 norm of weights: 2.240994707074152  
-----  
Iteration Number: 5476  
Loss: 12.639756784549329  
l2 norm of gradients: 0.11398679734040348  
l2 norm of weights: 2.240928231410846  
-----  
Iteration Number: 5477  
Loss: 12.638452475483199  
l2 norm of gradients: 0.11395290043794926  
l2 norm of weights: 2.2408618391796233  
-----  
Iteration Number: 5478  
Loss: 12.637149061546843  
l2 norm of gradients: 0.1139190228015107  
l2 norm of weights: 2.2407955303205536  
-----  
Iteration Number: 5479  
Loss: 12.635846542030873
```

```
12 norm of gradients: 0.11388516441901708
12 norm of weights: 2.240729304773746
-----
Iteration Number: 5480
Loss: 12.634544916226522
12 norm of gradients: 0.11385132527840065
12 norm of weights: 2.2406631624793514
-----
Iteration Number: 5481
Loss: 12.633244183425633
12 norm of gradients: 0.113817505367597
12 norm of weights: 2.240597103377557
-----
Iteration Number: 5482
Loss: 12.631944342920669
12 norm of gradients: 0.11378370467454471
12 norm of weights: 2.240531127408593
-----
Iteration Number: 5483
Loss: 12.630645394004704
12 norm of gradients: 0.11374992318718578
12 norm of weights: 2.2404652345127274
-----
Iteration Number: 5484
Loss: 12.629347335971417
12 norm of gradients: 0.1137161608934651
12 norm of weights: 2.2403994246302688
-----
Iteration Number: 5485
Loss: 12.628050168115113
12 norm of gradients: 0.11368241778133105
12 norm of weights: 2.240333697701564
-----
Iteration Number: 5486
Loss: 12.626753889730697
12 norm of gradients: 0.11364869383873488
12 norm of weights: 2.2402680536670023
-----
Iteration Number: 5487
Loss: 12.625458500113698
12 norm of gradients: 0.11361498905363132
12 norm of weights: 2.24020249246701
-----
Iteration Number: 5488
Loss: 12.624163998560249
12 norm of gradients: 0.11358130341397804
12 norm of weights: 2.2401370140420536
-----
```

Iteration Number: 5489
Loss: 12.622870384367088
l2 norm of gradients: 0.11354763690773602
l2 norm of weights: 2.2400716183326397

Iteration Number: 5490
Loss: 12.62157765683158
l2 norm of gradients: 0.11351398952286947
l2 norm of weights: 2.2400063052793144

Iteration Number: 5491
Loss: 12.62028581525168
l2 norm of gradients: 0.11348036124734576
l2 norm of weights: 2.239941074822663

Iteration Number: 5492
Loss: 12.61899485892597
l2 norm of gradients: 0.11344675206913539
l2 norm of weights: 2.239875926903311

Iteration Number: 5493
Loss: 12.617704787153626
l2 norm of gradients: 0.11341316197621218
l2 norm of weights: 2.2398108614619225

Iteration Number: 5494
Loss: 12.616415599234442
l2 norm of gradients: 0.1133795909565531
l2 norm of weights: 2.2397458784392015

Iteration Number: 5495
Loss: 12.61512729446882
l2 norm of gradients: 0.11334603899813829
l2 norm of weights: 2.2396809777758917

Iteration Number: 5496
Loss: 12.613839872157765
l2 norm of gradients: 0.11331250608895117
l2 norm of weights: 2.2396161594127757

Iteration Number: 5497
Loss: 12.612553331602886
l2 norm of gradients: 0.11327899221697833
l2 norm of weights: 2.239551423290676

Iteration Number: 5498
Loss: 12.611267672106413
l2 norm of gradients: 0.11324549737020968

```

12 norm of weights: 2.2394867693504543
-----
Iteration Number: 5499
Loss: 12.60998289297117
12 norm of gradients: 0.11321202153663817
12 norm of weights: 2.2394221975330124
-----
Iteration Number: 5500
Loss: 12.608698993500585
12 norm of gradients: 0.11317856470426013
12 norm of weights: 2.2393577077792894
-----
Iteration Number: 5501
Loss: 12.607415972998698
12 norm of gradients: 0.11314512686107503
12 norm of weights: 2.2392933000302664
-----
Iteration Number: 5502
Loss: 12.606133830770151
12 norm of gradients: 0.11311170799508569
12 norm of weights: 2.239228974226962
-----
Iteration Number: 5503
Loss: 12.60485256612019
12 norm of gradients: 0.11307830809429796
12 norm of weights: 2.239164730310434
-----
Iteration Number: 5504
Loss: 12.60357217835467
12 norm of gradients: 0.11304492714672121
12 norm of weights: 2.2391005682217817
-----
Iteration Number: 5505
Loss: 12.602292666780036
12 norm of gradients: 0.11301156514036773
12 norm of weights: 2.2390364879021405
-----
Iteration Number: 5506
Loss: 12.60101403070335
12 norm of gradients: 0.11297822206325335
12 norm of weights: 2.2389724892926877
-----
Iteration Number: 5507
Loss: 12.599736269432269
12 norm of gradients: 0.11294489790339696
12 norm of weights: 2.238908572334638
-----
Iteration Number: 5508

```

Loss: 12.598459382275056
l2 norm of gradients: 0.11291159264882077
l2 norm of weights: 2.2388447369692464

Iteration Number: 5509
Loss: 12.597183368540572
l2 norm of gradients: 0.11287830628755029
l2 norm of weights: 2.2387809831378065

Iteration Number: 5510
Loss: 12.595908227538281
l2 norm of gradients: 0.11284503880761419
l2 norm of weights: 2.2387173107816514

Iteration Number: 5511
Loss: 12.594633958578244
l2 norm of gradients: 0.11281179019704451
l2 norm of weights: 2.238653719842153

Iteration Number: 5512
Loss: 12.59336056097113
l2 norm of gradients: 0.11277856044387652
l2 norm of weights: 2.2385902102607225

Iteration Number: 5513
Loss: 12.5920880340282
l2 norm of gradients: 0.11274534953614873
l2 norm of weights: 2.2385267819788104

Iteration Number: 5514
Loss: 12.590816377061318
l2 norm of gradients: 0.11271215746190291
l2 norm of weights: 2.238463434937906

Iteration Number: 5515
Loss: 12.58954558938294
l2 norm of gradients: 0.11267898420918424
l2 norm of weights: 2.238400169079537

Iteration Number: 5516
Loss: 12.588275670306132
l2 norm of gradients: 0.11264582976604104
l2 norm of weights: 2.2383369843452723

Iteration Number: 5517
Loss: 12.58700661914455
l2 norm of gradients: 0.11261269412052496
l2 norm of weights: 2.238273880676717

```
-----  
Iteration Number: 5518  
Loss: 12.585738435212448  
l2 norm of gradients: 0.11257957726069098  
l2 norm of weights: 2.238210858015517  
-----  
Iteration Number: 5519  
Loss: 12.584471117824677  
l2 norm of gradients: 0.11254647917459742  
l2 norm of weights: 2.2381479163033564  
-----  
Iteration Number: 5520  
Loss: 12.583204666296687  
l2 norm of gradients: 0.11251339985030574  
l2 norm of weights: 2.238085055481959  
-----  
Iteration Number: 5521  
Loss: 12.58193907994452  
l2 norm of gradients: 0.11248033927588083  
l2 norm of weights: 2.2380222754930865  
-----  
Iteration Number: 5522  
Loss: 12.580674358084815  
l2 norm of gradients: 0.11244729743939089  
l2 norm of weights: 2.2379595762785405  
-----  
Iteration Number: 5523  
Loss: 12.57941050003481  
l2 norm of gradients: 0.11241427432890744  
l2 norm of weights: 2.2378969577801606  
-----  
Iteration Number: 5524  
Loss: 12.57814750511233  
l2 norm of gradients: 0.11238126993250518  
l2 norm of weights: 2.2378344199398263  
-----  
Iteration Number: 5525  
Loss: 12.576885372635799  
l2 norm of gradients: 0.11234828423826232  
l2 norm of weights: 2.2377719626994548  
-----  
Iteration Number: 5526  
Loss: 12.575624101924234  
l2 norm of gradients: 0.11231531723426032  
l2 norm of weights: 2.2377095860010026  
-----  
Iteration Number: 5527  
Loss: 12.574363692297245
```

```

12 norm of gradients: 0.1122823689085839
12 norm of weights: 2.2376472897864654
-----
Iteration Number: 5528
Loss: 12.573104143075035
12 norm of gradients: 0.11224943924932125
12 norm of weights: 2.237585073997877
-----
Iteration Number: 5529
Loss: 12.5718454535784
12 norm of gradients: 0.1122165282445638
12 norm of weights: 2.2375229385773108
-----
Iteration Number: 5530
Loss: 12.57058762312872
12 norm of gradients: 0.11218363588240635
12 norm of weights: 2.2374608834668774
-----
Iteration Number: 5531
Loss: 12.569330651047979
12 norm of gradients: 0.11215076215094709
12 norm of weights: 2.237398908608728
-----
Iteration Number: 5532
Loss: 12.568074536658747
12 norm of gradients: 0.11211790703828754
12 norm of weights: 2.237337013945051
-----
Iteration Number: 5533
Loss: 12.566819279284175
12 norm of gradients: 0.11208507053253254
12 norm of weights: 2.2372751994180744
-----
Iteration Number: 5534
Loss: 12.565564878248018
12 norm of gradients: 0.11205225262179035
12 norm of weights: 2.2372134649700643
-----
Iteration Number: 5535
Loss: 12.564311332874615
12 norm of gradients: 0.11201945329417255
12 norm of weights: 2.2371518105433266
-----
Iteration Number: 5536
Loss: 12.563058642488889
12 norm of gradients: 0.11198667253779417
12 norm of weights: 2.237090236080203
-----

```


Iteration Number: 5537
Loss: 12.561806806416358
l2 norm of gradients: 0.11195391034077347
l2 norm of weights: 2.2370287415230776

Iteration Number: 5538
Loss: 12.560555823983126
l2 norm of gradients: 0.11192116669123231
l2 norm of weights: 2.2369673268143697

Iteration Number: 5539
Loss: 12.559305694515885
l2 norm of gradients: 0.11188844157729573
l2 norm of weights: 2.2369059918965393

Iteration Number: 5540
Loss: 12.558056417341916
l2 norm of gradients: 0.1118557349870923
l2 norm of weights: 2.236844736712084

Iteration Number: 5541
Loss: 12.556807991789077
l2 norm of gradients: 0.11182304690875385
l2 norm of weights: 2.236783561203539

Iteration Number: 5542
Loss: 12.555560417185827
l2 norm of gradients: 0.11179037733041583
l2 norm of weights: 2.236722465313481

Iteration Number: 5543
Loss: 12.554313692861198
l2 norm of gradients: 0.11175772624021688
l2 norm of weights: 2.236661448984522

Iteration Number: 5544
Loss: 12.55306781814482
l2 norm of gradients: 0.11172509362629916
l2 norm of weights: 2.2366005121593138

Iteration Number: 5545
Loss: 12.551822792366892
l2 norm of gradients: 0.11169247947680819
l2 norm of weights: 2.2365396547805463

Iteration Number: 5546
Loss: 12.55057861485821
l2 norm of gradients: 0.11165988377989305

```

12 norm of weights: 2.236478876790948
-----
Iteration Number: 5547
Loss: 12.54933528495015
12 norm of gradients: 0.11162730652370609
12 norm of weights: 2.236418178133286
-----
Iteration Number: 5548
Loss: 12.548092801974668
12 norm of gradients: 0.11159474769640314
12 norm of weights: 2.236357558750365
-----
Iteration Number: 5549
Loss: 12.54685116526431
12 norm of gradients: 0.1115622072861435
12 norm of weights: 2.2362970185850295
-----
Iteration Number: 5550
Loss: 12.545610374152202
12 norm of gradients: 0.11152968528108988
12 norm of weights: 2.2362365575801597
-----
Iteration Number: 5551
Loss: 12.544370427972048
12 norm of gradients: 0.11149718166940846
12 norm of weights: 2.2361761756786773
-----
Iteration Number: 5552
Loss: 12.543131326058136
12 norm of gradients: 0.11146469643926887
12 norm of weights: 2.23611587282354
-----
Iteration Number: 5553
Loss: 12.541893067745338
12 norm of gradients: 0.11143222957884423
12 norm of weights: 2.2360556489577434
-----
Iteration Number: 5554
Loss: 12.540655652369106
12 norm of gradients: 0.11139978107631104
12 norm of weights: 2.2359955040243245
-----
Iteration Number: 5555
Loss: 12.539419079265464
12 norm of gradients: 0.11136735091984934
12 norm of weights: 2.2359354379663543
-----
Iteration Number: 5556

```

Loss: 12.538183347771026
l2 norm of gradients: 0.11133493909764267
l2 norm of weights: 2.2358754507269447

Iteration Number: 5557
Loss: 12.536948457222984
l2 norm of gradients: 0.11130254559787792
l2 norm of weights: 2.2358155422492456

Iteration Number: 5558
Loss: 12.535714406959105
l2 norm of gradients: 0.11127017040874564
l2 norm of weights: 2.235755712476444

Iteration Number: 5559
Loss: 12.534481196317737
l2 norm of gradients: 0.11123781351843975
l2 norm of weights: 2.235695961351765

Iteration Number: 5560
Loss: 12.5332488246378
l2 norm of gradients: 0.11120547491515775
l2 norm of weights: 2.2356362888184735

Iteration Number: 5561
Loss: 12.532017291258803
l2 norm of gradients: 0.11117315458710057
l2 norm of weights: 2.2355766948198705

Iteration Number: 5562
Loss: 12.530786595520823
l2 norm of gradients: 0.11114085252247263
l2 norm of weights: 2.235517179299296

Iteration Number: 5563
Loss: 12.529556736764516
l2 norm of gradients: 0.11110856870948195
l2 norm of weights: 2.235457742200128

Iteration Number: 5564
Loss: 12.528327714331112
l2 norm of gradients: 0.11107630313634004
l2 norm of weights: 2.235398383465782

Iteration Number: 5565
Loss: 12.527099527562424
l2 norm of gradients: 0.11104405579126196
l2 norm of weights: 2.2353391030397125

```

-----
Iteration Number: 5566
Loss: 12.52587217580083
l2 norm of gradients: 0.1110118266624662
l2 norm of weights: 2.2352799008654105
-----
Iteration Number: 5567
Loss: 12.524645658389288
l2 norm of gradients: 0.11097961573817494
l2 norm of weights: 2.2352207768864063
-----
Iteration Number: 5568
Loss: 12.523419974671334
l2 norm of gradients: 0.11094742300661374
l2 norm of weights: 2.2351617310462673
-----
Iteration Number: 5569
Loss: 12.522195123991066
l2 norm of gradients: 0.11091524845601183
l2 norm of weights: 2.2351027632885994
-----
Iteration Number: 5570
Loss: 12.520971105693173
l2 norm of gradients: 0.11088309207460197
l2 norm of weights: 2.2350438735570464
-----
Iteration Number: 5571
Loss: 12.519747919122903
l2 norm of gradients: 0.11085095385062042
l2 norm of weights: 2.234985061795289
-----
Iteration Number: 5572
Loss: 12.51852556362608
l2 norm of gradients: 0.11081883377230708
l2 norm of weights: 2.234926327947046
-----
Iteration Number: 5573
Loss: 12.517304038549101
l2 norm of gradients: 0.11078673182790545
l2 norm of weights: 2.234867671956075
-----
Iteration Number: 5574
Loss: 12.516083343238932
l2 norm of gradients: 0.1107546480056624
l2 norm of weights: 2.2348090937661715
-----
Iteration Number: 5575
Loss: 12.514863477043116

```

```

12 norm of gradients: 0.11072258229382866
12 norm of weights: 2.2347505933211664
-----
Iteration Number: 5576
Loss: 12.51364443930976
12 norm of gradients: 0.11069053468065841
12 norm of weights: 2.2346921705649314
-----
Iteration Number: 5577
Loss: 12.512426229387543
12 norm of gradients: 0.1106585051544094
12 norm of weights: 2.234633825441374
-----
Iteration Number: 5578
Loss: 12.511208846625717
12 norm of gradients: 0.1106264937033431
12 norm of weights: 2.2345755578944395
-----
Iteration Number: 5579
Loss: 12.5099922903741
12 norm of gradients: 0.1105945003157244
12 norm of weights: 2.234517367868112
-----
Iteration Number: 5580
Loss: 12.50877655998308
12 norm of gradients: 0.11056252497982198
12 norm of weights: 2.2344592553064118
-----
Iteration Number: 5581
Loss: 12.507561654803611
12 norm of gradients: 0.1105305676839081
12 norm of weights: 2.2344012201533987
-----
Iteration Number: 5582
Loss: 12.506347574187222
12 norm of gradients: 0.11049862841625857
12 norm of weights: 2.234343262353168
-----
Iteration Number: 5583
Loss: 12.505134317485997
12 norm of gradients: 0.11046670716515293
12 norm of weights: 2.2342853818498543
-----
Iteration Number: 5584
Loss: 12.503921884052598
12 norm of gradients: 0.1104348039188743
12 norm of weights: 2.234227578587629
-----

```

Iteration Number: 5585
Loss: 12.502710273240249
l2 norm of gradients: 0.11040291866570945
l2 norm of weights: 2.2341698525107008

Iteration Number: 5586
Loss: 12.501499484402741
l2 norm of gradients: 0.1103710513939488
l2 norm of weights: 2.234112203563317

Iteration Number: 5587
Loss: 12.500289516894432
l2 norm of gradients: 0.11033920209188644
l2 norm of weights: 2.234054631689761

Iteration Number: 5588
Loss: 12.499080370070237
l2 norm of gradients: 0.11030737074782015
l2 norm of weights: 2.233997136834355

Iteration Number: 5589
Loss: 12.49787204328565
l2 norm of gradients: 0.1102755573500513
l2 norm of weights: 2.233939718941457

Iteration Number: 5590
Loss: 12.496664535896716
l2 norm of gradients: 0.11024376188688502
l2 norm of weights: 2.233882377955465

Iteration Number: 5591
Loss: 12.495457847260052
l2 norm of gradients: 0.11021198434663008
l2 norm of weights: 2.2338251138208123

Iteration Number: 5592
Loss: 12.494251976732833
l2 norm of gradients: 0.11018022471759897
l2 norm of weights: 2.2337679264819696

Iteration Number: 5593
Loss: 12.493046923672804
l2 norm of gradients: 0.11014848298810782
l2 norm of weights: 2.2337108158834464

Iteration Number: 5594
Loss: 12.491842687438258
l2 norm of gradients: 0.11011675914647652

```

12 norm of weights: 2.2336537819697893
-----
Iteration Number: 5595
Loss: 12.490639267388067
12 norm of gradients: 0.11008505318102864
12 norm of weights: 2.2335968246855806
-----
Iteration Number: 5596
Loss: 12.489436662881653
12 norm of gradients: 0.11005336508009139
12 norm of weights: 2.233539943975441
-----
Iteration Number: 5597
Loss: 12.488234873279008
12 norm of gradients: 0.1100216948319959
12 norm of weights: 2.23348313978403
-----
Iteration Number: 5598
Loss: 12.487033897940677
12 norm of gradients: 0.1099900424250768
12 norm of weights: 2.233426412056041
-----
Iteration Number: 5599
Loss: 12.485833736227764
12 norm of gradients: 0.1099584078476726
12 norm of weights: 2.233369760736208
-----
Iteration Number: 5600
Loss: 12.484634387501943
12 norm of gradients: 0.10992679108812556
12 norm of weights: 2.2333131857693003
-----
Iteration Number: 5601
Loss: 12.483435851125432
12 norm of gradients: 0.10989519213478154
12 norm of weights: 2.2332566871001247
-----
Iteration Number: 5602
Loss: 12.482238126461027
12 norm of gradients: 0.10986361097599029
12 norm of weights: 2.233200264673526
-----
Iteration Number: 5603
Loss: 12.48104121287206
12 norm of gradients: 0.10983204760010529
12 norm of weights: 2.233143918434384
-----
Iteration Number: 5604

```

```

Loss: 12.47984510972244
12 norm of gradients: 0.10980050199548375
12 norm of weights: 2.23308764832762
-----
Iteration Number: 5605
Loss: 12.478649816376622
12 norm of gradients: 0.10976897415048664
12 norm of weights: 2.2330314542981866
-----
Iteration Number: 5606
Loss: 12.477455332199629
12 norm of gradients: 0.10973746405347878
12 norm of weights: 2.2329753362910787
-----
Iteration Number: 5607
Loss: 12.476261656557021
12 norm of gradients: 0.10970597169282875
12 norm of weights: 2.2329192942513245
-----
Iteration Number: 5608
Loss: 12.475068788814934
12 norm of gradients: 0.10967449705690893
12 norm of weights: 2.232863328123992
-----
Iteration Number: 5609
Loss: 12.473876728340045
12 norm of gradients: 0.10964304013409543
12 norm of weights: 2.2328074378541847
-----
Iteration Number: 5610
Loss: 12.472685474499603
12 norm of gradients: 0.1096116009127683
12 norm of weights: 2.2327516233870437
-----
Iteration Number: 5611
Loss: 12.471495026661394
12 norm of gradients: 0.10958017938131123
12 norm of weights: 2.2326958846677463
-----
Iteration Number: 5612
Loss: 12.470305384193768
12 norm of gradients: 0.10954877552811186
12 norm of weights: 2.2326402216415078
-----
Iteration Number: 5613
Loss: 12.469116546465626
12 norm of gradients: 0.1095173893415616
12 norm of weights: 2.2325846342535796

```



```

-----
Iteration Number: 5614
Loss: 12.467928512846424
l2 norm of gradients: 0.10948602081005573
l2 norm of weights: 2.2325291224492507
-----
Iteration Number: 5615
Loss: 12.466741282706163
l2 norm of gradients: 0.10945466992199335
l2 norm of weights: 2.2324736861738472
-----
Iteration Number: 5616
Loss: 12.465554855415409
l2 norm of gradients: 0.10942333666577739
l2 norm of weights: 2.2324183253727305
-----
Iteration Number: 5617
Loss: 12.464369230345273
l2 norm of gradients: 0.10939202102981462
l2 norm of weights: 2.232363039991301
-----
Iteration Number: 5618
Loss: 12.463184406867418
l2 norm of gradients: 0.10936072300251576
l2 norm of weights: 2.2323078299749937
-----
Iteration Number: 5619
Loss: 12.462000384354056
l2 norm of gradients: 0.10932944257229528
l2 norm of weights: 2.232252695269283
-----
Iteration Number: 5620
Loss: 12.460817162177953
l2 norm of gradients: 0.10929817972757162
l2 norm of weights: 2.232197635819678
-----
Iteration Number: 5621
Loss: 12.459634739712424
l2 norm of gradients: 0.10926693445676701
l2 norm of weights: 2.232142651571725
-----
Iteration Number: 5622
Loss: 12.458453116331334
l2 norm of gradients: 0.10923570674830763
l2 norm of weights: 2.2320877424710077
-----
Iteration Number: 5623
Loss: 12.457272291409096

```

```

12 norm of gradients: 0.10920449659062353
12 norm of weights: 2.2320329084631463
-----
Iteration Number: 5624
Loss: 12.456092264320672
12 norm of gradients: 0.1091733039721487
12 norm of weights: 2.231978149493797
-----
Iteration Number: 5625
Loss: 12.454913034441574
12 norm of gradients: 0.10914212888132101
12 norm of weights: 2.2319234655086535
-----
Iteration Number: 5626
Loss: 12.453734601147858
12 norm of gradients: 0.10911097130658219
12 norm of weights: 2.2318688564534463
-----
Iteration Number: 5627
Loss: 12.452556963816134
12 norm of gradients: 0.10907983123637796
12 norm of weights: 2.2318143222739417
-----
Iteration Number: 5628
Loss: 12.451380121823552
12 norm of gradients: 0.10904870865915801
12 norm of weights: 2.231759862915943
-----
Iteration Number: 5629
Loss: 12.450204074547813
12 norm of gradients: 0.10901760356337582
12 norm of weights: 2.23170547832529
-----
Iteration Number: 5630
Loss: 12.449028821367163
12 norm of gradients: 0.10898651593748897
12 norm of weights: 2.23165116844786
-----
Iteration Number: 5631
Loss: 12.447854361660385
12 norm of gradients: 0.10895544576995884
12 norm of weights: 2.231596933229565
-----
Iteration Number: 5632
Loss: 12.446680694806828
12 norm of gradients: 0.10892439304925093
12 norm of weights: 2.2315427726163555
-----

```

```

Iteration Number: 5633
Loss: 12.445507820186364
l2 norm of gradients: 0.10889335776383456
l2 norm of weights: 2.231488686554217
-----
Iteration Number: 5634
Loss: 12.44433573717942
l2 norm of gradients: 0.10886233990218308
l2 norm of weights: 2.2314346749891727
-----
Iteration Number: 5635
Loss: 12.443164445166964
l2 norm of gradients: 0.1088313394527738
l2 norm of weights: 2.231380737867281
-----
Iteration Number: 5636
Loss: 12.44199394353051
l2 norm of gradients: 0.10880035640408812
l2 norm of weights: 2.231326875134638
-----
Iteration Number: 5637
Loss: 12.440824231652115
l2 norm of gradients: 0.10876939074461127
l2 norm of weights: 2.231273086737375
-----
Iteration Number: 5638
Loss: 12.439655308914368
l2 norm of gradients: 0.10873844246283249
l2 norm of weights: 2.231219372621661
-----
Iteration Number: 5639
Loss: 12.43848717470042
l2 norm of gradients: 0.10870751154724519
l2 norm of weights: 2.2311657327337002
-----
Iteration Number: 5640
Loss: 12.43731982839394
l2 norm of gradients: 0.10867659798634663
l2 norm of weights: 2.2311121670197336
-----
Iteration Number: 5641
Loss: 12.436153269379158
l2 norm of gradients: 0.10864570176863818
l2 norm of weights: 2.231058675426039
-----
Iteration Number: 5642
Loss: 12.434987497040835
l2 norm of gradients: 0.10861482288262515

```

```
12 norm of weights: 2.2310052578989295
-----
Iteration Number: 5643
Loss: 12.433822510764273
12 norm of gradients: 0.10858396131681695
12 norm of weights: 2.2309519143847556
-----
Iteration Number: 5644
Loss: 12.432658309935311
12 norm of gradients: 0.10855311705972706
12 norm of weights: 2.2308986448299035
-----
Iteration Number: 5645
Loss: 12.431494893940334
12 norm of gradients: 0.10852229009987291
12 norm of weights: 2.2308454491807956
-----
Iteration Number: 5646
Loss: 12.430332262166264
12 norm of gradients: 0.10849148042577604
12 norm of weights: 2.2307923273838908
-----
Iteration Number: 5647
Loss: 12.429170414000556
12 norm of gradients: 0.10846068802596211
12 norm of weights: 2.2307392793856833
-----
Iteration Number: 5648
Loss: 12.428009348831209
12 norm of gradients: 0.10842991288896074
12 norm of weights: 2.230686305132705
-----
Iteration Number: 5649
Loss: 12.426849066046758
12 norm of gradients: 0.1083991550033057
12 norm of weights: 2.2306334045715226
-----
Iteration Number: 5650
Loss: 12.425689565036272
12 norm of gradients: 0.10836841435753479
12 norm of weights: 2.2305805776487397
-----
Iteration Number: 5651
Loss: 12.42453084518936
12 norm of gradients: 0.10833769094018995
12 norm of weights: 2.230527824310996
-----
Iteration Number: 5652
```

Loss: 12.423372905896162
l2 norm of gradients: 0.10830698473981716
l2 norm of weights: 2.2304751445049664

Iteration Number: 5653
Loss: 12.422215746547366
l2 norm of gradients: 0.1082762957449666
l2 norm of weights: 2.2304225381773635

Iteration Number: 5654
Loss: 12.421059366534179
l2 norm of gradients: 0.1082456239441925
l2 norm of weights: 2.230370005274934

Iteration Number: 5655
Loss: 12.419903765248353
l2 norm of gradients: 0.10821496932605318
l2 norm of weights: 2.230317545744463

Iteration Number: 5656
Loss: 12.418748942082171
l2 norm of gradients: 0.10818433187911113
l2 norm of weights: 2.2302651595327685

Iteration Number: 5657
Loss: 12.417594896428453
l2 norm of gradients: 0.10815371159193299
l2 norm of weights: 2.2302128465867073

Iteration Number: 5658
Loss: 12.416441627680546
l2 norm of gradients: 0.10812310845308948
l2 norm of weights: 2.2301606068531714

Iteration Number: 5659
Loss: 12.415289135232335
l2 norm of gradients: 0.10809252245115551
l2 norm of weights: 2.230108440279087

Iteration Number: 5660
Loss: 12.414137418478242
l2 norm of gradients: 0.10806195357471013
l2 norm of weights: 2.230056346811419

Iteration Number: 5661
Loss: 12.412986476813206
l2 norm of gradients: 0.10803140181233657
l2 norm of weights: 2.2300043263971663

```

-----
Iteration Number: 5662
Loss: 12.411836309632713
l2 norm of gradients: 0.10800086715262228
l2 norm of weights: 2.2299523789833633
-----
Iteration Number: 5663
Loss: 12.410686916332775
l2 norm of gradients: 0.10797034958415873
l2 norm of weights: 2.2299005045170825
-----
Iteration Number: 5664
Loss: 12.409538296309933
l2 norm of gradients: 0.10793984909554172
l2 norm of weights: 2.229848702945431
-----
Iteration Number: 5665
Loss: 12.408390448961256
l2 norm of gradients: 0.1079093656753712
l2 norm of weights: 2.22979697421555
-----
Iteration Number: 5666
Loss: 12.407243373684354
l2 norm of gradients: 0.10787889931225128
l2 norm of weights: 2.229745318274619
-----
Iteration Number: 5667
Loss: 12.406097069877347
l2 norm of gradients: 0.10784844999479033
l2 norm of weights: 2.229693735069852
-----
Iteration Number: 5668
Loss: 12.404951536938905
l2 norm of gradients: 0.10781801771160089
l2 norm of weights: 2.229642224548499
-----
Iteration Number: 5669
Loss: 12.403806774268215
l2 norm of gradients: 0.10778760245129979
l2 norm of weights: 2.2295907866578464
-----
Iteration Number: 5670
Loss: 12.402662781264986
l2 norm of gradients: 0.107757204202508
l2 norm of weights: 2.229539421345215
-----
Iteration Number: 5671
Loss: 12.401519557329475

```

```

12 norm of gradients: 0.10772682295385075
12 norm of weights: 2.2294881285579624
-----
Iteration Number: 5672
Loss: 12.400377101862443
12 norm of gradients: 0.10769645869395755
12 norm of weights: 2.229436908243481
-----
Iteration Number: 5673
Loss: 12.399235414265194
12 norm of gradients: 0.10766611141146211
12 norm of weights: 2.2293857603491993
-----
Iteration Number: 5674
Loss: 12.398094493939553
12 norm of gradients: 0.10763578109500249
12 norm of weights: 2.229334684822581
-----
Iteration Number: 5675
Loss: 12.396954340287863
12 norm of gradients: 0.10760546773322086
12 norm of weights: 2.229283681611126
-----
Iteration Number: 5676
Loss: 12.395814952713012
12 norm of gradients: 0.1075751713147638
12 norm of weights: 2.22923275066237
-----
Iteration Number: 5677
Loss: 12.394676330618388
12 norm of gradients: 0.10754489182828209
12 norm of weights: 2.2291818919238824
-----
Iteration Number: 5678
Loss: 12.393538473407922
12 norm of gradients: 0.10751462926243081
12 norm of weights: 2.2291311053432707
-----
Iteration Number: 5679
Loss: 12.392401380486065
12 norm of gradients: 0.10748438360586936
12 norm of weights: 2.2290803908681758
-----
Iteration Number: 5680
Loss: 12.391265051257788
12 norm of gradients: 0.1074541548472615
12 norm of weights: 2.229029748446275
-----

```

Iteration Number: 5681
Loss: 12.39012948512858
l2 norm of gradients: 0.10742394297527512
l2 norm of weights: 2.228979178025281

Iteration Number: 5682
Loss: 12.388994681504471
l2 norm of gradients: 0.10739374797858257
l2 norm of weights: 2.2289286795529417

Iteration Number: 5683
Loss: 12.38786063979199
l2 norm of gradients: 0.10736356984586046
l2 norm of weights: 2.228878252977041

Iteration Number: 5684
Loss: 12.386727359398208
l2 norm of gradients: 0.10733340856578978
l2 norm of weights: 2.228827898245397

Iteration Number: 5685
Loss: 12.385594839730707
l2 norm of gradients: 0.10730326412705579
l2 norm of weights: 2.2287776153058654

Iteration Number: 5686
Loss: 12.384463080197586
l2 norm of gradients: 0.10727313651834816
l2 norm of weights: 2.228727404106334

Iteration Number: 5687
Loss: 12.383332080207476
l2 norm of gradients: 0.1072430257283609
l2 norm of weights: 2.2286772645947286

Iteration Number: 5688
Loss: 12.382201839169518
l2 norm of gradients: 0.10721293174579229
l2 norm of weights: 2.228627196719009

Iteration Number: 5689
Loss: 12.381072356493378
l2 norm of gradients: 0.10718285455934515
l2 norm of weights: 2.2285772004271713

Iteration Number: 5690
Loss: 12.379943631589242
l2 norm of gradients: 0.10715279415772648


```
12 norm of weights: 2.228527275667245
-----
Iteration Number: 5691
Loss: 12.378815663867806
12 norm of gradients: 0.10712275052964783
12 norm of weights: 2.228477422387298
-----
Iteration Number: 5692
Loss: 12.3776884527403
12 norm of gradients: 0.10709272366382501
12 norm of weights: 2.2284276405354295
-----
Iteration Number: 5693
Loss: 12.376561997618452
12 norm of gradients: 0.10706271354897824
12 norm of weights: 2.2283779300597764
-----
Iteration Number: 5694
Loss: 12.375436297914524
12 norm of gradients: 0.10703272017383227
12 norm of weights: 2.228328290908511
-----
Iteration Number: 5695
Loss: 12.374311353041287
12 norm of gradients: 0.10700274352711608
12 norm of weights: 2.2282787230298386
-----
Iteration Number: 5696
Loss: 12.373187162412028
12 norm of gradients: 0.10697278359756321
12 norm of weights: 2.2282292263720023
-----
Iteration Number: 5697
Loss: 12.372063725440553
12 norm of gradients: 0.10694284037391157
12 norm of weights: 2.228179800883278
-----
Iteration Number: 5698
Loss: 12.370941041541183
12 norm of gradients: 0.10691291384490352
12 norm of weights: 2.2281304465119782
-----
Iteration Number: 5699
Loss: 12.369819110128754
12 norm of gradients: 0.10688300399928573
12 norm of weights: 2.2280811632064492
-----
Iteration Number: 5700
```

Loss: 12.368697930618616
l2 norm of gradients: 0.10685311082580956
l2 norm of weights: 2.2280319509150743

Iteration Number: 5701
Loss: 12.36757750242663
l2 norm of gradients: 0.10682323431323065
l2 norm of weights: 2.2279828095862695

Iteration Number: 5702
Loss: 12.366457824969178
l2 norm of gradients: 0.10679337445030915
l2 norm of weights: 2.2279337391684866

Iteration Number: 5703
Loss: 12.36533889766315
l2 norm of gradients: 0.10676353122580963
l2 norm of weights: 2.2278847396102135

Iteration Number: 5704
Loss: 12.36422071992595
l2 norm of gradients: 0.10673370462850122
l2 norm of weights: 2.227835810859972

Iteration Number: 5705
Loss: 12.363103291175495
l2 norm of gradients: 0.10670389464715745
l2 norm of weights: 2.2277869528663183

Iteration Number: 5706
Loss: 12.361986610830215
l2 norm of gradients: 0.1066741012705564
l2 norm of weights: 2.2277381655778448

Iteration Number: 5707
Loss: 12.360870678309047
l2 norm of gradients: 0.10664432448748064
l2 norm of weights: 2.2276894489431776

Iteration Number: 5708
Loss: 12.359755493031445
l2 norm of gradients: 0.10661456428671724
l2 norm of weights: 2.2276408029109787

Iteration Number: 5709
Loss: 12.358641054417372
l2 norm of gradients: 0.10658482065705772
l2 norm of weights: 2.2275922274299442

```
-----  
Iteration Number: 5710  
Loss: 12.3575273618873  
l2 norm of gradients: 0.10655509358729823  
l2 norm of weights: 2.227543722448806  
-----  
Iteration Number: 5711  
Loss: 12.356414414862208  
l2 norm of gradients: 0.1065253830662394  
l2 norm of weights: 2.227495287916329  
-----  
Iteration Number: 5712  
Loss: 12.35530221276359  
l2 norm of gradients: 0.1064956890826863  
l2 norm of weights: 2.2274469237813137  
-----  
Iteration Number: 5713  
Loss: 12.354190755013446  
l2 norm of gradients: 0.10646601162544876  
l2 norm of weights: 2.2273986299925967  
-----  
Iteration Number: 5714  
Loss: 12.353080041034284  
l2 norm of gradients: 0.10643635068334086  
l2 norm of weights: 2.2273504064990477  
-----  
Iteration Number: 5715  
Loss: 12.351970070249124  
l2 norm of gradients: 0.10640670624518152  
l2 norm of weights: 2.2273022532495714  
-----  
Iteration Number: 5716  
Loss: 12.350860842081483  
l2 norm of gradients: 0.10637707829979406  
l2 norm of weights: 2.2272541701931075  
-----  
Iteration Number: 5717  
Loss: 12.349752355955404  
l2 norm of gradients: 0.10634746683600645  
l2 norm of weights: 2.2272061572786304  
-----  
Iteration Number: 5718  
Loss: 12.348644611295416  
l2 norm of gradients: 0.10631787184265115  
l2 norm of weights: 2.227158214455149  
-----  
Iteration Number: 5719  
Loss: 12.347537607526567
```

```

12 norm of gradients: 0.1062882933085653
12 norm of weights: 2.2271103416717066
-----
Iteration Number: 5720
Loss: 12.346431344074405
12 norm of gradients: 0.10625873122259052
12 norm of weights: 2.227062538877381
-----
Iteration Number: 5721
Loss: 12.345325820364991
12 norm of gradients: 0.1062291855735732
12 norm of weights: 2.2270148060212853
-----
Iteration Number: 5722
Loss: 12.34422103582488
12 norm of gradients: 0.10619965635036419
12 norm of weights: 2.2269671430525673
-----
Iteration Number: 5723
Loss: 12.343116989881139
12 norm of gradients: 0.1061701435418189
12 norm of weights: 2.226919549920407
-----
Iteration Number: 5724
Loss: 12.342013681961339
12 norm of gradients: 0.10614064713679762
12 norm of weights: 2.2268720265740223
-----
Iteration Number: 5725
Loss: 12.340911111493552
12 norm of gradients: 0.10611116712416502
12 norm of weights: 2.2268245729626632
-----
Iteration Number: 5726
Loss: 12.339809277906353
12 norm of gradients: 0.10608170349279046
12 norm of weights: 2.226777189035615
-----
Iteration Number: 5727
Loss: 12.338708180628819
12 norm of gradients: 0.10605225623154804
12 norm of weights: 2.226729874742197
-----
Iteration Number: 5728
Loss: 12.337607819090536
12 norm of gradients: 0.10602282532931642
12 norm of weights: 2.226682630031764
-----

```

Iteration Number: 5729
Loss: 12.336508192721583
l2 norm of gradients: 0.10599341077497897
l2 norm of weights: 2.226635454853704

Iteration Number: 5730
Loss: 12.335409300952545
l2 norm of gradients: 0.10596401255742362
l2 norm of weights: 2.2265883491574394

Iteration Number: 5731
Loss: 12.334311143214507
l2 norm of gradients: 0.1059346306655431
l2 norm of weights: 2.226541312892428

Iteration Number: 5732
Loss: 12.333213718939056
l2 norm of gradients: 0.10590526508823477
l2 norm of weights: 2.226494346008161

Iteration Number: 5733
Loss: 12.332117027558278
l2 norm of gradients: 0.10587591581440063
l2 norm of weights: 2.2264474484541643

Iteration Number: 5734
Loss: 12.331021068504757
l2 norm of gradients: 0.10584658283294744
l2 norm of weights: 2.2264006201799984

Iteration Number: 5735
Loss: 12.329925841211583
l2 norm of gradients: 0.10581726613278668
l2 norm of weights: 2.226353861135257

Iteration Number: 5736
Loss: 12.328831345112333
l2 norm of gradients: 0.10578796570283444
l2 norm of weights: 2.2263071712695686

Iteration Number: 5737
Loss: 12.327737579641093
l2 norm of gradients: 0.1057586815320116
l2 norm of weights: 2.2262605505325967

Iteration Number: 5738
Loss: 12.32664454423244
l2 norm of gradients: 0.10572941360924365

```

12 norm of weights: 2.226213998874038
-----
Iteration Number: 5739
Loss: 12.325552238321453
12 norm of gradients: 0.10570016192346102
12 norm of weights: 2.2261675162436236
-----
Iteration Number: 5740
Loss: 12.324460661343705
12 norm of gradients: 0.10567092646359864
12 norm of weights: 2.226121102591119
-----
Iteration Number: 5741
Loss: 12.323369812735269
12 norm of gradients: 0.10564170721859634
12 norm of weights: 2.2260747578663236
-----
Iteration Number: 5742
Loss: 12.322279691932712
12 norm of gradients: 0.10561250417739869
12 norm of weights: 2.226028482019071
-----
Iteration Number: 5743
Loss: 12.321190298373098
12 norm of gradients: 0.10558331732895494
12 norm of weights: 2.2259822749992293
-----
Iteration Number: 5744
Loss: 12.320101631493985
12 norm of gradients: 0.10555414666221917
12 norm of weights: 2.2259361367566997
-----
Iteration Number: 5745
Loss: 12.31901369073342
12 norm of gradients: 0.1055249921661502
12 norm of weights: 2.2258900672414184
-----
Iteration Number: 5746
Loss: 12.317926475529957
12 norm of gradients: 0.1054958538297116
12 norm of weights: 2.2258440664033543
-----
Iteration Number: 5747
Loss: 12.316839985322634
12 norm of gradients: 0.1054667316418718
12 norm of weights: 2.225798134192513
-----
Iteration Number: 5748

```

```

Loss: 12.315754219550993
12 norm of gradients: 0.105437625591604
12 norm of weights: 2.2257522705589303
-----
Iteration Number: 5749
Loss: 12.314669177655052
12 norm of gradients: 0.1054085356678861
12 norm of weights: 2.2257064754526796
-----
Iteration Number: 5750
Loss: 12.313584859075338
12 norm of gradients: 0.10537946185970101
12 norm of weights: 2.2256607488238656
-----
Iteration Number: 5751
Loss: 12.312501263252862
12 norm of gradients: 0.10535040415603619
12 norm of weights: 2.2256150906226284
-----
Iteration Number: 5752
Loss: 12.311418389629129
12 norm of gradients: 0.10532136254588416
12 norm of weights: 2.2255695007991414
-----
Iteration Number: 5753
Loss: 12.310336237646137
12 norm of gradients: 0.10529233701824219
12 norm of weights: 2.225523979303612
-----
Iteration Number: 5754
Loss: 12.30925480674637
12 norm of gradients: 0.1052633275621123
12 norm of weights: 2.225478526086281
-----
Iteration Number: 5755
Loss: 12.308174096372806
12 norm of gradients: 0.10523433416650142
12 norm of weights: 2.225433141097424
-----
Iteration Number: 5756
Loss: 12.307094105968918
12 norm of gradients: 0.10520535682042137
12 norm of weights: 2.22538782428735
-----
Iteration Number: 5757
Loss: 12.30601483497866
12 norm of gradients: 0.10517639551288875
12 norm of weights: 2.225342575606401

```

```
-----
Iteration Number: 5758
Loss: 12.304936282846475
l2 norm of gradients: 0.10514745023292507
l2 norm of weights: 2.225297395004954
-----
Iteration Number: 5759
Loss: 12.303858449017302
l2 norm of gradients: 0.1051185209695567
l2 norm of weights: 2.2252522824334187
-----
Iteration Number: 5760
Loss: 12.302781332936561
l2 norm of gradients: 0.10508960771181483
l2 norm of weights: 2.2252072378422394
-----
Iteration Number: 5761
Loss: 12.301704934050171
l2 norm of gradients: 0.10506071044873565
l2 norm of weights: 2.2251622611818935
-----
Iteration Number: 5762
Loss: 12.300629251804526
l2 norm of gradients: 0.10503182916936016
l2 norm of weights: 2.225117352402892
-----
Iteration Number: 5763
Loss: 12.299554285646511
l2 norm of gradients: 0.10500296386273425
l2 norm of weights: 2.2250725114557803
-----
Iteration Number: 5764
Loss: 12.2984800350235
l2 norm of gradients: 0.10497411451790874
l2 norm of weights: 2.2250277382911365
-----
Iteration Number: 5765
Loss: 12.297406499383351
l2 norm of gradients: 0.10494528112393937
l2 norm of weights: 2.2249830328595728
-----
Iteration Number: 5766
Loss: 12.296333678174413
l2 norm of gradients: 0.10491646366988677
l2 norm of weights: 2.2249383951117347
-----
Iteration Number: 5767
Loss: 12.295261570845508
```



```

12 norm of gradients: 0.10488766214481655
12 norm of weights: 2.224893824998302
-----
Iteration Number: 5768
Loss: 12.294190176845957
12 norm of gradients: 0.10485887653779916
12 norm of weights: 2.224849322469987
-----
Iteration Number: 5769
Loss: 12.293119495625557
12 norm of gradients: 0.1048301068379101
12 norm of weights: 2.2248048874775366
-----
Iteration Number: 5770
Loss: 12.292049526634589
12 norm of gradients: 0.10480135303422972
12 norm of weights: 2.2247605199717304
-----
Iteration Number: 5771
Loss: 12.290980269323823
12 norm of gradients: 0.10477261511584332
12 norm of weights: 2.2247162199033816
-----
Iteration Number: 5772
Loss: 12.289911723144508
12 norm of gradients: 0.1047438930718413
12 norm of weights: 2.2246719872233376
-----
Iteration Number: 5773
Loss: 12.288843887548378
12 norm of gradients: 0.10471518689131885
12 norm of weights: 2.2246278218824775
-----
Iteration Number: 5774
Loss: 12.287776761987644
12 norm of gradients: 0.10468649656337622
12 norm of weights: 2.224583723831716
-----
Iteration Number: 5775
Loss: 12.286710345915
12 norm of gradients: 0.10465782207711867
12 norm of weights: 2.224539693021999
-----
Iteration Number: 5776
Loss: 12.285644638783632
12 norm of gradients: 0.10462916342165635
12 norm of weights: 2.224495729404308
-----

```

Iteration Number: 5777
Loss: 12.284579640047196
l2 norm of gradients: 0.10460052058610449
l2 norm of weights: 2.2244518329296565

Iteration Number: 5778
Loss: 12.28351534915983
l2 norm of gradients: 0.10457189355958328
l2 norm of weights: 2.224408003549091

Iteration Number: 5779
Loss: 12.282451765576154
l2 norm of gradients: 0.10454328233121794
l2 norm of weights: 2.224364241213693

Iteration Number: 5780
Loss: 12.281388888751273
l2 norm of gradients: 0.10451468689013868
l2 norm of weights: 2.224320545874575

Iteration Number: 5781
Loss: 12.280326718140754
l2 norm of gradients: 0.10448610722548073
l2 norm of weights: 2.2242769174828845

Iteration Number: 5782
Loss: 12.279265253200666
l2 norm of gradients: 0.10445754332638432
l2 norm of weights: 2.224233355989801

Iteration Number: 5783
Loss: 12.278204493387534
l2 norm of gradients: 0.10442899518199487
l2 norm of weights: 2.2241898613465394

Iteration Number: 5784
Loss: 12.277144438158386
l2 norm of gradients: 0.1044004627814626
l2 norm of weights: 2.2241464335043446

Iteration Number: 5785
Loss: 12.276085086970701
l2 norm of gradients: 0.10437194611394296
l2 norm of weights: 2.2241030724144975

Iteration Number: 5786
Loss: 12.275026439282453
l2 norm of gradients: 0.10434344516859638

```

12 norm of weights: 2.2240597780283102
-----
Iteration Number: 5787
Loss: 12.273968494552086
12 norm of gradients: 0.10431495993458836
12 norm of weights: 2.224016550297129
-----
Iteration Number: 5788
Loss: 12.272911252238517
12 norm of gradients: 0.10428649040108948
12 norm of weights: 2.2239733891723334
-----
Iteration Number: 5789
Loss: 12.271854711801153
12 norm of gradients: 0.10425803655727534
12 norm of weights: 2.2239302946053354
-----
Iteration Number: 5790
Loss: 12.270798872699862
12 norm of gradients: 0.10422959839232672
12 norm of weights: 2.2238872665475804
-----
Iteration Number: 5791
Loss: 12.269743734394986
12 norm of gradients: 0.10420117589542942
12 norm of weights: 2.2238443049505463
-----
Iteration Number: 5792
Loss: 12.268689296347356
12 norm of gradients: 0.10417276905577429
12 norm of weights: 2.223801409765745
-----
Iteration Number: 5793
Loss: 12.267635558018267
12 norm of gradients: 0.10414437786255742
12 norm of weights: 2.2237585809447205
-----
Iteration Number: 5794
Loss: 12.266582518869482
12 norm of gradients: 0.1041160023049799
12 norm of weights: 2.22371581843905
-----
Iteration Number: 5795
Loss: 12.26553017836325
12 norm of gradients: 0.1040876423722479
12 norm of weights: 2.2236731222003447
-----
Iteration Number: 5796

```

Loss: 12.264478535962287
l2 norm of gradients: 0.10405929805357281
l2 norm of weights: 2.223630492180247

Iteration Number: 5797
Loss: 12.263427591129778
l2 norm of gradients: 0.10403096933817108
l2 norm of weights: 2.223587928330434

Iteration Number: 5798
Loss: 12.262377343329387
l2 norm of gradients: 0.10400265621526436
l2 norm of weights: 2.2235454306026137

Iteration Number: 5799
Loss: 12.261327792025247
l2 norm of gradients: 0.10397435867407936
l2 norm of weights: 2.223502998948528

Iteration Number: 5800
Loss: 12.260278936681956
l2 norm of gradients: 0.10394607670384796
l2 norm of weights: 2.2234606333199527

Iteration Number: 5801
Loss: 12.259230776764593
l2 norm of gradients: 0.10391781029380726
l2 norm of weights: 2.2234183336686946

Iteration Number: 5802
Loss: 12.258183311738701
l2 norm of gradients: 0.10388955943319939
l2 norm of weights: 2.223376099946594

Iteration Number: 5803
Loss: 12.257136541070292
l2 norm of gradients: 0.10386132411127172
l2 norm of weights: 2.2233339321055254

Iteration Number: 5804
Loss: 12.25609046422585
l2 norm of gradients: 0.10383310431727684
l2 norm of weights: 2.223291830097393

Iteration Number: 5805
Loss: 12.255045080672328
l2 norm of gradients: 0.10380490004047244
l2 norm of weights: 2.2232497938741367

```

-----
Iteration Number: 5806
Loss: 12.25400038987715
l2 norm of gradients: 0.10377671127012147
l2 norm of weights: 2.223207823387727
-----
Iteration Number: 5807
Loss: 12.252956391308198
l2 norm of gradients: 0.10374853799549194
l2 norm of weights: 2.223165918590169
-----
Iteration Number: 5808
Loss: 12.251913084433836
l2 norm of gradients: 0.10372038020585715
l2 norm of weights: 2.2231240794334988
-----
Iteration Number: 5809
Loss: 12.250870468722882
l2 norm of gradients: 0.10369223789049563
l2 norm of weights: 2.2230823058697866
-----
Iteration Number: 5810
Loss: 12.249828543644634
l2 norm of gradients: 0.10366411103869108
l2 norm of weights: 2.223040597851134
-----
Iteration Number: 5811
Loss: 12.248787308668843
l2 norm of gradients: 0.10363599963973241
l2 norm of weights: 2.222998955329676
-----
Iteration Number: 5812
Loss: 12.247746763265736
l2 norm of gradients: 0.10360790368291377
l2 norm of weights: 2.222957378257579
-----
Iteration Number: 5813
Loss: 12.246706906906002
l2 norm of gradients: 0.10357982315753454
l2 norm of weights: 2.2229158665870443
-----
Iteration Number: 5814
Loss: 12.245667739060796
l2 norm of gradients: 0.10355175805289933
l2 norm of weights: 2.2228744202703035
-----
Iteration Number: 5815
Loss: 12.244629259201734

```

l2 norm of gradients: 0.10352370835831798
l2 norm of weights: 2.2228330392596214

Iteration Number: 5816
Loss: 12.2435914668009
l2 norm of gradients: 0.10349567406310556
l2 norm of weights: 2.222791723507296

Iteration Number: 5817
Loss: 12.242554361330843
l2 norm of gradients: 0.10346765515658243
l2 norm of weights: 2.222750472965657

Iteration Number: 5818
Loss: 12.241517942264576
l2 norm of gradients: 0.1034396516280743
l2 norm of weights: 2.222709287587066

Iteration Number: 5819
Loss: 12.240482209075571
l2 norm of gradients: 0.10341166346691187
l2 norm of weights: 2.2226681673239193

Iteration Number: 5820
Loss: 12.23944716123776
l2 norm of gradients: 0.10338369066243146
l2 norm of weights: 2.2226271121286434

Iteration Number: 5821
Loss: 12.238412798225546
l2 norm of gradients: 0.10335573320397441
l2 norm of weights: 2.2225861219536984

Iteration Number: 5822
Loss: 12.237379119513793
l2 norm of gradients: 0.10332779108088742
l2 norm of weights: 2.222545196751576

Iteration Number: 5823
Loss: 12.236346124577818
l2 norm of gradients: 0.10329986428252254
l2 norm of weights: 2.2225043364748007

Iteration Number: 5824
Loss: 12.235313812893404
l2 norm of gradients: 0.10327195279823707
l2 norm of weights: 2.2224635410759292

Iteration Number: 5825
Loss: 12.234282183936795
l2 norm of gradients: 0.10324405661739364
l2 norm of weights: 2.2224228105075508

Iteration Number: 5826
Loss: 12.233251237184698
l2 norm of gradients: 0.10321617572936015
l2 norm of weights: 2.2223821447222862

Iteration Number: 5827
Loss: 12.232220972114272
l2 norm of gradients: 0.10318831012350979
l2 norm of weights: 2.2223415436727896

Iteration Number: 5828
Loss: 12.23119138820314
l2 norm of gradients: 0.10316045978922114
l2 norm of weights: 2.222301007311747

Iteration Number: 5829
Loss: 12.230162484929386
l2 norm of gradients: 0.10313262471587814
l2 norm of weights: 2.2222605355918765

Iteration Number: 5830
Loss: 12.229134261771547
l2 norm of gradients: 0.10310480489286994
l2 norm of weights: 2.2222201284659278

Iteration Number: 5831
Loss: 12.228106718208625
l2 norm of gradients: 0.1030770003095911
l2 norm of weights: 2.2221797858866834

Iteration Number: 5832
Loss: 12.227079853720069
l2 norm of gradients: 0.10304921095544158
l2 norm of weights: 2.2221395078069586

Iteration Number: 5833
Loss: 12.226053667785798
l2 norm of gradients: 0.10302143681982659
l2 norm of weights: 2.2220992941796007

Iteration Number: 5834
Loss: 12.225028159886179
l2 norm of gradients: 0.10299367789215672

```

12 norm of weights: 2.222059144957487
-----
Iteration Number: 5835
Loss: 12.224003329502038
12 norm of gradients: 0.10296593416184802
12 norm of weights: 2.2220190600935297
-----
Iteration Number: 5836
Loss: 12.222979176114654
12 norm of gradients: 0.10293820561832172
12 norm of weights: 2.2219790395406718
-----
Iteration Number: 5837
Loss: 12.221955699205768
12 norm of gradients: 0.1029104922510047
12 norm of weights: 2.2219390832518884
-----
Iteration Number: 5838
Loss: 12.220932898257571
12 norm of gradients: 0.10288279404932893
12 norm of weights: 2.221899191180186
-----
Iteration Number: 5839
Loss: 12.219910772752709
12 norm of gradients: 0.10285511100273195
12 norm of weights: 2.221859363278605
-----
Iteration Number: 5840
Loss: 12.218889322174283
12 norm of gradients: 0.10282744310065668
12 norm of weights: 2.2218195995002157
-----
Iteration Number: 5841
Loss: 12.217868546005851
12 norm of gradients: 0.1027997903325514
12 norm of weights: 2.2217798997981215
-----
Iteration Number: 5842
Loss: 12.216848443731418
12 norm of gradients: 0.10277215268786977
12 norm of weights: 2.221740264125458
-----
Iteration Number: 5843
Loss: 12.215829014835448
12 norm of gradients: 0.10274453015607089
12 norm of weights: 2.221700692435392
-----
Iteration Number: 5844

```


Loss: 12.214810258802851
l2 norm of gradients: 0.1027169227266193
l2 norm of weights: 2.2216611846811225

Iteration Number: 5845
Loss: 12.213792175119002
l2 norm of gradients: 0.10268933038898491
l2 norm of weights: 2.22162174081588

Iteration Number: 5846
Loss: 12.212774763269708
l2 norm of gradients: 0.10266175313264313
l2 norm of weights: 2.2215823607929273

Iteration Number: 5847
Loss: 12.211758022741247
l2 norm of gradients: 0.10263419094707474
l2 norm of weights: 2.2215430445655597

Iteration Number: 5848
Loss: 12.210741953020333
l2 norm of gradients: 0.10260664382176601
l2 norm of weights: 2.2215037920871032

Iteration Number: 5849
Loss: 12.209726553594141
l2 norm of gradients: 0.10257911174620857
l2 norm of weights: 2.221464603310916

Iteration Number: 5850
Loss: 12.208711823950287
l2 norm of gradients: 0.10255159470989963
l2 norm of weights: 2.2214254781903873

Iteration Number: 5851
Loss: 12.207697763576848
l2 norm of gradients: 0.10252409270234178
l2 norm of weights: 2.2213864166789405

Iteration Number: 5852
Loss: 12.206684371962341
l2 norm of gradients: 0.10249660571304306
l2 norm of weights: 2.2213474187300277

Iteration Number: 5853
Loss: 12.205671648595732
l2 norm of gradients: 0.10246913373151699
l2 norm of weights: 2.221308484297135

```

-----
Iteration Number: 5854
Loss: 12.204659592966438
l2 norm of gradients: 0.1024416767472826
l2 norm of weights: 2.221269613333779
-----
Iteration Number: 5855
Loss: 12.20364820456433
l2 norm of gradients: 0.10241423474986436
l2 norm of weights: 2.2212308057935073
-----
Iteration Number: 5856
Loss: 12.202637482879714
l2 norm of gradients: 0.10238680772879225
l2 norm of weights: 2.221192061629902
-----
Iteration Number: 5857
Loss: 12.201627427403352
l2 norm of gradients: 0.10235939567360171
l2 norm of weights: 2.221153380796574
-----
Iteration Number: 5858
Loss: 12.20061803762645
l2 norm of gradients: 0.10233199857383372
l2 norm of weights: 2.2211147632471664
-----
Iteration Number: 5859
Loss: 12.199609313040662
l2 norm of gradients: 0.10230461641903474
l2 norm of weights: 2.2210762089353553
-----
Iteration Number: 5860
Loss: 12.19860125313809
l2 norm of gradients: 0.10227724919875672
l2 norm of weights: 2.2210377178148466
-----
Iteration Number: 5861
Loss: 12.197593857411274
l2 norm of gradients: 0.10224989690255715
l2 norm of weights: 2.220999289839379
-----
Iteration Number: 5862
Loss: 12.196587125353203
l2 norm of gradients: 0.10222255951999905
l2 norm of weights: 2.2209609249627214
-----
Iteration Number: 5863
Loss: 12.195581056457314

```

```

12 norm of gradients: 0.10219523704065091
12 norm of weights: 2.220922623138676
-----
Iteration Number: 5864
Loss: 12.194575650217482
12 norm of gradients: 0.10216792945408674
12 norm of weights: 2.2208843843210753
-----
Iteration Number: 5865
Loss: 12.193570906128034
12 norm of gradients: 0.10214063674988624
12 norm of weights: 2.2208462084637834
-----
Iteration Number: 5866
Loss: 12.192566823683736
12 norm of gradients: 0.10211335891763453
12 norm of weights: 2.2208080955206957
-----
Iteration Number: 5867
Loss: 12.191563402379792
12 norm of gradients: 0.10208609594692214
12 norm of weights: 2.2207700454457395
-----
Iteration Number: 5868
Loss: 12.190560641711855
12 norm of gradients: 0.10205884782734546
12 norm of weights: 2.2207320581928736
-----
Iteration Number: 5869
Loss: 12.18955854117602
12 norm of gradients: 0.10203161454850618
12 norm of weights: 2.220694133716087
-----
Iteration Number: 5870
Loss: 12.188557100268826
12 norm of gradients: 0.10200439610001165
12 norm of weights: 2.220656271969402
-----
Iteration Number: 5871
Loss: 12.187556318487244
12 norm of gradients: 0.10197719247147483
12 norm of weights: 2.2206184729068705
-----
Iteration Number: 5872
Loss: 12.186556195328695
12 norm of gradients: 0.10195000365251417
12 norm of weights: 2.2205807364825767
-----

```

Iteration Number: 5873
Loss: 12.185556730291038
l2 norm of gradients: 0.10192282963275374
l2 norm of weights: 2.2205430626506355

Iteration Number: 5874
Loss: 12.184557922872573
l2 norm of gradients: 0.10189567040182317
l2 norm of weights: 2.220505451365194

Iteration Number: 5875
Loss: 12.183559772572034
l2 norm of gradients: 0.10186852594935772
l2 norm of weights: 2.2204679025804284

Iteration Number: 5876
Loss: 12.182562278888604
l2 norm of gradients: 0.10184139626499822
l2 norm of weights: 2.220430416250549

Iteration Number: 5877
Loss: 12.181565441321895
l2 norm of gradients: 0.10181428133839107
l2 norm of weights: 2.220392992329796

Iteration Number: 5878
Loss: 12.18056925937197
l2 norm of gradients: 0.10178718115918832
l2 norm of weights: 2.2203556307724406

Iteration Number: 5879
Loss: 12.179573732539314
l2 norm of gradients: 0.1017600957170476
l2 norm of weights: 2.2203183315327846

Iteration Number: 5880
Loss: 12.178578860324865
l2 norm of gradients: 0.10173302500163221
l2 norm of weights: 2.2202810945651628

Iteration Number: 5881
Loss: 12.177584642229988
l2 norm of gradients: 0.10170596900261095
l2 norm of weights: 2.2202439198239388

Iteration Number: 5882
Loss: 12.17659107775649
l2 norm of gradients: 0.10167892770965839

```
12 norm of weights: 2.22020680726351
-----
Iteration Number: 5883
Loss: 12.175598166406614
12 norm of gradients: 0.10165190111245463
12 norm of weights: 2.220169756838302
-----
Iteration Number: 5884
Loss: 12.174605907683036
12 norm of gradients: 0.10162488920068546
12 norm of weights: 2.2201327685027734
-----
Iteration Number: 5885
Loss: 12.173614301088872
12 norm of gradients: 0.10159789196404227
12 norm of weights: 2.220095842211414
-----
Iteration Number: 5886
Loss: 12.172623346127668
12 norm of gradients: 0.10157090939222206
12 norm of weights: 2.220058977918743
-----
Iteration Number: 5887
Loss: 12.171633042303416
12 norm of gradients: 0.10154394147492764
12 norm of weights: 2.2200221755793126
-----
Iteration Number: 5888
Loss: 12.170643389120524
12 norm of gradients: 0.10151698820186729
12 norm of weights: 2.2199854351477044
-----
Iteration Number: 5889
Loss: 12.169654386083849
12 norm of gradients: 0.10149004956275506
12 norm of weights: 2.219948756578531
-----
Iteration Number: 5890
Loss: 12.168666032698681
12 norm of gradients: 0.10146312554731061
12 norm of weights: 2.219912139826437
-----
Iteration Number: 5891
Loss: 12.167678328470735
12 norm of gradients: 0.1014362161452593
12 norm of weights: 2.219875584846098
-----
Iteration Number: 5892
```

Loss: 12.166691272906167
l2 norm of gradients: 0.1014093213463322
l2 norm of weights: 2.2198390915922186

Iteration Number: 5893
Loss: 12.165704865511557
l2 norm of gradients: 0.10138244114026594
l2 norm of weights: 2.2198026600195364

Iteration Number: 5894
Loss: 12.164719105793925
l2 norm of gradients: 0.10135557551680302
l2 norm of weights: 2.219766290082819

Iteration Number: 5895
Loss: 12.16373399326072
l2 norm of gradients: 0.10132872446569144
l2 norm of weights: 2.219729981736865

Iteration Number: 5896
Loss: 12.162749527419821
l2 norm of gradients: 0.10130188797668498
l2 norm of weights: 2.2196937349365036

Iteration Number: 5897
Loss: 12.161765707779537
l2 norm of gradients: 0.10127506603954319
l2 norm of weights: 2.2196575496365947

Iteration Number: 5898
Loss: 12.160782533848614
l2 norm of gradients: 0.1012482586440312
l2 norm of weights: 2.2196214257920293

Iteration Number: 5899
Loss: 12.159800005136216
l2 norm of gradients: 0.10122146577991992
l2 norm of weights: 2.219585363357729

Iteration Number: 5900
Loss: 12.158818121151945
l2 norm of gradients: 0.10119468743698595
l2 norm of weights: 2.2195493622886464

Iteration Number: 5901
Loss: 12.157836881405837
l2 norm of gradients: 0.10116792360501163
l2 norm of weights: 2.2195134225397646

```

-----
Iteration Number: 5902
Loss: 12.156856285408347
l2 norm of gradients: 0.10114117427378501
l2 norm of weights: 2.219477544066097
-----
Iteration Number: 5903
Loss: 12.155876332670358
l2 norm of gradients: 0.10111443943309985
l2 norm of weights: 2.2194417268226885
-----
Iteration Number: 5904
Loss: 12.154897022703192
l2 norm of gradients: 0.10108771907275568
l2 norm of weights: 2.2194059707646137
-----
Iteration Number: 5905
Loss: 12.153918355018588
l2 norm of gradients: 0.10106101318255777
l2 norm of weights: 2.219370275846979
-----
Iteration Number: 5906
Loss: 12.152940329128715
l2 norm of gradients: 0.10103432175231707
l2 norm of weights: 2.2193346420249207
-----
Iteration Number: 5907
Loss: 12.151962944546172
l2 norm of gradients: 0.10100764477185037
l2 norm of weights: 2.2192990692536054
-----
Iteration Number: 5908
Loss: 12.150986200783981
l2 norm of gradients: 0.10098098223098013
l2 norm of weights: 2.219263557488231
-----
Iteration Number: 5909
Loss: 12.150010097355592
l2 norm of gradients: 0.10095433411953458
l2 norm of weights: 2.219228106684025
-----
Iteration Number: 5910
Loss: 12.149034633774875
l2 norm of gradients: 0.10092770042734779
l2 norm of weights: 2.219192716796247
-----
Iteration Number: 5911
Loss: 12.14805980955614

```

```
12 norm of gradients: 0.10090108114425948
12 norm of weights: 2.219157387780185
-----
Iteration Number: 5912
Loss: 12.1470856242141
12 norm of gradients: 0.10087447626011521
12 norm of weights: 2.2191221195911597
-----
Iteration Number: 5913
Loss: 12.146112077263911
12 norm of gradients: 0.10084788576476633
12 norm of weights: 2.2190869121845207
-----
Iteration Number: 5914
Loss: 12.145139168221144
12 norm of gradients: 0.10082130964806994
12 norm of weights: 2.219051765515648
-----
Iteration Number: 5915
Loss: 12.1441668966018
12 norm of gradients: 0.10079474789988888
12 norm of weights: 2.219016679539954
-----
Iteration Number: 5916
Loss: 12.143195261922287
12 norm of gradients: 0.10076820051009186
12 norm of weights: 2.218981654212879
-----
Iteration Number: 5917
Loss: 12.14222426369946
12 norm of gradients: 0.10074166746855334
12 norm of weights: 2.2189466894898944
-----
Iteration Number: 5918
Loss: 12.141253901450579
12 norm of gradients: 0.10071514876515358
12 norm of weights: 2.2189117853265037
-----
Iteration Number: 5919
Loss: 12.140284174693333
12 norm of gradients: 0.10068864438977868
12 norm of weights: 2.2188769416782383
-----
Iteration Number: 5920
Loss: 12.139315082945828
12 norm of gradients: 0.10066215433232044
12 norm of weights: 2.2188421585006615
-----
```


Iteration Number: 5921
Loss: 12.138346625726594
l2 norm of gradients: 0.10063567858267661
l2 norm of weights: 2.2188074357493663

Iteration Number: 5922
Loss: 12.137378802554585
l2 norm of gradients: 0.10060921713075067
l2 norm of weights: 2.218772773379976

Iteration Number: 5923
Loss: 12.136411612949168
l2 norm of gradients: 0.10058276996645188
l2 norm of weights: 2.2187381713481447

Iteration Number: 5924
Loss: 12.135445056430138
l2 norm of gradients: 0.10055633707969545
l2 norm of weights: 2.218703629609556

Iteration Number: 5925
Loss: 12.134479132517697
l2 norm of gradients: 0.10052991846040234
l2 norm of weights: 2.2186691481199237

Iteration Number: 5926
Loss: 12.133513840732487
l2 norm of gradients: 0.1005035140984993
l2 norm of weights: 2.2186347268349933

Iteration Number: 5927
Loss: 12.13254918059555
l2 norm of gradients: 0.10047712398391906
l2 norm of weights: 2.218600365710538

Iteration Number: 5928
Loss: 12.131585151628354
l2 norm of gradients: 0.10045074810659999
l2 norm of weights: 2.218566064702363

Iteration Number: 5929
Loss: 12.130621753352784
l2 norm of gradients: 0.10042438645648646
l2 norm of weights: 2.218531823766304

Iteration Number: 5930
Loss: 12.129658985291142
l2 norm of gradients: 0.10039803902352867

```

12 norm of weights: 2.2184976428582246
-----
Iteration Number: 5931
Loss: 12.128696846966147
12 norm of gradients: 0.10037170579768258
12 norm of weights: 2.2184635219340207
-----
Iteration Number: 5932
Loss: 12.12773533790094
12 norm of gradients: 0.10034538676891015
12 norm of weights: 2.218429460949617
-----
Iteration Number: 5933
Loss: 12.12677445761907
12 norm of gradients: 0.10031908192717905
12 norm of weights: 2.2183954598609694
-----
Iteration Number: 5934
Loss: 12.125814205644513
12 norm of gradients: 0.10029279126246295
12 norm of weights: 2.218361518624063
-----
Iteration Number: 5935
Loss: 12.124854581501644
12 norm of gradients: 0.10026651476474127
12 norm of weights: 2.2183276371949123
-----
Iteration Number: 5936
Loss: 12.123895584715275
12 norm of gradients: 0.10024025242399943
12 norm of weights: 2.218293815529563
-----
Iteration Number: 5937
Loss: 12.12293721481061
12 norm of gradients: 0.1002140042302286
12 norm of weights: 2.2182600535840913
-----
Iteration Number: 5938
Loss: 12.121979471313285
12 norm of gradients: 0.10018777017342594
12 norm of weights: 2.218226351314601
-----
Iteration Number: 5939
Loss: 12.121022353749341
12 norm of gradients: 0.10016155024359442
12 norm of weights: 2.2181927086772277
-----
Iteration Number: 5940

```

Loss: 12.120065861645237
l2 norm of gradients: 0.10013534443074296
l2 norm of weights: 2.218159125628137

Iteration Number: 5941
Loss: 12.119109994527843
l2 norm of gradients: 0.10010915272488628
l2 norm of weights: 2.218125602123524

Iteration Number: 5942
Loss: 12.118154751924443
l2 norm of gradients: 0.10008297511604512
l2 norm of weights: 2.218092138119613

Iteration Number: 5943
Loss: 12.117200133362735
l2 norm of gradients: 0.10005681159424606
l2 norm of weights: 2.2180587335726583

Iteration Number: 5944
Loss: 12.11624613837082
l2 norm of gradients: 0.10003066214952162
l2 norm of weights: 2.218025388438946

Iteration Number: 5945
Loss: 12.115292766477225
l2 norm of gradients: 0.10000452677191013
l2 norm of weights: 2.217992102674789

Iteration Number: 5946
Loss: 12.114340017210882
l2 norm of gradients: 0.0999784054514559
l2 norm of weights: 2.217958876236532

Iteration Number: 5947
Loss: 12.113387890101126
l2 norm of gradients: 0.09995229817820928
l2 norm of weights: 2.217925709080549

Iteration Number: 5948
Loss: 12.112436384677714
l2 norm of gradients: 0.0999262049422263
l2 norm of weights: 2.217892601163244

Iteration Number: 5949
Loss: 12.111485500470812
l2 norm of gradients: 0.09990012573356907
l2 norm of weights: 2.2178595524410505

```
-----
Iteration Number: 5950
Loss: 12.11053523701099
l2 norm of gradients: 0.09987406054230562
l2 norm of weights: 2.2178265628704312
-----
Iteration Number: 5951
Loss: 12.109585593829225
l2 norm of gradients: 0.09984800935850988
l2 norm of weights: 2.2177936324078797
-----
Iteration Number: 5952
Loss: 12.10863657045692
l2 norm of gradients: 0.09982197217226171
l2 norm of weights: 2.2177607610099175
-----
Iteration Number: 5953
Loss: 12.107688166425861
l2 norm of gradients: 0.09979594897364696
l2 norm of weights: 2.217727948633098
-----
Iteration Number: 5954
Loss: 12.106740381268263
l2 norm of gradients: 0.09976993975275736
l2 norm of weights: 2.2176951952340023
-----
Iteration Number: 5955
Loss: 12.105793214516746
l2 norm of gradients: 0.09974394449969066
l2 norm of weights: 2.2176625007692428
-----
Iteration Number: 5956
Loss: 12.104846665704324
l2 norm of gradients: 0.0997179632045505
l2 norm of weights: 2.2176298651954593
-----
Iteration Number: 5957
Loss: 12.103900734364432
l2 norm of gradients: 0.0996919958574465
l2 norm of weights: 2.217597288469323
-----
Iteration Number: 5958
Loss: 12.102955420030906
l2 norm of gradients: 0.09966604244849421
l2 norm of weights: 2.2175647705475345
-----
Iteration Number: 5959
Loss: 12.102010722237988
```

```

12 norm of gradients: 0.0996401029678152
12 norm of weights: 2.2175323113868237
-----
Iteration Number: 5960
Loss: 12.101066640520333
12 norm of gradients: 0.09961417740553699
12 norm of weights: 2.2174999109439493
-----
Iteration Number: 5961
Loss: 12.100123174412985
12 norm of gradients: 0.09958826575179296
12 norm of weights: 2.2174675691757004
-----
Iteration Number: 5962
Loss: 12.099180323451415
12 norm of gradients: 0.09956236799672263
12 norm of weights: 2.2174352860388953
-----
Iteration Number: 5963
Loss: 12.09823808717148
12 norm of gradients: 0.09953648413047149
12 norm of weights: 2.2174030614903812
-----
Iteration Number: 5964
Loss: 12.09729646510945
12 norm of gradients: 0.09951061414319082
12 norm of weights: 2.2173708954870364
-----
Iteration Number: 5965
Loss: 12.096355456801998
12 norm of gradients: 0.0994847580250381
12 norm of weights: 2.2173387879857662
-----
Iteration Number: 5966
Loss: 12.095415061786198
12 norm of gradients: 0.09945891576617663
12 norm of weights: 2.217306738943508
-----
Iteration Number: 5967
Loss: 12.094475279599534
12 norm of gradients: 0.09943308735677583
12 norm of weights: 2.2172747483172257
-----
Iteration Number: 5968
Loss: 12.093536109779883
12 norm of gradients: 0.09940727278701104
12 norm of weights: 2.2172428160639153
-----

```

Iteration Number: 5969
Loss: 12.092597551865534
l2 norm of gradients: 0.09938147204706364
l2 norm of weights: 2.2172109421406

Iteration Number: 5970
Loss: 12.091659605395172
l2 norm of gradients: 0.09935568512712097
l2 norm of weights: 2.2171791265043335

Iteration Number: 5971
Loss: 12.090722269907884
l2 norm of gradients: 0.0993299120173764
l2 norm of weights: 2.2171473691121983

Iteration Number: 5972
Loss: 12.089785544943158
l2 norm of gradients: 0.09930415270802931
l2 norm of weights: 2.2171156699213066

Iteration Number: 5973
Loss: 12.088849430040892
l2 norm of gradients: 0.09927840718928506
l2 norm of weights: 2.2170840288888

Iteration Number: 5974
Loss: 12.08791392474137
l2 norm of gradients: 0.09925267545135505
l2 norm of weights: 2.2170524459718486

Iteration Number: 5975
Loss: 12.086979028585283
l2 norm of gradients: 0.09922695748445674
l2 norm of weights: 2.2170209211276517

Iteration Number: 5976
Loss: 12.08604474111372
l2 norm of gradients: 0.09920125327881356
l2 norm of weights: 2.216989454313439

Iteration Number: 5977
Loss: 12.08511106186818
l2 norm of gradients: 0.09917556282465491
l2 norm of weights: 2.2169580454864684

Iteration Number: 5978
Loss: 12.08417799039054
l2 norm of gradients: 0.09914988611221633

```
12 norm of weights: 2.2169266946040267
-----
Iteration Number: 5979
Loss: 12.083245526223099
12 norm of gradients: 0.0991242231317393
12 norm of weights: 2.2168954016234306
-----
Iteration Number: 5980
Loss: 12.08231366890853
12 norm of gradients: 0.09909857387347143
12 norm of weights: 2.2168641665020257
-----
Iteration Number: 5981
Loss: 12.08138241798993
12 norm of gradients: 0.09907293832766628
12 norm of weights: 2.2168329891971865
-----
Iteration Number: 5982
Loss: 12.080451773010768
12 norm of gradients: 0.09904731648458344
12 norm of weights: 2.2168018696663165
-----
Iteration Number: 5983
Loss: 12.079521733514927
12 norm of gradients: 0.09902170833448871
12 norm of weights: 2.216770807866849
-----
Iteration Number: 5984
Loss: 12.078592299046687
12 norm of gradients: 0.0989961138676537
12 norm of weights: 2.216739803756245
-----
Iteration Number: 5985
Loss: 12.077663469150712
12 norm of gradients: 0.0989705330743562
12 norm of weights: 2.2167088572919966
-----
Iteration Number: 5986
Loss: 12.07673524337207
12 norm of gradients: 0.09894496594488
12 norm of weights: 2.2166779684316222
-----
Iteration Number: 5987
Loss: 12.075807621256219
12 norm of gradients: 0.09891941246951509
12 norm of weights: 2.216647137132672
-----
Iteration Number: 5988
```

Loss: 12.074880602349026
l2 norm of gradients: 0.09889387263855733
l2 norm of weights: 2.2166163633527227

Iteration Number: 5989
Loss: 12.073954186196735
l2 norm of gradients: 0.09886834644230874
l2 norm of weights: 2.2165856470493814

Iteration Number: 5990
Loss: 12.073028372346002
l2 norm of gradients: 0.09884283387107738
l2 norm of weights: 2.216554988180284

Iteration Number: 5991
Loss: 12.072103160343854
l2 norm of gradients: 0.09881733491517738
l2 norm of weights: 2.216524386703095

Iteration Number: 5992
Loss: 12.071178549737734
l2 norm of gradients: 0.098791849564929
l2 norm of weights: 2.216493842575508

Iteration Number: 5993
Loss: 12.07025454007547
l2 norm of gradients: 0.09876637781065843
l2 norm of weights: 2.2164633557552444

Iteration Number: 5994
Loss: 12.06933113090528
l2 norm of gradients: 0.09874091964269809
l2 norm of weights: 2.2164329262000564

Iteration Number: 5995
Loss: 12.068408321775776
l2 norm of gradients: 0.09871547505138638
l2 norm of weights: 2.216402553867724

Iteration Number: 5996
Loss: 12.067486112235965
l2 norm of gradients: 0.09869004402706785
l2 norm of weights: 2.2163722387160556

Iteration Number: 5997
Loss: 12.066564501835241
l2 norm of gradients: 0.09866462656009307
l2 norm of weights: 2.216341980702889


```

-----
Iteration Number: 5998
Loss: 12.065643490123394
l2 norm of gradients: 0.09863922264081876
l2 norm of weights: 2.21631177978609
-----
Iteration Number: 5999
Loss: 12.064723076650601
l2 norm of gradients: 0.09861383225960765
l2 norm of weights: 2.2162816359235546
-----
Iteration Number: 6000
Loss: 12.063803260967436
l2 norm of gradients: 0.09858845540682874
l2 norm of weights: 2.2162515490732067
-----
Iteration Number: 6001
Loss: 12.062884042624853
l2 norm of gradients: 0.09856309207285688
l2 norm of weights: 2.2162215191929984
-----
Iteration Number: 6002
Loss: 12.061965421174207
l2 norm of gradients: 0.09853774224807323
l2 norm of weights: 2.2161915462409105
-----
Iteration Number: 6003
Loss: 12.061047396167234
l2 norm of gradients: 0.09851240592286485
l2 norm of weights: 2.216161630174954
-----
Iteration Number: 6004
Loss: 12.060129967156062
l2 norm of gradients: 0.09848708308762515
l2 norm of weights: 2.216131770953167
-----
Iteration Number: 6005
Loss: 12.05921313369321
l2 norm of gradients: 0.09846177373275346
l2 norm of weights: 2.2161019685336165
-----
Iteration Number: 6006
Loss: 12.058296895331578
l2 norm of gradients: 0.09843647784865528
l2 norm of weights: 2.2160722228743985
-----
Iteration Number: 6007
Loss: 12.057381251624467

```

```

12 norm of gradients: 0.09841119542574223
12 norm of weights: 2.216042533933638
-----
Iteration Number: 6008
Loss: 12.05646620212555
12 norm of gradients: 0.09838592645443202
12 norm of weights: 2.216012901669487
-----
Iteration Number: 6009
Loss: 12.0555517463889
12 norm of gradients: 0.09836067092514854
12 norm of weights: 2.2159833260401283
-----
Iteration Number: 6010
Loss: 12.054637883968972
12 norm of gradients: 0.09833542882832168
12 norm of weights: 2.2159538070037703
-----
Iteration Number: 6011
Loss: 12.05372461442061
12 norm of gradients: 0.09831020015438756
12 norm of weights: 2.215924344518654
-----
Iteration Number: 6012
Loss: 12.052811937299033
12 norm of gradients: 0.09828498489378841
12 norm of weights: 2.215894938543044
-----
Iteration Number: 6013
Loss: 12.05189985215986
12 norm of gradients: 0.09825978303697261
12 norm of weights: 2.215865589035238
-----
Iteration Number: 6014
Loss: 12.050988358559087
12 norm of gradients: 0.09823459457439457
12 norm of weights: 2.2158362959535585
-----
Iteration Number: 6015
Loss: 12.050077456053105
12 norm of gradients: 0.09820941949651496
12 norm of weights: 2.2158070592563592
-----
Iteration Number: 6016
Loss: 12.04916714419867
12 norm of gradients: 0.09818425779380047
12 norm of weights: 2.2157778789020206
-----

```

Iteration Number: 6017
Loss: 12.048257422552945
l2 norm of gradients: 0.098159109456724
l2 norm of weights: 2.2157487548489514

Iteration Number: 6018
Loss: 12.047348290673462
l2 norm of gradients: 0.09813397447576458
l2 norm of weights: 2.21571968705559

Iteration Number: 6019
Loss: 12.04643974811814
l2 norm of gradients: 0.0981088528414074
l2 norm of weights: 2.215690675480403

Iteration Number: 6020
Loss: 12.04553179444528
l2 norm of gradients: 0.09808374454414374
l2 norm of weights: 2.215661720081884

Iteration Number: 6021
Loss: 12.044624429213577
l2 norm of gradients: 0.0980586495744711
l2 norm of weights: 2.2156328208185556

Iteration Number: 6022
Loss: 12.043717651982087
l2 norm of gradients: 0.09803356792289304
l2 norm of weights: 2.215603977648969

Iteration Number: 6023
Loss: 12.042811462310272
l2 norm of gradients: 0.09800849957991936
l2 norm of weights: 2.2155751905317045

Iteration Number: 6024
Loss: 12.041905859757954
l2 norm of gradients: 0.09798344453606597
l2 norm of weights: 2.2155464594253687

Iteration Number: 6025
Loss: 12.04100084388535
l2 norm of gradients: 0.09795840278185497
l2 norm of weights: 2.2155177842885982

Iteration Number: 6026
Loss: 12.040096414253053
l2 norm of gradients: 0.09793337430781457

```

12 norm of weights: 2.215489165080056
-----
Iteration Number: 6027
Loss: 12.039192570422042
12 norm of gradients: 0.09790835910447916
12 norm of weights: 2.2154606017584357
-----
Iteration Number: 6028
Loss: 12.038289311953665
12 norm of gradients: 0.09788335716238934
12 norm of weights: 2.215432094282457
-----
Iteration Number: 6029
Loss: 12.03738663840966
12 norm of gradients: 0.0978583684720918
12 norm of weights: 2.2154036426108683
-----
Iteration Number: 6030
Loss: 12.036484549352142
12 norm of gradients: 0.0978333930241395
12 norm of weights: 2.2153752467024477
-----
Iteration Number: 6031
Loss: 12.035583044343603
12 norm of gradients: 0.09780843080909148
12 norm of weights: 2.215346906515999
-----
Iteration Number: 6032
Loss: 12.03468212294691
12 norm of gradients: 0.09778348181751292
12 norm of weights: 2.2153186220103556
-----
Iteration Number: 6033
Loss: 12.033781784725322
12 norm of gradients: 0.09775854603997536
12 norm of weights: 2.215290393144379
-----
Iteration Number: 6034
Loss: 12.032882029242458
12 norm of gradients: 0.09773362346705633
12 norm of weights: 2.215262219876958
-----
Iteration Number: 6035
Loss: 12.03198285606233
12 norm of gradients: 0.09770871408933961
12 norm of weights: 2.2152341021670097
-----
Iteration Number: 6036

```

Loss: 12.031084264749317
l2 norm of gradients: 0.09768381789741512
l2 norm of weights: 2.2152060399734803

Iteration Number: 6037
Loss: 12.030186254868184
l2 norm of gradients: 0.0976589348818791
l2 norm of weights: 2.215178033255343

Iteration Number: 6038
Loss: 12.029288825984061
l2 norm of gradients: 0.09763406503333376
l2 norm of weights: 2.2151500819715984

Iteration Number: 6039
Loss: 12.028391977662467
l2 norm of gradients: 0.0976092083423877
l2 norm of weights: 2.2151221860812766

Iteration Number: 6040
Loss: 12.027495709469283
l2 norm of gradients: 0.09758436479965561
l2 norm of weights: 2.2150943455434344

Iteration Number: 6041
Loss: 12.026600020970777
l2 norm of gradients: 0.09755953439575836
l2 norm of weights: 2.2150665603171578

Iteration Number: 6042
Loss: 12.025704911733586
l2 norm of gradients: 0.09753471712132301
l2 norm of weights: 2.2150388303615594

Iteration Number: 6043
Loss: 12.024810381324723
l2 norm of gradients: 0.09750991296698291
l2 norm of weights: 2.2150111556357803

Iteration Number: 6044
Loss: 12.023916429311578
l2 norm of gradients: 0.09748512192337756
l2 norm of weights: 2.2149835360989902

Iteration Number: 6045
Loss: 12.023023055261907
l2 norm of gradients: 0.09746034398115257
l2 norm of weights: 2.214955971710385

```

-----
Iteration Number: 6046
Loss: 12.022130258743852
l2 norm of gradients: 0.0974355791309598
l2 norm of weights: 2.2149284624291896
-----
Iteration Number: 6047
Loss: 12.02123803932592
l2 norm of gradients: 0.09741082736345746
l2 norm of weights: 2.214901008214657
-----
Iteration Number: 6048
Loss: 12.02034639657699
l2 norm of gradients: 0.09738608866930978
l2 norm of weights: 2.2148736090260672
-----
Iteration Number: 6049
Loss: 12.019455330066316
l2 norm of gradients: 0.09736136303918724
l2 norm of weights: 2.2148462648227287
-----
Iteration Number: 6050
Loss: 12.018564839363524
l2 norm of gradients: 0.09733665046376658
l2 norm of weights: 2.2148189755639773
-----
Iteration Number: 6051
Loss: 12.017674924038612
l2 norm of gradients: 0.09731195093373067
l2 norm of weights: 2.214791741209176
-----
Iteration Number: 6052
Loss: 12.016785583661951
l2 norm of gradients: 0.0972872644397687
l2 norm of weights: 2.2147645617177174
-----
Iteration Number: 6053
Loss: 12.015896817804277
l2 norm of gradients: 0.09726259097257597
l2 norm of weights: 2.2147374370490196
-----
Iteration Number: 6054
Loss: 12.015008626036705
l2 norm of gradients: 0.09723793052285405
l2 norm of weights: 2.2147103671625303
-----
Iteration Number: 6055
Loss: 12.014121007930715

```

```

12 norm of gradients: 0.09721328308131075
12 norm of weights: 2.2146833520177234
-----
Iteration Number: 6056
Loss: 12.013233963058157
12 norm of gradients: 0.09718864863866004
12 norm of weights: 2.2146563915741013
-----
Iteration Number: 6057
Loss: 12.012347490991253
12 norm of gradients: 0.09716402718562216
12 norm of weights: 2.2146294857911935
-----
Iteration Number: 6058
Loss: 12.011461591302592
12 norm of gradients: 0.09713941871292354
12 norm of weights: 2.214602634628558
-----
Iteration Number: 6059
Loss: 12.010576263565133
12 norm of gradients: 0.09711482321129684
12 norm of weights: 2.2145758380457794
-----
Iteration Number: 6060
Loss: 12.009691507352207
12 norm of gradients: 0.09709024067148096
12 norm of weights: 2.2145490960024707
-----
Iteration Number: 6061
Loss: 12.008807322237503
12 norm of gradients: 0.09706567108422101
12 norm of weights: 2.2145224084582718
-----
Iteration Number: 6062
Loss: 12.00792370779509
12 norm of gradients: 0.09704111444026835
12 norm of weights: 2.2144957753728507
-----
Iteration Number: 6063
Loss: 12.007040663599398
12 norm of gradients: 0.09701657073038054
12 norm of weights: 2.2144691967059025
-----
Iteration Number: 6064
Loss: 12.006158189225221
12 norm of gradients: 0.09699203994532145
12 norm of weights: 2.21444267241715
-----

```

Iteration Number: 6065
Loss: 12.005276284247728
l2 norm of gradients: 0.09696752207586104
l2 norm of weights: 2.214416202466344

Iteration Number: 6066
Loss: 12.004394948242451
l2 norm of gradients: 0.09694301711277564
l2 norm of weights: 2.214389786813262

Iteration Number: 6067
Loss: 12.003514180785288
l2 norm of gradients: 0.0969185250468478
l2 norm of weights: 2.2143634254177087

Iteration Number: 6068
Loss: 12.002633981452494
l2 norm of gradients: 0.09689404586886623
l2 norm of weights: 2.2143371182395173

Iteration Number: 6069
Loss: 12.001754349820704
l2 norm of gradients: 0.0968695795696259
l2 norm of weights: 2.2143108652385477

Iteration Number: 6070
Loss: 12.000875285466911
l2 norm of gradients: 0.0968451261399281
l2 norm of weights: 2.214284666374688

Iteration Number: 6071
Loss: 11.999996787968474
l2 norm of gradients: 0.0968206855705803
l2 norm of weights: 2.214258521607852

Iteration Number: 6072
Loss: 11.999118856903104
l2 norm of gradients: 0.0967962578523962
l2 norm of weights: 2.214232430897983

Iteration Number: 6073
Loss: 11.998241491848903
l2 norm of gradients: 0.0967718429761958
l2 norm of weights: 2.21420639420505

Iteration Number: 6074
Loss: 11.997364692384311
l2 norm of gradients: 0.0967474409328053


```

12 norm of weights: 2.2141804114890498
-----
Iteration Number: 6075
Loss: 11.996488458088143
12 norm of gradients: 0.09672305171305715
12 norm of weights: 2.214154482710007
-----
Iteration Number: 6076
Loss: 11.995612788539574
12 norm of gradients: 0.09669867530779008
12 norm of weights: 2.214128607827973
-----
Iteration Number: 6077
Loss: 11.994737683318142
12 norm of gradients: 0.09667431170784904
12 norm of weights: 2.2141027868030267
-----
Iteration Number: 6078
Loss: 11.993863142003747
12 norm of gradients: 0.09664996090408529
12 norm of weights: 2.214077019595274
-----
Iteration Number: 6079
Loss: 11.992989164176652
12 norm of gradients: 0.09662562288735621
12 norm of weights: 2.2140513061648477
-----
Iteration Number: 6080
Loss: 11.99211574941748
12 norm of gradients: 0.09660129764852561
12 norm of weights: 2.2140256464719092
-----
Iteration Number: 6081
Loss: 11.991242897307215
12 norm of gradients: 0.09657698517846346
12 norm of weights: 2.214000040476646
-----
Iteration Number: 6082
Loss: 11.9903706074272
12 norm of gradients: 0.0965526854680459
12 norm of weights: 2.2139744881392724
-----
Iteration Number: 6083
Loss: 11.989498879359143
12 norm of gradients: 0.09652839850815559
12 norm of weights: 2.213948989420031
-----
Iteration Number: 6084

```

Loss: 11.988627712685108
l2 norm of gradients: 0.09650412428968115
l2 norm of weights: 2.2139235442791905

Iteration Number: 6085
Loss: 11.987757106987523
l2 norm of gradients: 0.09647986280351767
l2 norm of weights: 2.213898152677048

Iteration Number: 6086
Loss: 11.98688706184917
l2 norm of gradients: 0.09645561404056634
l2 norm of weights: 2.213872814573926

Iteration Number: 6087
Loss: 11.986017576853191
l2 norm of gradients: 0.09643137799173482
l2 norm of weights: 2.213847529930175

Iteration Number: 6088
Loss: 11.985148651583092
l2 norm of gradients: 0.09640715464793682
l2 norm of weights: 2.213822298706174

Iteration Number: 6089
Loss: 11.98428028562273
l2 norm of gradients: 0.09638294400009241
l2 norm of weights: 2.213797120862326

Iteration Number: 6090
Loss: 11.983412478556325
l2 norm of gradients: 0.096358746039128
l2 norm of weights: 2.2137719963590636

Iteration Number: 6091
Loss: 11.982545229968453
l2 norm of gradients: 0.09633456075597613
l2 norm of weights: 2.213746925156845

Iteration Number: 6092
Loss: 11.98167853944404
l2 norm of gradients: 0.09631038814157564
l2 norm of weights: 2.2137219072161556

Iteration Number: 6093
Loss: 11.980812406568388
l2 norm of gradients: 0.09628622818687173
l2 norm of weights: 2.213696942497509

```

-----
Iteration Number: 6094
Loss: 11.979946830927133
l2 norm of gradients: 0.09626208088281578
l2 norm of weights: 2.213672030961444
-----
Iteration Number: 6095
Loss: 11.979081812106283
l2 norm of gradients: 0.09623794622036547
l2 norm of weights: 2.2136471725685274
-----
Iteration Number: 6096
Loss: 11.978217349692196
l2 norm of gradients: 0.09621382419048483
l2 norm of weights: 2.2136223672793527
-----
Iteration Number: 6097
Loss: 11.977353443271578
l2 norm of gradients: 0.09618971478414393
l2 norm of weights: 2.21359761505454
-----
Iteration Number: 6098
Loss: 11.97649009243151
l2 norm of gradients: 0.09616561799231937
l2 norm of weights: 2.213572915854737
-----
Iteration Number: 6099
Loss: 11.975627296759406
l2 norm of gradients: 0.0961415338059939
l2 norm of weights: 2.2135482696406172
-----
Iteration Number: 6100
Loss: 11.974765055843047
l2 norm of gradients: 0.09611746221615662
l2 norm of weights: 2.2135236763728825
-----
Iteration Number: 6101
Loss: 11.973903369270566
l2 norm of gradients: 0.09609340321380277
l2 norm of weights: 2.2134991360122602
-----
Iteration Number: 6102
Loss: 11.973042236630452
l2 norm of gradients: 0.09606935678993403
l2 norm of weights: 2.2134746485195054
-----
Iteration Number: 6103
Loss: 11.972181657511538

```

```

12 norm of gradients: 0.09604532293555824
12 norm of weights: 2.213450213855399
-----
Iteration Number: 6104
Loss: 11.971321631503015
12 norm of gradients: 0.09602130164168951
12 norm of weights: 2.2134258319807487
-----
Iteration Number: 6105
Loss: 11.970462158194438
12 norm of gradients: 0.09599729289934841
12 norm of weights: 2.213401502856391
-----
Iteration Number: 6106
Loss: 11.969603237175692
12 norm of gradients: 0.09597329669956156
12 norm of weights: 2.2133772264431864
-----
Iteration Number: 6107
Loss: 11.968744868037035
12 norm of gradients: 0.095949313033362
12 norm of weights: 2.213353002702023
-----
Iteration Number: 6108
Loss: 11.967887050369065
12 norm of gradients: 0.09592534189178904
12 norm of weights: 2.2133288315938175
-----
Iteration Number: 6109
Loss: 11.967029783762731
12 norm of gradients: 0.09590138326588817
12 norm of weights: 2.2133047130795105
-----
Iteration Number: 6110
Loss: 11.966173067809343
12 norm of gradients: 0.09587743714671128
12 norm of weights: 2.213280647120071
-----
Iteration Number: 6111
Loss: 11.96531690210055
12 norm of gradients: 0.09585350352531655
12 norm of weights: 2.213256633676494
-----
Iteration Number: 6112
Loss: 11.964461286228353
12 norm of gradients: 0.09582958239276833
12 norm of weights: 2.2132326727098013
-----

```

Iteration Number: 6113
Loss: 11.963606219785113
l2 norm of gradients: 0.09580567374013739
l2 norm of weights: 2.213208764181041

Iteration Number: 6114
Loss: 11.962751702363532
l2 norm of gradients: 0.09578177755850069
l2 norm of weights: 2.213184908051289

Iteration Number: 6115
Loss: 11.961897733556658
l2 norm of gradients: 0.09575789383894151
l2 norm of weights: 2.213161104281646

Iteration Number: 6116
Loss: 11.9610443129579
l2 norm of gradients: 0.09573402257254947
l2 norm of weights: 2.2131373528332405

Iteration Number: 6117
Loss: 11.960191440161006
l2 norm of gradients: 0.09571016375042034
l2 norm of weights: 2.2131136536672273

Iteration Number: 6118
Loss: 11.959339114760068
l2 norm of gradients: 0.0956863173636563
l2 norm of weights: 2.2130900067447876

Iteration Number: 6119
Loss: 11.958487336349537
l2 norm of gradients: 0.09566248340336579
l2 norm of weights: 2.213066412027129

Iteration Number: 6120
Loss: 11.957636104524212
l2 norm of gradients: 0.09563866186066357
l2 norm of weights: 2.213042869475485

Iteration Number: 6121
Loss: 11.956785418879225
l2 norm of gradients: 0.09561485272667057
l2 norm of weights: 2.2130193790511177

Iteration Number: 6122
Loss: 11.955935279010069
l2 norm of gradients: 0.09559105599251415

```
l2 norm of weights: 2.212995940715314
-----
Iteration Number: 6123
Loss: 11.955085684512579
l2 norm of gradients: 0.09556727164932793
l2 norm of weights: 2.2129725544293866
-----
Iteration Number: 6124
Loss: 11.954236634982934
l2 norm of gradients: 0.09554349968825175
l2 norm of weights: 2.212949220154676
-----
Iteration Number: 6125
Loss: 11.953388130017661
l2 norm of gradients: 0.0955197401004318
l2 norm of weights: 2.2129259378525483
-----
Iteration Number: 6126
Loss: 11.95254016921363
l2 norm of gradients: 0.09549599287702056
l2 norm of weights: 2.212902707484397
-----
Iteration Number: 6127
Loss: 11.951692752168064
l2 norm of gradients: 0.09547225800917677
l2 norm of weights: 2.212879529011641
-----
Iteration Number: 6128
Loss: 11.95084587847852
l2 norm of gradients: 0.09544853548806558
l2 norm of weights: 2.2128564023957256
-----
Iteration Number: 6129
Loss: 11.949999547742905
l2 norm of gradients: 0.0954248253048582
l2 norm of weights: 2.2128333275981227
-----
Iteration Number: 6130
Loss: 11.949153759559472
l2 norm of gradients: 0.09540112745073238
l2 norm of weights: 2.21281030458033
-----
Iteration Number: 6131
Loss: 11.948308513526815
l2 norm of gradients: 0.09537744191687207
l2 norm of weights: 2.212787333303873
-----
Iteration Number: 6132
```

Loss: 11.947463809243866
l2 norm of gradients: 0.09535376869446743
l2 norm of weights: 2.2127644137303015

Iteration Number: 6133
Loss: 11.946619646309912
l2 norm of gradients: 0.09533010777471505
l2 norm of weights: 2.212741545821193

Iteration Number: 6134
Loss: 11.94577602432458
l2 norm of gradients: 0.09530645914881777
l2 norm of weights: 2.2127187295381505

Iteration Number: 6135
Loss: 11.944932942887828
l2 norm of gradients: 0.09528282280798465
l2 norm of weights: 2.2126959648428035

Iteration Number: 6136
Loss: 11.944090401599968
l2 norm of gradients: 0.09525919874343115
l2 norm of weights: 2.212673251696808

Iteration Number: 6137
Loss: 11.943248400061652
l2 norm of gradients: 0.09523558694637899
l2 norm of weights: 2.212650590061845

Iteration Number: 6138
Loss: 11.942406937873871
l2 norm of gradients: 0.09521198740805614
l2 norm of weights: 2.2126279798996236

Iteration Number: 6139
Loss: 11.941566014637957
l2 norm of gradients: 0.09518840011969697
l2 norm of weights: 2.212605421171877

Iteration Number: 6140
Loss: 11.940725629955583
l2 norm of gradients: 0.09516482507254205
l2 norm of weights: 2.212582913840366

Iteration Number: 6141
Loss: 11.93988578342876
l2 norm of gradients: 0.0951412622578383
l2 norm of weights: 2.2125604578668767

```

-----
Iteration Number: 6142
Loss: 11.939046474659852
l2 norm of gradients: 0.0951177116668389
l2 norm of weights: 2.212538053213222
-----
Iteration Number: 6143
Loss: 11.938207703251543
l2 norm of gradients: 0.09509417329080333
l2 norm of weights: 2.212515699841241
-----
Iteration Number: 6144
Loss: 11.93736946880687
l2 norm of gradients: 0.09507064712099743
l2 norm of weights: 2.212493397712797
-----
Iteration Number: 6145
Loss: 11.936531770929205
l2 norm of gradients: 0.09504713314869329
l2 norm of weights: 2.212471146789782
-----
Iteration Number: 6146
Loss: 11.93569460922226
l2 norm of gradients: 0.09502363136516925
l2 norm of weights: 2.212448947034112
-----
Iteration Number: 6147
Loss: 11.934857983290085
l2 norm of gradients: 0.09500014176171007
l2 norm of weights: 2.21242679840773
-----
Iteration Number: 6148
Loss: 11.934021892737064
l2 norm of gradients: 0.09497666432960669
l2 norm of weights: 2.2124047008726047
-----
Iteration Number: 6149
Loss: 11.93318633716793
l2 norm of gradients: 0.0949531990601564
l2 norm of weights: 2.2123826543907312
-----
Iteration Number: 6150
Loss: 11.93235131618774
l2 norm of gradients: 0.09492974594466279
l2 norm of weights: 2.21236065892413
-----
Iteration Number: 6151
Loss: 11.931516829401895

```



```

12 norm of gradients: 0.09490630497443574
12 norm of weights: 2.212338714434847
-----
Iteration Number: 6152
Loss: 11.930682876416132
12 norm of gradients: 0.09488287614079143
12 norm of weights: 2.212316820884956
-----
Iteration Number: 6153
Loss: 11.929849456836525
12 norm of gradients: 0.0948594594350523
12 norm of weights: 2.212294978236555
-----
Iteration Number: 6154
Loss: 11.929016570269486
12 norm of gradients: 0.09483605484854722
12 norm of weights: 2.212273186451768
-----
Iteration Number: 6155
Loss: 11.928184216321755
12 norm of gradients: 0.09481266237261113
12 norm of weights: 2.2122514454927455
-----
Iteration Number: 6156
Loss: 11.927352394600415
12 norm of gradients: 0.09478928199858548
12 norm of weights: 2.2122297553216637
-----
Iteration Number: 6157
Loss: 11.926521104712887
12 norm of gradients: 0.09476591371781792
12 norm of weights: 2.2122081159007245
-----
Iteration Number: 6158
Loss: 11.925690346266911
12 norm of gradients: 0.09474255752166247
12 norm of weights: 2.2121865271921557
-----
Iteration Number: 6159
Loss: 11.924860118870578
12 norm of gradients: 0.09471921340147929
12 norm of weights: 2.2121649891582105
-----
Iteration Number: 6160
Loss: 11.924030422132311
12 norm of gradients: 0.09469588134863502
12 norm of weights: 2.212143501761169
-----

```

Iteration Number: 6161
Loss: 11.923201255660858
l2 norm of gradients: 0.09467256135450249
l2 norm of weights: 2.2121220649633355

Iteration Number: 6162
Loss: 11.922372619065309
l2 norm of gradients: 0.09464925341046082
l2 norm of weights: 2.2121006787270416

Iteration Number: 6163
Loss: 11.921544511955078
l2 norm of gradients: 0.09462595750789557
l2 norm of weights: 2.212079343014643

Iteration Number: 6164
Loss: 11.920716933939925
l2 norm of gradients: 0.09460267363819835
l2 norm of weights: 2.212058057788523

Iteration Number: 6165
Loss: 11.91988988462993
l2 norm of gradients: 0.0945794017927673
l2 norm of weights: 2.212036823011089

Iteration Number: 6166
Loss: 11.91906336363551
l2 norm of gradients: 0.0945561419630067
l2 norm of weights: 2.2120156386447754

Iteration Number: 6167
Loss: 11.918237370567416
l2 norm of gradients: 0.09453289414032726
l2 norm of weights: 2.21199450465204

Iteration Number: 6168
Loss: 11.91741190503673
l2 norm of gradients: 0.09450965831614584
l2 norm of weights: 2.2119734209953696

Iteration Number: 6169
Loss: 11.916586966654862
l2 norm of gradients: 0.09448643448188575
l2 norm of weights: 2.211952387637274

Iteration Number: 6170
Loss: 11.915762555033554
l2 norm of gradients: 0.09446322262897643

```

l2 norm of weights: 2.2119314045402896
-----
Iteration Number: 6171
Loss: 11.914938669784881
l2 norm of gradients: 0.09444002274885374
l2 norm of weights: 2.2119104716669784
-----
Iteration Number: 6172
Loss: 11.91411531052124
l2 norm of gradients: 0.09441683483295979
l2 norm of weights: 2.2118895889799273
-----
Iteration Number: 6173
Loss: 11.913292476855377
l2 norm of gradients: 0.094393658872743
l2 norm of weights: 2.2118687564417496
-----
Iteration Number: 6174
Loss: 11.912470168400343
l2 norm of gradients: 0.09437049485965804
l2 norm of weights: 2.2118479740150843
-----
Iteration Number: 6175
Loss: 11.91164838476954
l2 norm of gradients: 0.09434734278516592
l2 norm of weights: 2.211827241662595
-----
Iteration Number: 6176
Loss: 11.91082712557668
l2 norm of gradients: 0.09432420264073393
l2 norm of weights: 2.2118065593469716
-----
Iteration Number: 6177
Loss: 11.910006390435818
l2 norm of gradients: 0.09430107441783565
l2 norm of weights: 2.2117859270309284
-----
Iteration Number: 6178
Loss: 11.909186178961331
l2 norm of gradients: 0.09427795810795093
l2 norm of weights: 2.2117653446772065
-----
Iteration Number: 6179
Loss: 11.908366490767925
l2 norm of gradients: 0.09425485370256598
l2 norm of weights: 2.211744812248572
-----
Iteration Number: 6180

```

```

Loss: 11.90754732547063
l2 norm of gradients: 0.09423176119317325
l2 norm of weights: 2.2117243297078164
-----
Iteration Number: 6181
Loss: 11.906728682684811
l2 norm of gradients: 0.09420868057127149
l2 norm of weights: 2.2117038970177565
-----
Iteration Number: 6182
Loss: 11.905910562026156
l2 norm of gradients: 0.09418561182836568
l2 norm of weights: 2.2116835141412343
-----
Iteration Number: 6183
Loss: 11.905092963110677
l2 norm of gradients: 0.09416255495596722
l2 norm of weights: 2.211663181041118
-----
Iteration Number: 6184
Loss: 11.904275885554714
l2 norm of gradients: 0.09413950994559367
l2 norm of weights: 2.2116428976802998
-----
Iteration Number: 6185
Loss: 11.90345932897494
l2 norm of gradients: 0.09411647678876904
l2 norm of weights: 2.211622664021699
-----
Iteration Number: 6186
Loss: 11.90264329298834
l2 norm of gradients: 0.09409345547702341
l2 norm of weights: 2.211602480028259
-----
Iteration Number: 6187
Loss: 11.901827777212237
l2 norm of gradients: 0.09407044600189331
l2 norm of weights: 2.211582345662949
-----
Iteration Number: 6188
Loss: 11.901012781264269
l2 norm of gradients: 0.09404744835492153
l2 norm of weights: 2.2115622608887624
-----
Iteration Number: 6189
Loss: 11.900198304762409
l2 norm of gradients: 0.0940244625276571
l2 norm of weights: 2.2115422256687203

```

```

-----
Iteration Number: 6190
Loss: 11.899384347324949
l2 norm of gradients: 0.09400148851165538
l2 norm of weights: 2.2115222399658667
-----
Iteration Number: 6191
Loss: 11.8985709085705
l2 norm of gradients: 0.093978526298478
l2 norm of weights: 2.211502303743272
-----
Iteration Number: 6192
Loss: 11.897757988118004
l2 norm of gradients: 0.09395557587969285
l2 norm of weights: 2.2114824169640324
-----
Iteration Number: 6193
Loss: 11.896945585586726
l2 norm of gradients: 0.09393263724687415
l2 norm of weights: 2.211462579591267
-----
Iteration Number: 6194
Loss: 11.896133700596252
l2 norm of gradients: 0.09390971039160241
l2 norm of weights: 2.2114427915881234
-----
Iteration Number: 6195
Loss: 11.89532233276649
l2 norm of gradients: 0.09388679530546434
l2 norm of weights: 2.211423052917771
-----
Iteration Number: 6196
Loss: 11.894511481717677
l2 norm of gradients: 0.09386389198005307
l2 norm of weights: 2.211403363543407
-----
Iteration Number: 6197
Loss: 11.893701147070363
l2 norm of gradients: 0.09384100040696786
l2 norm of weights: 2.2113837234282525
-----
Iteration Number: 6198
Loss: 11.892891328445423
l2 norm of gradients: 0.09381812057781434
l2 norm of weights: 2.211364132535554
-----
Iteration Number: 6199
Loss: 11.892082025464056

```

```

12 norm of gradients: 0.09379525248420441
12 norm of weights: 2.2113445908285834
-----
Iteration Number: 6200
Loss: 11.891273237747779
12 norm of gradients: 0.09377239611775626
12 norm of weights: 2.2113250982706365
-----
Iteration Number: 6201
Loss: 11.890464964918431
12 norm of gradients: 0.0937495514700943
12 norm of weights: 2.2113056548250363
-----
Iteration Number: 6202
Loss: 11.889657206598178
12 norm of gradients: 0.09372671853284928
12 norm of weights: 2.211286260455129
-----
Iteration Number: 6203
Loss: 11.888849962409495
12 norm of gradients: 0.09370389729765823
12 norm of weights: 2.211266915124286
-----
Iteration Number: 6204
Loss: 11.888043231975185
12 norm of gradients: 0.09368108775616435
12 norm of weights: 2.2112476187959054
-----
Iteration Number: 6205
Loss: 11.887237014918366
12 norm of gradients: 0.09365828990001732
12 norm of weights: 2.2112283714334087
-----
Iteration Number: 6206
Loss: 11.886431310862482
12 norm of gradients: 0.09363550372087286
12 norm of weights: 2.211209173000243
-----
Iteration Number: 6207
Loss: 11.885626119431283
12 norm of gradients: 0.09361272921039317
12 norm of weights: 2.2111900234598796
-----
Iteration Number: 6208
Loss: 11.884821440248857
12 norm of gradients: 0.09358996636024655
12 norm of weights: 2.2111709227758163
-----

```

Iteration Number: 6209
Loss: 11.884017272939591
l2 norm of gradients: 0.09356721516210771
l2 norm of weights: 2.2111518709115745

Iteration Number: 6210
Loss: 11.8832136171282
l2 norm of gradients: 0.09354447560765757
l2 norm of weights: 2.211132867830701

Iteration Number: 6211
Loss: 11.882410472439721
l2 norm of gradients: 0.09352174768858325
l2 norm of weights: 2.211113913496768

Iteration Number: 6212
Loss: 11.881607838499496
l2 norm of gradients: 0.09349903139657828
l2 norm of weights: 2.2110950078733715

Iteration Number: 6213
Loss: 11.880805714933194
l2 norm of gradients: 0.09347632672334244
l2 norm of weights: 2.2110761509241335

Iteration Number: 6214
Loss: 11.880004101366799
l2 norm of gradients: 0.0934536336605816
l2 norm of weights: 2.2110573426127003

Iteration Number: 6215
Loss: 11.879202997426608
l2 norm of gradients: 0.09343095220000817
l2 norm of weights: 2.211038582902743

Iteration Number: 6216
Loss: 11.878402402739239
l2 norm of gradients: 0.09340828233334059
l2 norm of weights: 2.2110198717579577

Iteration Number: 6217
Loss: 11.877602316931624
l2 norm of gradients: 0.09338562405230368
l2 norm of weights: 2.2110012091420654

Iteration Number: 6218
Loss: 11.876802739631007
l2 norm of gradients: 0.0933629773486285

```

12 norm of weights: 2.2109825950188116
-----
Iteration Number: 6219
Loss: 11.876003670464952
12 norm of gradients: 0.0933403422140524
12 norm of weights: 2.210964029351967
-----
Iteration Number: 6220
Loss: 11.875205109061334
12 norm of gradients: 0.09331771864031894
12 norm of weights: 2.2109455121053267
-----
Iteration Number: 6221
Loss: 11.874407055048348
12 norm of gradients: 0.09329510661917795
12 norm of weights: 2.210927043242711
-----
Iteration Number: 6222
Loss: 11.873609508054496
12 norm of gradients: 0.09327250614238558
12 norm of weights: 2.210908622727964
-----
Iteration Number: 6223
Loss: 11.872812467708608
12 norm of gradients: 0.09324991720170417
12 norm of weights: 2.210890250524956
-----
Iteration Number: 6224
Loss: 11.872015933639805
12 norm of gradients: 0.09322733978890235
12 norm of weights: 2.2108719265975805
-----
Iteration Number: 6225
Loss: 11.871219905477544
12 norm of gradients: 0.093204773895755
12 norm of weights: 2.210853650909757
-----
Iteration Number: 6226
Loss: 11.87042438285158
12 norm of gradients: 0.09318221951404325
12 norm of weights: 2.210835423425428
-----
Iteration Number: 6227
Loss: 11.869629365391987
12 norm of gradients: 0.09315967663555448
12 norm of weights: 2.2108172441085627
-----
Iteration Number: 6228

```


Loss: 11.868834852729155
l2 norm of gradients: 0.09313714525208235
l2 norm of weights: 2.210799112923153

Iteration Number: 6229
Loss: 11.868040844493782
l2 norm of gradients: 0.09311462535542675
l2 norm of weights: 2.210781029833217

Iteration Number: 6230
Loss: 11.867247340316869
l2 norm of gradients: 0.09309211693739379
l2 norm of weights: 2.2107629948027965

Iteration Number: 6231
Loss: 11.866454339829744
l2 norm of gradients: 0.09306961998979593
l2 norm of weights: 2.2107450077959583

Iteration Number: 6232
Loss: 11.865661842664043
l2 norm of gradients: 0.09304713450445176
l2 norm of weights: 2.2107270687767935

Iteration Number: 6233
Loss: 11.864869848451704
l2 norm of gradients: 0.0930246604731862
l2 norm of weights: 2.2107091777094174

Iteration Number: 6234
Loss: 11.864078356824981
l2 norm of gradients: 0.09300219788783039
l2 norm of weights: 2.2106913345579704

Iteration Number: 6235
Loss: 11.86328736741644
l2 norm of gradients: 0.09297974674022168
l2 norm of weights: 2.210673539286618

Iteration Number: 6236
Loss: 11.862496879858957
l2 norm of gradients: 0.0929573070222037
l2 norm of weights: 2.210655791859549

Iteration Number: 6237
Loss: 11.861706893785716
l2 norm of gradients: 0.09293487872562638
l2 norm of weights: 2.210638092240977

```

-----
Iteration Number: 6238
Loss: 11.860917408830206
l2 norm of gradients: 0.09291246184234574
l2 norm of weights: 2.2106204403951404
-----
Iteration Number: 6239
Loss: 11.860128424626238
l2 norm of gradients: 0.09289005636422416
l2 norm of weights: 2.210602836286302
-----
Iteration Number: 6240
Loss: 11.859339940807915
l2 norm of gradients: 0.09286766228313026
l2 norm of weights: 2.210585279878749
-----
Iteration Number: 6241
Loss: 11.858551957009666
l2 norm of gradients: 0.09284527959093883
l2 norm of weights: 2.2105677711367933
-----
Iteration Number: 6242
Loss: 11.857764472866211
l2 norm of gradients: 0.09282290827953099
l2 norm of weights: 2.21055031002477
-----
Iteration Number: 6243
Loss: 11.856977488012587
l2 norm of gradients: 0.09280054834079389
l2 norm of weights: 2.2105328965070403
-----
Iteration Number: 6244
Loss: 11.856191002084145
l2 norm of gradients: 0.09277819976662124
l2 norm of weights: 2.2105155305479887
-----
Iteration Number: 6245
Loss: 11.855405014716531
l2 norm of gradients: 0.0927558625489127
l2 norm of weights: 2.210498212112024
-----
Iteration Number: 6246
Loss: 11.8546195255457
l2 norm of gradients: 0.09273353667957426
l2 norm of weights: 2.2104809411635795
-----
Iteration Number: 6247
Loss: 11.853834534207925

```

```
12 norm of gradients: 0.09271122215051815
12 norm of weights: 2.210463717667114
-----
Iteration Number: 6248
Loss: 11.85305004033977
12 norm of gradients: 0.09268891895366285
12 norm of weights: 2.2104465415871086
-----
Iteration Number: 6249
Loss: 11.852266043578114
12 norm of gradients: 0.09266662708093296
12 norm of weights: 2.2104294128880704
-----
Iteration Number: 6250
Loss: 11.851482543560142
12 norm of gradients: 0.09264434652425951
12 norm of weights: 2.2104123315345294
-----
Iteration Number: 6251
Loss: 11.850699539923337
12 norm of gradients: 0.09262207727557946
12 norm of weights: 2.210395297491041
-----
Iteration Number: 6252
Loss: 11.849917032305498
12 norm of gradients: 0.09259981932683625
12 norm of weights: 2.210378310722184
-----
Iteration Number: 6253
Loss: 11.849135020344725
12 norm of gradients: 0.09257757266997946
12 norm of weights: 2.210361371192562
-----
Iteration Number: 6254
Loss: 11.848353503679412
12 norm of gradients: 0.0925553372969648
12 norm of weights: 2.2103444788668027
-----
Iteration Number: 6255
Loss: 11.847572481948275
12 norm of gradients: 0.09253311319975432
12 norm of weights: 2.2103276337095576
-----
Iteration Number: 6256
Loss: 11.846791954790321
12 norm of gradients: 0.09251090037031628
12 norm of weights: 2.2103108356855024
-----
```

Iteration Number: 6257
Loss: 11.84601192184487
l2 norm of gradients: 0.092488698800625
l2 norm of weights: 2.2102940847593384

Iteration Number: 6258
Loss: 11.845232382751533
l2 norm of gradients: 0.09246650848266129
l2 norm of weights: 2.2102773808957887

Iteration Number: 6259
Loss: 11.844453337150236
l2 norm of gradients: 0.09244432940841182
l2 norm of weights: 2.2102607240596024

Iteration Number: 6260
Loss: 11.843674784681202
l2 norm of gradients: 0.09242216156986978
l2 norm of weights: 2.210244114215552

Iteration Number: 6261
Loss: 11.842896724984955
l2 norm of gradients: 0.09240000495903442
l2 norm of weights: 2.2102275513284333

Iteration Number: 6262
Loss: 11.842119157702328
l2 norm of gradients: 0.09237785956791122
l2 norm of weights: 2.210211035363068

Iteration Number: 6263
Loss: 11.84134208247445
l2 norm of gradients: 0.0923557253885119
l2 norm of weights: 2.2101945662843003

Iteration Number: 6264
Loss: 11.840565498942752
l2 norm of gradients: 0.09233360241285431
l2 norm of weights: 2.210178144057

Iteration Number: 6265
Loss: 11.839789406748968
l2 norm of gradients: 0.09231149063296255
l2 norm of weights: 2.210161768646059

Iteration Number: 6266
Loss: 11.839013805535133
l2 norm of gradients: 0.0922893900408669

```

12 norm of weights: 2.210145440016394
-----
Iteration Number: 6267
Loss: 11.838238694943586
12 norm of gradients: 0.09226730062860389
12 norm of weights: 2.2101291581329465
-----
Iteration Number: 6268
Loss: 11.837464074616953
12 norm of gradients: 0.09224522238821624
12 norm of weights: 2.2101129229606813
-----
Iteration Number: 6269
Loss: 11.83668994419818
12 norm of gradients: 0.09222315531175275
12 norm of weights: 2.210096734464587
-----
Iteration Number: 6270
Loss: 11.835916303330496
12 norm of gradients: 0.09220109939126861
12 norm of weights: 2.210080592609677
-----
Iteration Number: 6271
Loss: 11.835143151657434
12 norm of gradients: 0.09217905461882502
12 norm of weights: 2.210064497360987
-----
Iteration Number: 6272
Loss: 11.834370488822838
12 norm of gradients: 0.09215702098648941
12 norm of weights: 2.2100484486835787
-----
Iteration Number: 6273
Loss: 11.83359831447083
12 norm of gradients: 0.09213499848633554
12 norm of weights: 2.2100324465425363
-----
Iteration Number: 6274
Loss: 11.832826628245847
12 norm of gradients: 0.09211298711044316
12 norm of weights: 2.210016490902968
-----
Iteration Number: 6275
Loss: 11.832055429792618
12 norm of gradients: 0.09209098685089837
12 norm of weights: 2.210000581730007
-----
Iteration Number: 6276

```

Loss: 11.831284718756173
l2 norm of gradients: 0.09206899769979335
l2 norm of weights: 2.209984718988808

Iteration Number: 6277
Loss: 11.830514494781836
l2 norm of gradients: 0.09204701964922644
l2 norm of weights: 2.2099689026445524

Iteration Number: 6278
Loss: 11.829744757515229
l2 norm of gradients: 0.0920250526913023
l2 norm of weights: 2.2099531326624438

Iteration Number: 6279
Loss: 11.828975506602275
l2 norm of gradients: 0.09200309681813164
l2 norm of weights: 2.209937409007709

Iteration Number: 6280
Loss: 11.82820674168919
l2 norm of gradients: 0.09198115202183137
l2 norm of weights: 2.209921731645601

Iteration Number: 6281
Loss: 11.827438462422487
l2 norm of gradients: 0.09195921829452464
l2 norm of weights: 2.2099061005413936

Iteration Number: 6282
Loss: 11.826670668448974
l2 norm of gradients: 0.09193729562834069
l2 norm of weights: 2.209890515660387

Iteration Number: 6283
Loss: 11.825903359415763
l2 norm of gradients: 0.091915384015415
l2 norm of weights: 2.209874976967903

Iteration Number: 6284
Loss: 11.825136534970252
l2 norm of gradients: 0.09189348344788913
l2 norm of weights: 2.209859484429289

Iteration Number: 6285
Loss: 11.824370194760139
l2 norm of gradients: 0.09187159391791096
l2 norm of weights: 2.2098440380099142

```

-----
Iteration Number: 6286
Loss: 11.823604338433414
l2 norm of gradients: 0.09184971541763433
l2 norm of weights: 2.2098286376751735
-----
Iteration Number: 6287
Loss: 11.822838965638363
l2 norm of gradients: 0.09182784793921944
l2 norm of weights: 2.2098132833904844
-----
Iteration Number: 6288
Loss: 11.822074076023574
l2 norm of gradients: 0.09180599147483252
l2 norm of weights: 2.2097979751212873
-----
Iteration Number: 6289
Loss: 11.821309669237916
l2 norm of gradients: 0.09178414601664604
l2 norm of weights: 2.2097827128330483
-----
Iteration Number: 6290
Loss: 11.820545744930564
l2 norm of gradients: 0.0917623115568386
l2 norm of weights: 2.209767496491255
-----
Iteration Number: 6291
Loss: 11.819782302750978
l2 norm of gradients: 0.09174048808759493
l2 norm of weights: 2.2097523260614205
-----
Iteration Number: 6292
Loss: 11.819019342348913
l2 norm of gradients: 0.09171867560110594
l2 norm of weights: 2.20973720150908
-----
Iteration Number: 6293
Loss: 11.818256863374422
l2 norm of gradients: 0.0916968740895687
l2 norm of weights: 2.209722122799793
-----
Iteration Number: 6294
Loss: 11.817494865477846
l2 norm of gradients: 0.09167508354518643
l2 norm of weights: 2.2097070898991427
-----
Iteration Number: 6295
Loss: 11.816733348309823

```

```

12 norm of gradients: 0.09165330396016842
12 norm of weights: 2.2096921027727356
-----
Iteration Number: 6296
Loss: 11.815972311521278
12 norm of gradients: 0.0916315353267303
12 norm of weights: 2.209677161386202
-----
Iteration Number: 6297
Loss: 11.815211754763428
12 norm of gradients: 0.09160977763709365
12 norm of weights: 2.2096622657051945
-----
Iteration Number: 6298
Loss: 11.814451677687785
12 norm of gradients: 0.09158803088348619
12 norm of weights: 2.209647415695391
-----
Iteration Number: 6299
Loss: 11.813692079946152
12 norm of gradients: 0.09156629505814202
12 norm of weights: 2.209632611322492
-----
Iteration Number: 6300
Loss: 11.812932961190626
12 norm of gradients: 0.09154457015330109
12 norm of weights: 2.209617852552222
-----
Iteration Number: 6301
Loss: 11.812174321073586
12 norm of gradients: 0.09152285616120964
12 norm of weights: 2.2096031393503277
-----
Iteration Number: 6302
Loss: 11.811416159247706
12 norm of gradients: 0.09150115307412
12 norm of weights: 2.209588471682581
-----
Iteration Number: 6303
Loss: 11.810658475365953
12 norm of gradients: 0.09147946088429072
12 norm of weights: 2.2095738495147756
-----
Iteration Number: 6304
Loss: 11.809901269081582
12 norm of gradients: 0.09145777958398636
12 norm of weights: 2.2095592728127302
-----

```


Iteration Number: 6305
Loss: 11.809144540048132
l2 norm of gradients: 0.09143610916547765
l2 norm of weights: 2.209544741542286

Iteration Number: 6306
Loss: 11.808388287919444
l2 norm of gradients: 0.09141444962104146
l2 norm of weights: 2.2095302556693066

Iteration Number: 6307
Loss: 11.807632512349636
l2 norm of gradients: 0.09139280094296076
l2 norm of weights: 2.209515815159681

Iteration Number: 6308
Loss: 11.806877212993124
l2 norm of gradients: 0.09137116312352468
l2 norm of weights: 2.209501419979321

Iteration Number: 6309
Loss: 11.8061223895046
l2 norm of gradients: 0.0913495361550285
l2 norm of weights: 2.20948707009416

Iteration Number: 6310
Loss: 11.805368041539056
l2 norm of gradients: 0.09132792002977352
l2 norm of weights: 2.209472765470157

Iteration Number: 6311
Loss: 11.80461416875177
l2 norm of gradients: 0.0913063147400672
l2 norm of weights: 2.209458506073293

Iteration Number: 6312
Loss: 11.803860770798302
l2 norm of gradients: 0.09128472027822315
l2 norm of weights: 2.209444291869573

Iteration Number: 6313
Loss: 11.803107847334507
l2 norm of gradients: 0.09126313663656103
l2 norm of weights: 2.209430122825025

Iteration Number: 6314
Loss: 11.80235539801652
l2 norm of gradients: 0.09124156380740667

```

12 norm of weights: 2.2094159989057003
-----
Iteration Number: 6315
Loss: 11.801603422500767
12 norm of gradients: 0.09122000178309195
12 norm of weights: 2.209401920077673
-----
Iteration Number: 6316
Loss: 11.800851920443957
12 norm of gradients: 0.09119845055595495
12 norm of weights: 2.20938788630704
-----
Iteration Number: 6317
Loss: 11.80010089150309
12 norm of gradients: 0.09117691011833974
12 norm of weights: 2.2093738975599244
-----
Iteration Number: 6318
Loss: 11.79935033533545
12 norm of gradients: 0.0911553804625965
12 norm of weights: 2.2093599538024686
-----
Iteration Number: 6319
Loss: 11.798600251598605
12 norm of gradients: 0.09113386158108158
12 norm of weights: 2.2093460550008404
-----
Iteration Number: 6320
Loss: 11.79785063995041
12 norm of gradients: 0.09111235346615744
12 norm of weights: 2.2093322011212306
-----
Iteration Number: 6321
Loss: 11.797101500049
12 norm of gradients: 0.09109085611019253
12 norm of weights: 2.2093183921298527
-----
Iteration Number: 6322
Loss: 11.796352831552804
12 norm of gradients: 0.09106936950556152
12 norm of weights: 2.209304627992943
-----
Iteration Number: 6323
Loss: 11.795604634120535
12 norm of gradients: 0.09104789364464506
12 norm of weights: 2.209290908676762
-----
Iteration Number: 6324

```

Loss: 11.79485690741118
l2 norm of gradients: 0.09102642851982991
l2 norm of weights: 2.2092772341475926

Iteration Number: 6325
Loss: 11.794109651084016
l2 norm of gradients: 0.09100497412350898
l2 norm of weights: 2.209263604371741

Iteration Number: 6326
Loss: 11.793362864798608
l2 norm of gradients: 0.09098353044808116
l2 norm of weights: 2.209250019315536

Iteration Number: 6327
Loss: 11.792616548214795
l2 norm of gradients: 0.09096209748595153
l2 norm of weights: 2.2092364789453303

Iteration Number: 6328
Loss: 11.791870700992707
l2 norm of gradients: 0.0909406752295312
l2 norm of weights: 2.209222983227499

Iteration Number: 6329
Loss: 11.791125322792755
l2 norm of gradients: 0.09091926367123732
l2 norm of weights: 2.209209532128441

Iteration Number: 6330
Loss: 11.790380413275628
l2 norm of gradients: 0.09089786280349316
l2 norm of weights: 2.2091961256145765

Iteration Number: 6331
Loss: 11.78963597210231
l2 norm of gradients: 0.09087647261872805
l2 norm of weights: 2.2091827636523513

Iteration Number: 6332
Loss: 11.78889199893405
l2 norm of gradients: 0.09085509310937741
l2 norm of weights: 2.2091694462082314

Iteration Number: 6333
Loss: 11.788148493432391
l2 norm of gradients: 0.0908337242678827
l2 norm of weights: 2.2091561732487084

```

-----
Iteration Number: 6334
Loss: 11.787405455259151
l2 norm of gradients: 0.09081236608669142
l2 norm of weights: 2.209142944740295
-----
Iteration Number: 6335
Loss: 11.786662884076433
l2 norm of gradients: 0.09079101855825718
l2 norm of weights: 2.2091297606495273
-----
Iteration Number: 6336
Loss: 11.785920779546618
l2 norm of gradients: 0.09076968167503965
l2 norm of weights: 2.209116620942965
-----
Iteration Number: 6337
Loss: 11.785179141332371
l2 norm of gradients: 0.09074835542950453
l2 norm of weights: 2.20910352558719
-----
Iteration Number: 6338
Loss: 11.784437969096633
l2 norm of gradients: 0.09072703981412358
l2 norm of weights: 2.2090904745488067
-----
Iteration Number: 6339
Loss: 11.783697262502633
l2 norm of gradients: 0.09070573482137466
l2 norm of weights: 2.209077467794444
-----
Iteration Number: 6340
Loss: 11.782957021213864
l2 norm of gradients: 0.09068444044374158
l2 norm of weights: 2.209064505290752
-----
Iteration Number: 6341
Loss: 11.78221724489412
l2 norm of gradients: 0.0906631566737143
l2 norm of weights: 2.2090515870044043
-----
Iteration Number: 6342
Loss: 11.781477933207459
l2 norm of gradients: 0.09064188350378875
l2 norm of weights: 2.2090387129020974
-----
Iteration Number: 6343
Loss: 11.780739085818222

```

```

12 norm of gradients: 0.09062062092646696
12 norm of weights: 2.209025882950551
-----
Iteration Number: 6344
Loss: 11.780000702391026
12 norm of gradients: 0.09059936893425692
12 norm of weights: 2.2090130971165065
-----
Iteration Number: 6345
Loss: 11.779262782590774
12 norm of gradients: 0.09057812751967284
12 norm of weights: 2.2090003553667295
-----
Iteration Number: 6346
Loss: 11.77852532608264
12 norm of gradients: 0.09055689667523474
12 norm of weights: 2.2089876576680068
-----
Iteration Number: 6347
Loss: 11.77778833253208
12 norm of gradients: 0.09053567639346878
12 norm of weights: 2.2089750039871494
-----
Iteration Number: 6348
Loss: 11.77705180160482
12 norm of gradients: 0.0905144666669072
12 norm of weights: 2.20896239429099
-----
Iteration Number: 6349
Loss: 11.776315732966879
12 norm of gradients: 0.09049326748808818
12 norm of weights: 2.2089498285463853
-----
Iteration Number: 6350
Loss: 11.775580126284538
12 norm of gradients: 0.09047207884955596
12 norm of weights: 2.208937306720213
-----
Iteration Number: 6351
Loss: 11.77484498122436
12 norm of gradients: 0.09045090074386081
12 norm of weights: 2.208924828779375
-----
Iteration Number: 6352
Loss: 11.774110297453188
12 norm of gradients: 0.090429733163559
12 norm of weights: 2.208912394690795
-----

```

Iteration Number: 6353
Loss: 11.773376074638133
l2 norm of gradients: 0.09040857610121286
l2 norm of weights: 2.2089000044214195

Iteration Number: 6354
Loss: 11.772642312446592
l2 norm of gradients: 0.09038742954939068
l2 norm of weights: 2.2088876579382184

Iteration Number: 6355
Loss: 11.77190901054623
l2 norm of gradients: 0.09036629350066681
l2 norm of weights: 2.2088753552081832

Iteration Number: 6356
Loss: 11.771176168604985
l2 norm of gradients: 0.09034516794762158
l2 norm of weights: 2.2088630961983284

Iteration Number: 6357
Loss: 11.770443786291088
l2 norm of gradients: 0.09032405288284136
l2 norm of weights: 2.208850880875692

Iteration Number: 6358
Loss: 11.76971186327302
l2 norm of gradients: 0.0903029482989185
l2 norm of weights: 2.208838709207333

Iteration Number: 6359
Loss: 11.768980399219558
l2 norm of gradients: 0.09028185418845137
l2 norm of weights: 2.2088265811603343

Iteration Number: 6360
Loss: 11.768249393799735
l2 norm of gradients: 0.09026077054404427
l2 norm of weights: 2.2088144967018

Iteration Number: 6361
Loss: 11.767518846682876
l2 norm of gradients: 0.09023969735830764
l2 norm of weights: 2.208802455798859

Iteration Number: 6362
Loss: 11.766788757538567
l2 norm of gradients: 0.09021863462385778

```

12 norm of weights: 2.20879045841866
-----
Iteration Number: 6363
Loss: 11.766059126036668
12 norm of gradients: 0.09019758233331704
12 norm of weights: 2.208778504528376
-----
Iteration Number: 6364
Loss: 11.765329951847328
12 norm of gradients: 0.0901765404793138
12 norm of weights: 2.2087665940952026
-----
Iteration Number: 6365
Loss: 11.764601234640942
12 norm of gradients: 0.09015550905448233
12 norm of weights: 2.208754727086357
-----
Iteration Number: 6366
Loss: 11.763872974088203
12 norm of gradients: 0.09013448805146297
12 norm of weights: 2.208742903469079
-----
Iteration Number: 6367
Loss: 11.76314516986006
12 norm of gradients: 0.09011347746290199
12 norm of weights: 2.2087311232106317
-----
Iteration Number: 6368
Loss: 11.762417821627748
12 norm of gradients: 0.0900924772814517
12 norm of weights: 2.2087193862783
-----
Iteration Number: 6369
Loss: 11.761690929062757
12 norm of gradients: 0.09007148749977029
12 norm of weights: 2.2087076926393907
-----
Iteration Number: 6370
Loss: 11.760964491836868
12 norm of gradients: 0.09005050811052202
12 norm of weights: 2.2086960422612343
-----
Iteration Number: 6371
Loss: 11.760238509622113
12 norm of gradients: 0.09002953910637707
12 norm of weights: 2.208684435111182
-----
Iteration Number: 6372

```

```

Loss: 11.759512982090815
l2 norm of gradients: 0.0900085804800116
l2 norm of weights: 2.20867287115661
-----
Iteration Number: 6373
Loss: 11.75878790891555
l2 norm of gradients: 0.08998763222410776
l2 norm of weights: 2.208661350364914
-----
Iteration Number: 6374
Loss: 11.75806328976918
l2 norm of gradients: 0.08996669433135362
l2 norm of weights: 2.208649872703514
-----
Iteration Number: 6375
Loss: 11.757339124324826
l2 norm of gradients: 0.08994576679444322
l2 norm of weights: 2.208638438139852
-----
Iteration Number: 6376
Loss: 11.756615412255888
l2 norm of gradients: 0.08992484960607657
l2 norm of weights: 2.2086270466413906
-----
Iteration Number: 6377
Loss: 11.755892153236028
l2 norm of gradients: 0.08990394275895967
l2 norm of weights: 2.2086156981756178
-----
Iteration Number: 6378
Loss: 11.75516934693918
l2 norm of gradients: 0.08988304624580445
l2 norm of weights: 2.2086043927100416
-----
Iteration Number: 6379
Loss: 11.754446993039553
l2 norm of gradients: 0.08986216005932873
l2 norm of weights: 2.2085931302121926
-----
Iteration Number: 6380
Loss: 11.75372509121161
l2 norm of gradients: 0.0898412841922563
l2 norm of weights: 2.2085819106496247
-----
Iteration Number: 6381
Loss: 11.753003641130109
l2 norm of gradients: 0.089820418637317
l2 norm of weights: 2.208570733989913

```



```

-----
Iteration Number: 6382
Loss: 11.752282642470044
l2 norm of gradients: 0.08979956338724646
l2 norm of weights: 2.2085596002006556
-----
Iteration Number: 6383
Loss: 11.751562094906706
l2 norm of gradients: 0.08977871843478637
l2 norm of weights: 2.2085485092494714
-----
Iteration Number: 6384
Loss: 11.750841998115636
l2 norm of gradients: 0.0897578837726843
l2 norm of weights: 2.2085374611040036
-----
Iteration Number: 6385
Loss: 11.750122351772648
l2 norm of gradients: 0.08973705939369375
l2 norm of weights: 2.2085264557319166
-----
Iteration Number: 6386
Loss: 11.749403155553823
l2 norm of gradients: 0.08971624529057416
l2 norm of weights: 2.2085154931008963
-----
Iteration Number: 6387
Loss: 11.748684409135514
l2 norm of gradients: 0.08969544145609086
l2 norm of weights: 2.2085045731786512
-----
Iteration Number: 6388
Loss: 11.747966112194334
l2 norm of gradients: 0.08967464788301523
l2 norm of weights: 2.2084936959329133
-----
Iteration Number: 6389
Loss: 11.747248264407167
l2 norm of gradients: 0.08965386456412439
l2 norm of weights: 2.208482861331435
-----
Iteration Number: 6390
Loss: 11.746530865451163
l2 norm of gradients: 0.08963309149220151
l2 norm of weights: 2.2084720693419917
-----
Iteration Number: 6391
Loss: 11.745813915003733

```

```

12 norm of gradients: 0.08961232866003564
12 norm of weights: 2.2084613199323804
-----
Iteration Number: 6392
Loss: 11.74509741274256
12 norm of gradients: 0.08959157606042176
12 norm of weights: 2.2084506130704202
-----
Iteration Number: 6393
Loss: 11.744381358345594
12 norm of gradients: 0.08957083368616073
12 norm of weights: 2.208439948723953
-----
Iteration Number: 6394
Loss: 11.743665751491044
12 norm of gradients: 0.08955010153005931
12 norm of weights: 2.2084293268608417
-----
Iteration Number: 6395
Loss: 11.742950591857385
12 norm of gradients: 0.08952937958493022
12 norm of weights: 2.208418747448973
-----
Iteration Number: 6396
Loss: 11.742235879123363
12 norm of gradients: 0.08950866784359203
12 norm of weights: 2.2084082104562532
-----
Iteration Number: 6397
Loss: 11.741521612967981
12 norm of gradients: 0.08948796629886925
12 norm of weights: 2.208397715850613
-----
Iteration Number: 6398
Loss: 11.74080779307051
12 norm of gradients: 0.0894672749435922
12 norm of weights: 2.2083872636000037
-----
Iteration Number: 6399
Loss: 11.74009441911049
12 norm of gradients: 0.08944659377059724
12 norm of weights: 2.208376853672399
-----
Iteration Number: 6400
Loss: 11.739381490767713
12 norm of gradients: 0.08942592277272651
12 norm of weights: 2.2083664860357937
-----

```

Iteration Number: 6401
Loss: 11.738669007722246
l2 norm of gradients: 0.08940526194282804
l2 norm of weights: 2.2083561606582065

Iteration Number: 6402
Loss: 11.737956969654407
l2 norm of gradients: 0.08938461127375581
l2 norm of weights: 2.2083458775076763

Iteration Number: 6403
Loss: 11.737245376244795
l2 norm of gradients: 0.08936397075836956
l2 norm of weights: 2.2083356365522646

Iteration Number: 6404
Loss: 11.736534227174257
l2 norm of gradients: 0.08934334038953508
l2 norm of weights: 2.2083254377600556

Iteration Number: 6405
Loss: 11.735823522123903
l2 norm of gradients: 0.08932272016012394
l2 norm of weights: 2.2083152810991535

Iteration Number: 6406
Loss: 11.735113260775114
l2 norm of gradients: 0.08930211006301353
l2 norm of weights: 2.2083051665376865

Iteration Number: 6407
Loss: 11.734403442809523
l2 norm of gradients: 0.08928151009108723
l2 norm of weights: 2.2082950940438026

Iteration Number: 6408
Loss: 11.733694067909033
l2 norm of gradients: 0.0892609202372342
l2 norm of weights: 2.208285063585674

Iteration Number: 6409
Loss: 11.732985135755806
l2 norm of gradients: 0.08924034049434948
l2 norm of weights: 2.2082750751314926

Iteration Number: 6410
Loss: 11.732276646032265
l2 norm of gradients: 0.08921977085533402

```

12 norm of weights: 2.208265128649473
-----
Iteration Number: 6411
Loss: 11.73156859842109
12 norm of gradients: 0.08919921131309454
12 norm of weights: 2.2082552241078526
-----
Iteration Number: 6412
Loss: 11.730860992605226
12 norm of gradients: 0.08917866186054373
12 norm of weights: 2.208245361474889
-----
Iteration Number: 6413
Loss: 11.73015382826788
12 norm of gradients: 0.08915812249059997
12 norm of weights: 2.208235540718862
-----
Iteration Number: 6414
Loss: 11.729447105092516
12 norm of gradients: 0.08913759319618769
12 norm of weights: 2.208225761808074
-----
Iteration Number: 6415
Loss: 11.728740822762859
12 norm of gradients: 0.08911707397023699
12 norm of weights: 2.2082160247108487
-----
Iteration Number: 6416
Loss: 11.728034980962889
12 norm of gradients: 0.08909656480568395
12 norm of weights: 2.208206329395531
-----
Iteration Number: 6417
Loss: 11.727329579376855
12 norm of gradients: 0.08907606569547036
12 norm of weights: 2.2081966758304885
-----
Iteration Number: 6418
Loss: 11.726624617689257
12 norm of gradients: 0.08905557663254393
12 norm of weights: 2.20818706398411
-----
Iteration Number: 6419
Loss: 11.725920095584861
12 norm of gradients: 0.08903509760985824
12 norm of weights: 2.2081774938248055
-----
Iteration Number: 6420

```

Loss: 11.725216012748685
l2 norm of gradients: 0.08901462862037257
l2 norm of weights: 2.208167965321008

Iteration Number: 6421
Loss: 11.724512368866007
l2 norm of gradients: 0.08899416965705216
l2 norm of weights: 2.208158478441171

Iteration Number: 6422
Loss: 11.72380916362237
l2 norm of gradients: 0.088973720712868
l2 norm of weights: 2.20814903315377

Iteration Number: 6423
Loss: 11.723106396703562
l2 norm of gradients: 0.08895328178079696
l2 norm of weights: 2.208139629427303

Iteration Number: 6424
Loss: 11.72240406779564
l2 norm of gradients: 0.08893285285382162
l2 norm of weights: 2.2081302672302874

Iteration Number: 6425
Loss: 11.721702176584909
l2 norm of gradients: 0.08891243392493052
l2 norm of weights: 2.208120946531265

Iteration Number: 6426
Loss: 11.721000722757948
l2 norm of gradients: 0.08889202498711786
l2 norm of weights: 2.2081116672987977

Iteration Number: 6427
Loss: 11.720299706001573
l2 norm of gradients: 0.0888716260333838
l2 norm of weights: 2.2081024295014693

Iteration Number: 6428
Loss: 11.71959912600286
l2 norm of gradients: 0.08885123705673423
l2 norm of weights: 2.2080932331078844

Iteration Number: 6429
Loss: 11.71889898244916
l2 norm of gradients: 0.08883085805018082
l2 norm of weights: 2.208084078086671

```
-----
Iteration Number: 6430
Loss: 11.718199275028056
l2 norm of gradients: 0.08881048900674107
l2 norm of weights: 2.2080749644064763
-----
Iteration Number: 6431
Loss: 11.717500003427402
l2 norm of gradients: 0.0887901299194383
l2 norm of weights: 2.2080658920359717
-----
Iteration Number: 6432
Loss: 11.716801167335303
l2 norm of gradients: 0.08876978078130157
l2 norm of weights: 2.2080568609438473
-----
Iteration Number: 6433
Loss: 11.716102766440118
l2 norm of gradients: 0.08874944158536581
l2 norm of weights: 2.2080478710988167
-----
Iteration Number: 6434
Loss: 11.715404800430465
l2 norm of gradients: 0.08872911232467164
l2 norm of weights: 2.2080389224696146
-----
Iteration Number: 6435
Loss: 11.714707268995214
l2 norm of gradients: 0.08870879299226556
l2 norm of weights: 2.2080300150249967
-----
Iteration Number: 6436
Loss: 11.714010171823487
l2 norm of gradients: 0.08868848358119975
l2 norm of weights: 2.20802114873374
-----
Iteration Number: 6437
Loss: 11.713313508604672
l2 norm of gradients: 0.08866818408453228
l2 norm of weights: 2.2080123235646445
-----
Iteration Number: 6438
Loss: 11.712617279028397
l2 norm of gradients: 0.08864789449532691
l2 norm of weights: 2.20800353948653
-----
Iteration Number: 6439
Loss: 11.71192148278455
```

```

12 norm of gradients: 0.08862761480665321
12 norm of weights: 2.2079947964682383
-----
Iteration Number: 6440
Loss: 11.711226119563275
12 norm of gradients: 0.08860734501158646
12 norm of weights: 2.2079860944786325
-----
Iteration Number: 6441
Loss: 11.710531189054965
12 norm of gradients: 0.08858708510320785
12 norm of weights: 2.2079774334865974
-----
Iteration Number: 6442
Loss: 11.709836690950269
12 norm of gradients: 0.08856683507460418
12 norm of weights: 2.207968813461039
-----
Iteration Number: 6443
Loss: 11.709142624940089
12 norm of gradients: 0.08854659491886807
12 norm of weights: 2.2079602343708844
-----
Iteration Number: 6444
Loss: 11.708448990715578
12 norm of gradients: 0.0885263646290979
12 norm of weights: 2.2079516961850825
-----
Iteration Number: 6445
Loss: 11.707755787968141
12 norm of gradients: 0.08850614419839778
12 norm of weights: 2.2079431988726035
-----
Iteration Number: 6446
Loss: 11.70706301638944
12 norm of gradients: 0.0884859336198776
12 norm of weights: 2.207934742402438
-----
Iteration Number: 6447
Loss: 11.706370675671378
12 norm of gradients: 0.08846573288665298
12 norm of weights: 2.2079263267436
-----
Iteration Number: 6448
Loss: 11.705678765506129
12 norm of gradients: 0.0884455419918453
12 norm of weights: 2.2079179518651224
-----

```

Iteration Number: 6449
Loss: 11.704987285586098
l2 norm of gradients: 0.08842536092858165
l2 norm of weights: 2.207909617736061

Iteration Number: 6450
Loss: 11.704296235603948
l2 norm of gradients: 0.08840518968999489
l2 norm of weights: 2.207901324325492

Iteration Number: 6451
Loss: 11.703605615252606
l2 norm of gradients: 0.08838502826922355
l2 norm of weights: 2.2078930716025136

Iteration Number: 6452
Loss: 11.70291542422523
l2 norm of gradients: 0.088364876659412
l2 norm of weights: 2.2078848595362452

Iteration Number: 6453
Loss: 11.702225662215234
l2 norm of gradients: 0.08834473485371025
l2 norm of weights: 2.207876688095826

Iteration Number: 6454
Loss: 11.701536328916292
l2 norm of gradients: 0.08832460284527405
l2 norm of weights: 2.2078685572504186

Iteration Number: 6455
Loss: 11.700847424022314
l2 norm of gradients: 0.08830448062726486
l2 norm of weights: 2.2078604669692052

Iteration Number: 6456
Loss: 11.700158947227479
l2 norm of gradients: 0.0882843681928499
l2 norm of weights: 2.20785241722139

Iteration Number: 6457
Loss: 11.69947089822619
l2 norm of gradients: 0.08826426553520207
l2 norm of weights: 2.2078444079761974

Iteration Number: 6458
Loss: 11.698783276713124
l2 norm of gradients: 0.08824417264750002


```

12 norm of weights: 2.207836439202875
-----
Iteration Number: 6459
Loss: 11.698096082383191
12 norm of gradients: 0.08822408952292808
12 norm of weights: 2.207828510870689
-----
Iteration Number: 6460
Loss: 11.697409314931553
12 norm of gradients: 0.08820401615467623
12 norm of weights: 2.2078206229489283
-----
Iteration Number: 6461
Loss: 11.696722974053621
12 norm of gradients: 0.08818395253594019
12 norm of weights: 2.2078127754069024
-----
Iteration Number: 6462
Loss: 11.696037059445057
12 norm of gradients: 0.08816389865992148
12 norm of weights: 2.207804968213943
-----
Iteration Number: 6463
Loss: 11.695351570801773
12 norm of gradients: 0.08814385451982716
12 norm of weights: 2.2077972013394005
-----
Iteration Number: 6464
Loss: 11.694666507819923
12 norm of gradients: 0.08812382010887004
12 norm of weights: 2.2077894747526488
-----
Iteration Number: 6465
Loss: 11.693981870195909
12 norm of gradients: 0.08810379542026868
12 norm of weights: 2.207781788423082
-----
Iteration Number: 6466
Loss: 11.69329765762638
12 norm of gradients: 0.08808378044724723
12 norm of weights: 2.2077741423201145
-----
Iteration Number: 6467
Loss: 11.692613869808241
12 norm of gradients: 0.08806377518303553
12 norm of weights: 2.2077665364131835
-----
Iteration Number: 6468

```

Loss: 11.691930506438634
l2 norm of gradients: 0.0880437796208692
l2 norm of weights: 2.207758970671745

Iteration Number: 6469
Loss: 11.691247567214948
l2 norm of gradients: 0.08802379375398943
l2 norm of weights: 2.2077514450652775

Iteration Number: 6470
Loss: 11.690565051834822
l2 norm of gradients: 0.08800381757564307
l2 norm of weights: 2.20774395956328

Iteration Number: 6471
Loss: 11.689882959996144
l2 norm of gradients: 0.08798385107908274
l2 norm of weights: 2.2077365141352736

Iteration Number: 6472
Loss: 11.68920129139704
l2 norm of gradients: 0.08796389425756665
l2 norm of weights: 2.207729108750798

Iteration Number: 6473
Loss: 11.688520045735888
l2 norm of gradients: 0.08794394710435866
l2 norm of weights: 2.207721743379416

Iteration Number: 6474
Loss: 11.687839222711306
l2 norm of gradients: 0.08792400961272832
l2 norm of weights: 2.207714417990711

Iteration Number: 6475
Loss: 11.687158822022164
l2 norm of gradients: 0.08790408177595087
l2 norm of weights: 2.207707132554286

Iteration Number: 6476
Loss: 11.686478843367574
l2 norm of gradients: 0.08788416358730713
l2 norm of weights: 2.2076998870397664

Iteration Number: 6477
Loss: 11.685799286446887
l2 norm of gradients: 0.08786425504008356
l2 norm of weights: 2.207692681416798

```
-----
Iteration Number: 6478
Loss: 11.685120150959703
l2 norm of gradients: 0.08784435612757235
l2 norm of weights: 2.2076855156550472
-----
Iteration Number: 6479
Loss: 11.684441436605871
l2 norm of gradients: 0.08782446684307128
l2 norm of weights: 2.2076783897242014
-----
Iteration Number: 6480
Loss: 11.68376314308548
l2 norm of gradients: 0.08780458717988378
l2 norm of weights: 2.2076713035939695
-----
Iteration Number: 6481
Loss: 11.683085270098855
l2 norm of gradients: 0.08778471713131884
l2 norm of weights: 2.20766425723408
-----
Iteration Number: 6482
Loss: 11.682407817346585
l2 norm of gradients: 0.08776485669069119
l2 norm of weights: 2.2076572506142833
-----
Iteration Number: 6483
Loss: 11.681730784529472
l2 norm of gradients: 0.08774500585132114
l2 norm of weights: 2.2076502837043512
-----
Iteration Number: 6484
Loss: 11.681054171348595
l2 norm of gradients: 0.0877251646065346
l2 norm of weights: 2.207643356474074
-----
Iteration Number: 6485
Loss: 11.680377977505245
l2 norm of gradients: 0.08770533294966318
l2 norm of weights: 2.2076364688932646
-----
Iteration Number: 6486
Loss: 11.679702202700978
l2 norm of gradients: 0.08768551087404403
l2 norm of weights: 2.207629620931756
-----
Iteration Number: 6487
Loss: 11.679026846637582
```

```
12 norm of gradients: 0.08766569837301987
12 norm of weights: 2.2076228125594035
-----
Iteration Number: 6488
Loss: 11.678351909017088
12 norm of gradients: 0.08764589543993916
12 norm of weights: 2.2076160437460803
-----
Iteration Number: 6489
Loss: 11.67767738954177
12 norm of gradients: 0.08762610206815592
12 norm of weights: 2.207609314461683
-----
Iteration Number: 6490
Loss: 11.67700328791414
12 norm of gradients: 0.08760631825102971
12 norm of weights: 2.2076026246761273
-----
Iteration Number: 6491
Loss: 11.676329603836962
12 norm of gradients: 0.08758654398192575
12 norm of weights: 2.2075959743593505
-----
Iteration Number: 6492
Loss: 11.675656337013228
12 norm of gradients: 0.08756677925421488
12 norm of weights: 2.20758936348131
-----
Iteration Number: 6493
Loss: 11.674983487146175
12 norm of gradients: 0.08754702406127339
12 norm of weights: 2.2075827920119844
-----
Iteration Number: 6494
Loss: 11.674311053939292
12 norm of gradients: 0.08752727839648336
12 norm of weights: 2.2075762599213724
-----
Iteration Number: 6495
Loss: 11.67363903709629
12 norm of gradients: 0.08750754225323236
12 norm of weights: 2.2075697671794936
-----
Iteration Number: 6496
Loss: 11.672967436321134
12 norm of gradients: 0.0874878156249135
12 norm of weights: 2.2075633137563884
-----
```

Iteration Number: 6497
Loss: 11.67229625131802
l2 norm of gradients: 0.08746809850492553
l2 norm of weights: 2.207556899622119

Iteration Number: 6498
Loss: 11.671625481791393
l2 norm of gradients: 0.08744839088667278
l2 norm of weights: 2.2075505247467646

Iteration Number: 6499
Loss: 11.670955127445927
l2 norm of gradients: 0.08742869276356507
l2 norm of weights: 2.2075441891004295

Iteration Number: 6500
Loss: 11.670285187986547
l2 norm of gradients: 0.08740900412901795
l2 norm of weights: 2.207537892653235

Iteration Number: 6501
Loss: 11.669615663118405
l2 norm of gradients: 0.08738932497645231
l2 norm of weights: 2.207531635375326

Iteration Number: 6502
Loss: 11.668946552546902
l2 norm of gradients: 0.08736965529929482
l2 norm of weights: 2.2075254172368646

Iteration Number: 6503
Loss: 11.66827785597767
l2 norm of gradients: 0.08734999509097759
l2 norm of weights: 2.207519238208037

Iteration Number: 6504
Loss: 11.66760957311658
l2 norm of gradients: 0.08733034434493826
l2 norm of weights: 2.207513098259047

Iteration Number: 6505
Loss: 11.666941703669755
l2 norm of gradients: 0.08731070305462012
l2 norm of weights: 2.2075069973601207

Iteration Number: 6506
Loss: 11.66627424734353
l2 norm of gradients: 0.08729107121347193

```

12 norm of weights: 2.2075009354815034
-----
Iteration Number: 6507
Loss: 11.665607203844505
12 norm of gradients: 0.08727144881494807
12 norm of weights: 2.2074949125934626
-----
Iteration Number: 6508
Loss: 11.664940572879496
12 norm of gradients: 0.0872518358525084
12 norm of weights: 2.2074889286662853
-----
Iteration Number: 6509
Loss: 11.664274354155568
12 norm of gradients: 0.08723223231961828
12 norm of weights: 2.2074829836702783
-----
Iteration Number: 6510
Loss: 11.66360854738002
12 norm of gradients: 0.0872126382097487
12 norm of weights: 2.2074770775757697
-----
Iteration Number: 6511
Loss: 11.662943152260386
12 norm of gradients: 0.08719305351637605
12 norm of weights: 2.2074712103531082
-----
Iteration Number: 6512
Loss: 11.662278168504447
12 norm of gradients: 0.08717347823298248
12 norm of weights: 2.207465381972663
-----
Iteration Number: 6513
Loss: 11.661613595820198
12 norm of gradients: 0.08715391235305539
12 norm of weights: 2.2074595924048226
-----
Iteration Number: 6514
Loss: 11.660949433915889
12 norm of gradients: 0.08713435587008783
12 norm of weights: 2.207453841619997
-----
Iteration Number: 6515
Loss: 11.660285682500001
12 norm of gradients: 0.08711480877757842
12 norm of weights: 2.207448129588616
-----
Iteration Number: 6516

```

Loss: 11.659622341281253
l2 norm of gradients: 0.08709527106903117
l2 norm of weights: 2.2074424562811306

Iteration Number: 6517
Loss: 11.658959409968594
l2 norm of gradients: 0.08707574273795568
l2 norm of weights: 2.207436821668011

Iteration Number: 6518
Loss: 11.658296888271208
l2 norm of gradients: 0.08705622377786709
l2 norm of weights: 2.207431225719749

Iteration Number: 6519
Loss: 11.65763477589852
l2 norm of gradients: 0.08703671418228587
l2 norm of weights: 2.2074256684068554

Iteration Number: 6520
Loss: 11.656973072560184
l2 norm of gradients: 0.08701721394473817
l2 norm of weights: 2.207420149699862

Iteration Number: 6521
Loss: 11.65631177796609
l2 norm of gradients: 0.0869977230587556
l2 norm of weights: 2.2074146695693218

Iteration Number: 6522
Loss: 11.655650891826365
l2 norm of gradients: 0.08697824151787517
l2 norm of weights: 2.2074092279858064

Iteration Number: 6523
Loss: 11.654990413851367
l2 norm of gradients: 0.08695876931563948
l2 norm of weights: 2.2074038249199086

Iteration Number: 6524
Loss: 11.654330343751688
l2 norm of gradients: 0.08693930644559655
l2 norm of weights: 2.207398460342242

Iteration Number: 6525
Loss: 11.653670681238157
l2 norm of gradients: 0.08691985290129987
l2 norm of weights: 2.207393134223439

```

-----
Iteration Number: 6526
Loss: 11.653011426021832
l2 norm of gradients: 0.08690040867630855
l2 norm of weights: 2.2073878465341537
-----
Iteration Number: 6527
Loss: 11.652352577814003
l2 norm of gradients: 0.08688097376418687
l2 norm of weights: 2.2073825972450596
-----
Iteration Number: 6528
Loss: 11.651694136326203
l2 norm of gradients: 0.08686154815850497
l2 norm of weights: 2.207377386326851
-----
Iteration Number: 6529
Loss: 11.65103610127018
l2 norm of gradients: 0.08684213185283811
l2 norm of weights: 2.207372213750242
-----
Iteration Number: 6530
Loss: 11.650378472357932
l2 norm of gradients: 0.08682272484076722
l2 norm of weights: 2.2073670794859654
-----
Iteration Number: 6531
Loss: 11.64972124930168
l2 norm of gradients: 0.08680332711587867
l2 norm of weights: 2.207361983504778
-----
Iteration Number: 6532
Loss: 11.64906443181388
l2 norm of gradients: 0.08678393867176412
l2 norm of weights: 2.2073569257774537
-----
Iteration Number: 6533
Loss: 11.648408019607215
l2 norm of gradients: 0.08676455950202093
l2 norm of weights: 2.2073519062747873
-----
Iteration Number: 6534
Loss: 11.647752012394605
l2 norm of gradients: 0.0867451896002517
l2 norm of weights: 2.2073469249675934
-----
Iteration Number: 6535
Loss: 11.647096409889206

```



```

12 norm of gradients: 0.08672582896006459
12 norm of weights: 2.2073419818267075
-----
Iteration Number: 6536
Loss: 11.646441211804389
12 norm of gradients: 0.08670647757507316
12 norm of weights: 2.2073370768229847
-----
Iteration Number: 6537
Loss: 11.64578641785377
12 norm of gradients: 0.0866871354388964
12 norm of weights: 2.2073322099273005
-----
Iteration Number: 6538
Loss: 11.645132027751188
12 norm of gradients: 0.08666780254515877
12 norm of weights: 2.20732738111055
-----
Iteration Number: 6539
Loss: 11.644478041210725
12 norm of gradients: 0.0866484788874901
12 norm of weights: 2.2073225903436495
-----
Iteration Number: 6540
Loss: 11.643824457946671
12 norm of gradients: 0.0866291644595257
12 norm of weights: 2.207317837597534
-----
Iteration Number: 6541
Loss: 11.64317127767357
12 norm of gradients: 0.08660985925490626
12 norm of weights: 2.2073131228431593
-----
Iteration Number: 6542
Loss: 11.642518500106178
12 norm of gradients: 0.08659056326727797
12 norm of weights: 2.207308446051501
-----
Iteration Number: 6543
Loss: 11.641866124959487
12 norm of gradients: 0.08657127649029235
12 norm of weights: 2.2073038071935547
-----
Iteration Number: 6544
Loss: 11.641214151948722
12 norm of gradients: 0.08655199891760636
12 norm of weights: 2.2072992062403363
-----

```

Iteration Number: 6545
Loss: 11.64056258078933
l2 norm of gradients: 0.08653273054288232
l2 norm of weights: 2.2072946431628813

Iteration Number: 6546
Loss: 11.639911411196994
l2 norm of gradients: 0.08651347135978808
l2 norm of weights: 2.2072901179322457

Iteration Number: 6547
Loss: 11.63926064288762
l2 norm of gradients: 0.08649422136199676
l2 norm of weights: 2.2072856305195048

Iteration Number: 6548
Loss: 11.638610275577348
l2 norm of gradients: 0.0864749805431869
l2 norm of weights: 2.2072811808957544

Iteration Number: 6549
Loss: 11.637960308982533
l2 norm of gradients: 0.08645574889704256
l2 norm of weights: 2.2072767690321102

Iteration Number: 6550
Loss: 11.637310742819777
l2 norm of gradients: 0.08643652641725297
l2 norm of weights: 2.2072723948997077

Iteration Number: 6551
Loss: 11.636661576805901
l2 norm of gradients: 0.08641731309751297
l2 norm of weights: 2.207268058469702

Iteration Number: 6552
Loss: 11.636012810657947
l2 norm of gradients: 0.08639810893152261
l2 norm of weights: 2.207263759713269

Iteration Number: 6553
Loss: 11.635364444093195
l2 norm of gradients: 0.08637891391298741
l2 norm of weights: 2.207259498601603

Iteration Number: 6554
Loss: 11.634716476829148
l2 norm of gradients: 0.08635972803561823

```

12 norm of weights: 2.2072552751059202
-----
Iteration Number: 6555
Loss: 11.63406890858353
12 norm of gradients: 0.08634055129313135
12 norm of weights: 2.2072510891974555
-----
Iteration Number: 6556
Loss: 11.633421739074299
12 norm of gradients: 0.08632138367924834
12 norm of weights: 2.207246940847463
-----
Iteration Number: 6557
Loss: 11.632774968019643
12 norm of gradients: 0.08630222518769617
12 norm of weights: 2.2072428300272176
-----
Iteration Number: 6558
Loss: 11.632128595137967
12 norm of gradients: 0.0862830758122072
12 norm of weights: 2.2072387567080143
-----
Iteration Number: 6559
Loss: 11.631482620147906
12 norm of gradients: 0.08626393554651911
12 norm of weights: 2.207234720861167
-----
Iteration Number: 6560
Loss: 11.630837042768322
12 norm of gradients: 0.08624480438437494
12 norm of weights: 2.20723072245801
-----
Iteration Number: 6561
Loss: 11.630191862718299
12 norm of gradients: 0.08622568231952309
12 norm of weights: 2.2072267614698973
-----
Iteration Number: 6562
Loss: 11.629547079717147
12 norm of gradients: 0.08620656934571727
12 norm of weights: 2.2072228378682026
-----
Iteration Number: 6563
Loss: 11.628902693484413
12 norm of gradients: 0.08618746545671659
12 norm of weights: 2.207218951624319
-----
Iteration Number: 6564

```

```

Loss: 11.628258703739846
l2 norm of gradients: 0.08616837064628542
l2 norm of weights: 2.2072151027096605
-----
Iteration Number: 6565
Loss: 11.627615110203443
l2 norm of gradients: 0.08614928490819354
l2 norm of weights: 2.2072112910956596
-----
Iteration Number: 6566
Loss: 11.626971912595405
l2 norm of gradients: 0.08613020823621605
l2 norm of weights: 2.207207516753769
-----
Iteration Number: 6567
Loss: 11.626329110636174
l2 norm of gradients: 0.08611114062413332
l2 norm of weights: 2.2072037796554618
-----
Iteration Number: 6568
Loss: 11.62568670404641
l2 norm of gradients: 0.08609208206573105
l2 norm of weights: 2.2072000797722287
-----
Iteration Number: 6569
Loss: 11.625044692546995
l2 norm of gradients: 0.08607303255480037
l2 norm of weights: 2.207196417075583
-----
Iteration Number: 6570
Loss: 11.624403075859036
l2 norm of gradients: 0.08605399208513755
l2 norm of weights: 2.2071927915370555
-----
Iteration Number: 6571
Loss: 11.62376185370386
l2 norm of gradients: 0.08603496065054435
l2 norm of weights: 2.2071892031281974
-----
Iteration Number: 6572
Loss: 11.623121025803025
l2 norm of gradients: 0.08601593824482769
l2 norm of weights: 2.2071856518205797
-----
Iteration Number: 6573
Loss: 11.62248059187831
l2 norm of gradients: 0.08599692486179993
l2 norm of weights: 2.207182137585793

```

```

-----
Iteration Number: 6574
Loss: 11.621840551651706
l2 norm of gradients: 0.08597792049527855
l2 norm of weights: 2.2071786603954466
-----
Iteration Number: 6575
Loss: 11.621200904845441
l2 norm of gradients: 0.08595892513908651
l2 norm of weights: 2.207175220221171
-----
Iteration Number: 6576
Loss: 11.62056165118196
l2 norm of gradients: 0.08593993878705194
l2 norm of weights: 2.2071718170346157
-----
Iteration Number: 6577
Loss: 11.619922790383926
l2 norm of gradients: 0.08592096143300837
l2 norm of weights: 2.2071684508074485
-----
Iteration Number: 6578
Loss: 11.619284322174233
l2 norm of gradients: 0.08590199307079448
l2 norm of weights: 2.207165121511359
-----
Iteration Number: 6579
Loss: 11.618646246275986
l2 norm of gradients: 0.08588303369425436
l2 norm of weights: 2.2071618291180544
-----
Iteration Number: 6580
Loss: 11.618008562412518
l2 norm of gradients: 0.08586408329723724
l2 norm of weights: 2.207158573599263
-----
Iteration Number: 6581
Loss: 11.617371270307384
l2 norm of gradients: 0.08584514187359782
l2 norm of weights: 2.2071553549267318
-----
Iteration Number: 6582
Loss: 11.616734369684359
l2 norm of gradients: 0.08582620941719585
l2 norm of weights: 2.207152173072227
-----
Iteration Number: 6583
Loss: 11.616097860267432

```

```

12 norm of gradients: 0.08580728592189647
12 norm of weights: 2.207149028007535
-----
Iteration Number: 6584
Loss: 11.615461741780825
12 norm of gradients: 0.08578837138157012
12 norm of weights: 2.2071459197044616
-----
Iteration Number: 6585
Loss: 11.614826013948973
12 norm of gradients: 0.08576946579009241
12 norm of weights: 2.207142848134832
-----
Iteration Number: 6586
Loss: 11.61419067649653
12 norm of gradients: 0.08575056914134421
12 norm of weights: 2.2071398132704907
-----
Iteration Number: 6587
Loss: 11.613555729148375
12 norm of gradients: 0.08573168142921168
12 norm of weights: 2.2071368150833015
-----
Iteration Number: 6588
Loss: 11.612921171629603
12 norm of gradients: 0.08571280264758625
12 norm of weights: 2.2071338535451486
-----
Iteration Number: 6589
Loss: 11.612287003665532
12 norm of gradients: 0.08569393279036457
12 norm of weights: 2.2071309286279344
-----
Iteration Number: 6590
Loss: 11.611653224981694
12 norm of gradients: 0.08567507185144844
12 norm of weights: 2.207128040303582
-----
Iteration Number: 6591
Loss: 11.611019835303846
12 norm of gradients: 0.08565621982474504
12 norm of weights: 2.2071251885440324
-----
Iteration Number: 6592
Loss: 11.61038683435796
12 norm of gradients: 0.0856373767041667
12 norm of weights: 2.207122373321247
-----

```

Iteration Number: 6593
Loss: 11.609754221870231
l2 norm of gradients: 0.08561854248363103
l2 norm of weights: 2.207119594607207

Iteration Number: 6594
Loss: 11.609121997567065
l2 norm of gradients: 0.0855997171570608
l2 norm of weights: 2.207116852373912

Iteration Number: 6595
Loss: 11.6084901611751
l2 norm of gradients: 0.08558090071838403
l2 norm of weights: 2.207114146593381

Iteration Number: 6596
Loss: 11.607858712421173
l2 norm of gradients: 0.08556209316153403
l2 norm of weights: 2.207111477237653

Iteration Number: 6597
Loss: 11.607227651032355
l2 norm of gradients: 0.08554329448044912
l2 norm of weights: 2.2071088442787867

Iteration Number: 6598
Loss: 11.60659697673593
l2 norm of gradients: 0.08552450466907309
l2 norm of weights: 2.2071062476888583

Iteration Number: 6599
Loss: 11.605966689259397
l2 norm of gradients: 0.08550572372135479
l2 norm of weights: 2.207103687439966

Iteration Number: 6600
Loss: 11.605336788330472
l2 norm of gradients: 0.08548695163124823
l2 norm of weights: 2.2071011635042237

Iteration Number: 6601
Loss: 11.604707273677095
l2 norm of gradients: 0.08546818839271272
l2 norm of weights: 2.2070986758537687

Iteration Number: 6602
Loss: 11.604078145027412
l2 norm of gradients: 0.08544943399971278

```
12 norm of weights: 2.207096224460754
-----
Iteration Number: 6603
Loss: 11.603449402109797
12 norm of gradients: 0.08543068844621796
12 norm of weights: 2.207093809297354
-----
Iteration Number: 6604
Loss: 11.602821044652833
12 norm of gradients: 0.08541195172620318
12 norm of weights: 2.2070914303357623
-----
Iteration Number: 6605
Loss: 11.602193072385319
12 norm of gradients: 0.08539322383364845
12 norm of weights: 2.2070890875481903
-----
Iteration Number: 6606
Loss: 11.601565485036273
12 norm of gradients: 0.08537450476253901
12 norm of weights: 2.20708678090687
-----
Iteration Number: 6607
Loss: 11.600938282334933
12 norm of gradients: 0.08535579450686515
12 norm of weights: 2.2070845103840524
-----
Iteration Number: 6608
Loss: 11.600311464010746
12 norm of gradients: 0.08533709306062248
12 norm of weights: 2.207082275952007
-----
Iteration Number: 6609
Loss: 11.599685029793372
12 norm of gradients: 0.08531840041781175
12 norm of weights: 2.2070800775830226
-----
Iteration Number: 6610
Loss: 11.599058979412698
12 norm of gradients: 0.08529971657243879
12 norm of weights: 2.207077915249408
-----
Iteration Number: 6611
Loss: 11.59843331259881
12 norm of gradients: 0.08528104151851466
12 norm of weights: 2.207075788923491
-----
Iteration Number: 6612
```


Loss: 11.597808029082028
l2 norm of gradients: 0.08526237525005559
l2 norm of weights: 2.207073698577618

Iteration Number: 6613
Loss: 11.597183128592869
l2 norm of gradients: 0.08524371776108289
l2 norm of weights: 2.207071644184154

Iteration Number: 6614
Loss: 11.596558610862074
l2 norm of gradients: 0.08522506904562313
l2 norm of weights: 2.2070696257154845

Iteration Number: 6615
Loss: 11.595934475620595
l2 norm of gradients: 0.08520642909770788
l2 norm of weights: 2.207067643144014

Iteration Number: 6616
Loss: 11.5953107225996
l2 norm of gradients: 0.08518779791137399
l2 norm of weights: 2.2070656964421644

Iteration Number: 6617
Loss: 11.594687351530467
l2 norm of gradients: 0.0851691754806633
l2 norm of weights: 2.2070637855823785

Iteration Number: 6618
Loss: 11.594064362144797
l2 norm of gradients: 0.08515056179962298
l2 norm of weights: 2.207061910537118

Iteration Number: 6619
Loss: 11.593441754174393
l2 norm of gradients: 0.08513195686230515
l2 norm of weights: 2.2070600712788626

Iteration Number: 6620
Loss: 11.592819527351274
l2 norm of gradients: 0.08511336066276712
l2 norm of weights: 2.2070582677801123

Iteration Number: 6621
Loss: 11.592197681407674
l2 norm of gradients: 0.08509477319507136
l2 norm of weights: 2.2070565000133846

```

-----
Iteration Number: 6622
Loss: 11.591576216076048
l2 norm of gradients: 0.08507619445328543
l2 norm of weights: 2.207054767951218
-----
Iteration Number: 6623
Loss: 11.590955131089048
l2 norm of gradients: 0.08505762443148193
l2 norm of weights: 2.2070530715661683
-----
Iteration Number: 6624
Loss: 11.590334426179545
l2 norm of gradients: 0.08503906312373873
l2 norm of weights: 2.207051410830811
-----
Iteration Number: 6625
Loss: 11.589714101080629
l2 norm of gradients: 0.08502051052413861
l2 norm of weights: 2.2070497857177407
-----
Iteration Number: 6626
Loss: 11.589094155525592
l2 norm of gradients: 0.08500196662676966
l2 norm of weights: 2.207048196199571
-----
Iteration Number: 6627
Loss: 11.588474589247943
l2 norm of gradients: 0.08498343142572491
l2 norm of weights: 2.207046642248934
-----
Iteration Number: 6628
Loss: 11.5878554019814
l2 norm of gradients: 0.08496490491510258
l2 norm of weights: 2.207045123838482
-----
Iteration Number: 6629
Loss: 11.587236593459894
l2 norm of gradients: 0.08494638708900584
l2 norm of weights: 2.207043640940884
-----
Iteration Number: 6630
Loss: 11.586618163417565
l2 norm of gradients: 0.08492787794154315
l2 norm of weights: 2.20704219352883
-----
Iteration Number: 6631
Loss: 11.586000111588774

```

```

12 norm of gradients: 0.08490937746682793
12 norm of weights: 2.207040781575028
-----
Iteration Number: 6632
Loss: 11.585382437708077
12 norm of gradients: 0.08489088565897869
12 norm of weights: 2.207039405052205
-----
Iteration Number: 6633
Loss: 11.584765141510246
12 norm of gradients: 0.08487240251211901
12 norm of weights: 2.2070380639331066
-----
Iteration Number: 6634
Loss: 11.584148222730278
12 norm of gradients: 0.0848539280203776
12 norm of weights: 2.207036758190499
-----
Iteration Number: 6635
Loss: 11.583531681103358
12 norm of gradients: 0.08483546217788814
12 norm of weights: 2.2070354877971643
-----
Iteration Number: 6636
Loss: 11.582915516364887
12 norm of gradients: 0.08481700497878948
12 norm of weights: 2.2070342527259057
-----
Iteration Number: 6637
Loss: 11.582299728250492
12 norm of gradients: 0.08479855641722545
12 norm of weights: 2.2070330529495448
-----
Iteration Number: 6638
Loss: 11.58168431649599
12 norm of gradients: 0.08478011648734494
12 norm of weights: 2.207031888440922
-----
Iteration Number: 6639
Loss: 11.581069280837415
12 norm of gradients: 0.08476168518330199
12 norm of weights: 2.2070307591728957
-----
Iteration Number: 6640
Loss: 11.580454621011013
12 norm of gradients: 0.08474326249925558
12 norm of weights: 2.2070296651183443
-----

```

Iteration Number: 6641
Loss: 11.579840336753229
l2 norm of gradients: 0.08472484842936974
l2 norm of weights: 2.2070286062501645

Iteration Number: 6642
Loss: 11.579226427800728
l2 norm of gradients: 0.08470644296781359
l2 norm of weights: 2.207027582541272

Iteration Number: 6643
Loss: 11.578612893890382
l2 norm of gradients: 0.08468804610876127
l2 norm of weights: 2.2070265939646

Iteration Number: 6644
Loss: 11.577999734759263
l2 norm of gradients: 0.08466965784639198
l2 norm of weights: 2.207025640493103

Iteration Number: 6645
Loss: 11.577386950144664
l2 norm of gradients: 0.08465127817488989
l2 norm of weights: 2.207024722099751

Iteration Number: 6646
Loss: 11.57677453978407
l2 norm of gradients: 0.0846329070884442
l2 norm of weights: 2.207023838757536

Iteration Number: 6647
Loss: 11.576162503415189
l2 norm of gradients: 0.08461454458124924
l2 norm of weights: 2.2070229904394663

Iteration Number: 6648
Loss: 11.57555084077593
l2 norm of gradients: 0.08459619064750418
l2 norm of weights: 2.2070221771185703

Iteration Number: 6649
Loss: 11.574939551604405
l2 norm of gradients: 0.08457784528141336
l2 norm of weights: 2.2070213987678944

Iteration Number: 6650
Loss: 11.574328635638942
l2 norm of gradients: 0.084559508477186

```
l2 norm of weights: 2.2070206553605036
-----
Iteration Number: 6651
Loss: 11.57371809261807
l2 norm of gradients: 0.08454118022903645
l2 norm of weights: 2.2070199468694827
-----
Iteration Number: 6652
Loss: 11.573107922280526
l2 norm of gradients: 0.084522860531184
l2 norm of weights: 2.207019273267934
-----
Iteration Number: 6653
Loss: 11.572498124365259
l2 norm of gradients: 0.0845045493778529
l2 norm of weights: 2.2070186345289784
-----
Iteration Number: 6654
Loss: 11.571888698611412
l2 norm of gradients: 0.08448624676327249
l2 norm of weights: 2.2070180306257563
-----
Iteration Number: 6655
Loss: 11.571279644758347
l2 norm of gradients: 0.084467952681677
l2 norm of weights: 2.207017461531426
-----
Iteration Number: 6656
Loss: 11.57067096254563
l2 norm of gradients: 0.0844496671273057
l2 norm of weights: 2.207016927219165
-----
Iteration Number: 6657
Loss: 11.570062651713025
l2 norm of gradients: 0.08443139009440286
l2 norm of weights: 2.2070164276621687
-----
Iteration Number: 6658
Loss: 11.569454712000505
l2 norm of gradients: 0.08441312157721764
l2 norm of weights: 2.2070159628336516
-----
Iteration Number: 6659
Loss: 11.568847143148256
l2 norm of gradients: 0.08439486157000431
l2 norm of weights: 2.207015532706847
-----
Iteration Number: 6660
```

Loss: 11.568239944896657
l2 norm of gradients: 0.084376610067022
l2 norm of weights: 2.207015137255006

Iteration Number: 6661
Loss: 11.567633116986304
l2 norm of gradients: 0.08435836706253483
l2 norm of weights: 2.207014776451399

Iteration Number: 6662
Loss: 11.56702665915799
l2 norm of gradients: 0.08434013255081194
l2 norm of weights: 2.2070144502693148

Iteration Number: 6663
Loss: 11.56642057115271
l2 norm of gradients: 0.08432190652612734
l2 norm of weights: 2.20701415868206

Iteration Number: 6664
Loss: 11.565814852711675
l2 norm of gradients: 0.08430368898276007
l2 norm of weights: 2.2070139016629606

Iteration Number: 6665
Loss: 11.56520950357629
l2 norm of gradients: 0.08428547991499408
l2 norm of weights: 2.207013679185361

Iteration Number: 6666
Loss: 11.564604523488168
l2 norm of gradients: 0.08426727931711829
l2 norm of weights: 2.2070134912226234

Iteration Number: 6667
Loss: 11.563999912189127
l2 norm of gradients: 0.08424908718342652
l2 norm of weights: 2.2070133377481294

Iteration Number: 6668
Loss: 11.56339566942119
l2 norm of gradients: 0.08423090350821759
l2 norm of weights: 2.2070132187352787

Iteration Number: 6669
Loss: 11.562791794926575
l2 norm of gradients: 0.08421272828579524
l2 norm of weights: 2.207013134157489

```
-----  
Iteration Number: 6670  
Loss: 11.562188288447713  
l2 norm of gradients: 0.08419456151046807  
l2 norm of weights: 2.207013083988197  
-----  
Iteration Number: 6671  
Loss: 11.561585149727229  
l2 norm of gradients: 0.08417640317654972  
l2 norm of weights: 2.207013068200858  
-----  
Iteration Number: 6672  
Loss: 11.560982378507964  
l2 norm of gradients: 0.08415825327835869  
l2 norm of weights: 2.2070130867689453  
-----  
Iteration Number: 6673  
Loss: 11.560379974532948  
l2 norm of gradients: 0.08414011181021838  
l2 norm of weights: 2.2070131396659503  
-----  
Iteration Number: 6674  
Loss: 11.559777937545423  
l2 norm of gradients: 0.08412197876645715  
l2 norm of weights: 2.207013226865384  
-----  
Iteration Number: 6675  
Loss: 11.559176267288827  
l2 norm of gradients: 0.08410385414140827  
l2 norm of weights: 2.207013348340775  
-----  
Iteration Number: 6676  
Loss: 11.558574963506807  
l2 norm of gradients: 0.08408573792940988  
l2 norm of weights: 2.207013504065669  
-----  
Iteration Number: 6677  
Loss: 11.557974025943205  
l2 norm of gradients: 0.08406763012480506  
l2 norm of weights: 2.207013694013633  
-----  
Iteration Number: 6678  
Loss: 11.557373454342072  
l2 norm of gradients: 0.08404953072194181  
l2 norm of weights: 2.2070139181582498  
-----  
Iteration Number: 6679  
Loss: 11.55677324844765
```

```

12 norm of gradients: 0.08403143971517293
12 norm of weights: 2.2070141764731224
-----
Iteration Number: 6680
Loss: 11.556173408004392
12 norm of gradients: 0.08401335709885625
12 norm of weights: 2.20701446893187
-----
Iteration Number: 6681
Loss: 11.555573932756952
12 norm of gradients: 0.08399528286735435
12 norm of weights: 2.207014795508132
-----
Iteration Number: 6682
Loss: 11.554974822450177
12 norm of gradients: 0.08397721701503479
12 norm of weights: 2.2070151561755655
-----
Iteration Number: 6683
Loss: 11.554376076829122
12 norm of gradients: 0.08395915953626994
12 norm of weights: 2.2070155509078457
-----
Iteration Number: 6684
Loss: 11.553777695639042
12 norm of gradients: 0.08394111042543718
12 norm of weights: 2.207015979678666
-----
Iteration Number: 6685
Loss: 11.553179678625392
12 norm of gradients: 0.0839230696769186
12 norm of weights: 2.2070164424617387
-----
Iteration Number: 6686
Loss: 11.552582025533825
12 norm of gradients: 0.08390503728510126
12 norm of weights: 2.207016939230794
-----
Iteration Number: 6687
Loss: 11.551984736110192
12 norm of gradients: 0.08388701324437702
12 norm of weights: 2.2070174699595793
-----
Iteration Number: 6688
Loss: 11.551387810100557
12 norm of gradients: 0.08386899754914269
12 norm of weights: 2.2070180346218624
-----

```


Iteration Number: 6689
Loss: 11.550791247251162
l2 norm of gradients: 0.0838509901937998
l2 norm of weights: 2.207018633191428

Iteration Number: 6690
Loss: 11.550195047308467
l2 norm of gradients: 0.08383299117275486
l2 norm of weights: 2.2070192656420784

Iteration Number: 6691
Loss: 11.549599210019124
l2 norm of gradients: 0.08381500048041923
l2 norm of weights: 2.207019931947636

Iteration Number: 6692
Loss: 11.54900373512999
l2 norm of gradients: 0.08379701811120899
l2 norm of weights: 2.207020632081939

Iteration Number: 6693
Loss: 11.548408622388108
l2 norm of gradients: 0.08377904405954519
l2 norm of weights: 2.2070213660188465

Iteration Number: 6694
Loss: 11.547813871540734
l2 norm of gradients: 0.08376107831985365
l2 norm of weights: 2.207022133732233

Iteration Number: 6695
Loss: 11.547219482335311
l2 norm of gradients: 0.08374312088656502
l2 norm of weights: 2.2070229351959934

Iteration Number: 6696
Loss: 11.546625454519493
l2 norm of gradients: 0.08372517175411483
l2 norm of weights: 2.207023770384039

Iteration Number: 6697
Loss: 11.546031787841118
l2 norm of gradients: 0.08370723091694339
l2 norm of weights: 2.207024639270301

Iteration Number: 6698
Loss: 11.545438482048233
l2 norm of gradients: 0.08368929836949589

```
12 norm of weights: 2.2070255418287266
-----
Iteration Number: 6699
Loss: 11.54484553688908
12 norm of gradients: 0.08367137410622225
12 norm of weights: 2.2070264780332836
-----
Iteration Number: 6700
Loss: 11.544252952112094
12 norm of gradients: 0.08365345812157729
12 norm of weights: 2.2070274478579557
-----
Iteration Number: 6701
Loss: 11.54366072746591
12 norm of gradients: 0.08363555041002055
12 norm of weights: 2.2070284512767455
-----
Iteration Number: 6702
Loss: 11.54306886269937
12 norm of gradients: 0.08361765096601646
12 norm of weights: 2.2070294882636743
-----
Iteration Number: 6703
Loss: 11.542477357561495
12 norm of gradients: 0.08359975978403424
12 norm of weights: 2.2070305587927805
-----
Iteration Number: 6704
Loss: 11.541886211801518
12 norm of gradients: 0.0835818768585478
12 norm of weights: 2.2070316628381206
-----
Iteration Number: 6705
Loss: 11.54129542516886
12 norm of gradients: 0.08356400218403602
12 norm of weights: 2.2070328003737703
-----
Iteration Number: 6706
Loss: 11.540704997413144
12 norm of gradients: 0.08354613575498245
12 norm of weights: 2.2070339713738223
-----
Iteration Number: 6707
Loss: 11.540114928284186
12 norm of gradients: 0.08352827756587541
12 norm of weights: 2.2070351758123876
-----
Iteration Number: 6708
```

Loss: 11.539525217531999
l2 norm of gradients: 0.08351042761120815
l2 norm of weights: 2.2070364136635945

Iteration Number: 6709
Loss: 11.53893586490679
l2 norm of gradients: 0.08349258588547846
l2 norm of weights: 2.2070376849015902

Iteration Number: 6710
Loss: 11.538346870158968
l2 norm of gradients: 0.08347475238318919
l2 norm of weights: 2.20703898950054

Iteration Number: 6711
Loss: 11.537758233039135
l2 norm of gradients: 0.0834569270988477
l2 norm of weights: 2.2070403274346266

Iteration Number: 6712
Loss: 11.537169953298081
l2 norm of gradients: 0.08343911002696629
l2 norm of weights: 2.207041698678051

Iteration Number: 6713
Loss: 11.536582030686802
l2 norm of gradients: 0.08342130116206194
l2 norm of weights: 2.207043103205031

Iteration Number: 6714
Loss: 11.535994464956485
l2 norm of gradients: 0.08340350049865641
l2 norm of weights: 2.2070445409898047

Iteration Number: 6715
Loss: 11.535407255858505
l2 norm of gradients: 0.0833857080312762
l2 norm of weights: 2.2070460120066255

Iteration Number: 6716
Loss: 11.534820403144442
l2 norm of gradients: 0.08336792375445265
l2 norm of weights: 2.207047516229767

Iteration Number: 6717
Loss: 11.534233906566069
l2 norm of gradients: 0.08335014766272167
l2 norm of weights: 2.2070490536335194

```

-----
Iteration Number: 6718
Loss: 11.533647765875347
l2 norm of gradients: 0.0833323797506241
l2 norm of weights: 2.2070506241921906
-----
Iteration Number: 6719
Loss: 11.53306198082444
l2 norm of gradients: 0.08331462001270537
l2 norm of weights: 2.2070522278801072
-----
Iteration Number: 6720
Loss: 11.532476551165695
l2 norm of gradients: 0.08329686844351572
l2 norm of weights: 2.207053864671613
-----
Iteration Number: 6721
Loss: 11.531891476651662
l2 norm of gradients: 0.08327912503761017
l2 norm of weights: 2.2070555345410705
-----
Iteration Number: 6722
Loss: 11.531306757035082
l2 norm of gradients: 0.08326138978954835
l2 norm of weights: 2.207057237462859
-----
Iteration Number: 6723
Loss: 11.530722392068888
l2 norm of gradients: 0.0832436626938947
l2 norm of weights: 2.2070589734113764
-----
Iteration Number: 6724
Loss: 11.530138381506207
l2 norm of gradients: 0.08322594374521834
l2 norm of weights: 2.2070607423610378
-----
Iteration Number: 6725
Loss: 11.52955472510036
l2 norm of gradients: 0.08320823293809312
l2 norm of weights: 2.207062544286276
-----
Iteration Number: 6726
Loss: 11.52897142260486
l2 norm of gradients: 0.0831905302670976
l2 norm of weights: 2.2070643791615434
-----
Iteration Number: 6727
Loss: 11.52838847377341

```

l2 norm of gradients: 0.08317283572681508
l2 norm of weights: 2.207066246961308

Iteration Number: 6728
Loss: 11.52780587835991
l2 norm of gradients: 0.08315514931183345
l2 norm of weights: 2.2070681476600567

Iteration Number: 6729
Loss: 11.527223636118453
l2 norm of gradients: 0.08313747101674546
l2 norm of weights: 2.207070081232293

Iteration Number: 6730
Loss: 11.526641746803325
l2 norm of gradients: 0.08311980083614844
l2 norm of weights: 2.2070720476525407

Iteration Number: 6731
Loss: 11.526060210168989
l2 norm of gradients: 0.08310213876464438
l2 norm of weights: 2.207074046895338

Iteration Number: 6732
Loss: 11.525479025970125
l2 norm of gradients: 0.08308448479684018
l2 norm of weights: 2.2070760789352435

Iteration Number: 6733
Loss: 11.52489819396158
l2 norm of gradients: 0.08306683892734712
l2 norm of weights: 2.207078143746832

Iteration Number: 6734
Loss: 11.524317713898412
l2 norm of gradients: 0.0830492011507814
l2 norm of weights: 2.2070802413046966

Iteration Number: 6735
Loss: 11.523737585535859
l2 norm of gradients: 0.08303157146176376
l2 norm of weights: 2.207082371583448

Iteration Number: 6736
Loss: 11.523157808629353
l2 norm of gradients: 0.08301394985491968
l2 norm of weights: 2.207084534557715

Iteration Number: 6737
Loss: 11.522578382934519
l2 norm of gradients: 0.0829963363248793
l2 norm of weights: 2.2070867302021435

Iteration Number: 6738
Loss: 11.521999308207167
l2 norm of gradients: 0.08297873086627734
l2 norm of weights: 2.207088958491397

Iteration Number: 6739
Loss: 11.521420584203307
l2 norm of gradients: 0.08296113347375329
l2 norm of weights: 2.2070912194001573

Iteration Number: 6740
Loss: 11.520842210679126
l2 norm of gradients: 0.08294354414195132
l2 norm of weights: 2.2070935129031226

Iteration Number: 6741
Loss: 11.52026418739102
l2 norm of gradients: 0.08292596286552009
l2 norm of weights: 2.2070958389750106

Iteration Number: 6742
Loss: 11.519686514095552
l2 norm of gradients: 0.08290838963911307
l2 norm of weights: 2.207098197590555

Iteration Number: 6743
Loss: 11.519109190549496
l2 norm of gradients: 0.08289082445738827
l2 norm of weights: 2.207100588724507

Iteration Number: 6744
Loss: 11.5185322165098
l2 norm of gradients: 0.08287326731500841
l2 norm of weights: 2.207103012351637

Iteration Number: 6745
Loss: 11.517955591733614
l2 norm of gradients: 0.08285571820664084
l2 norm of weights: 2.2071054684467324

Iteration Number: 6746
Loss: 11.51737931597827
l2 norm of gradients: 0.08283817712695744

```

12 norm of weights: 2.2071079569845966
-----
Iteration Number: 6747
Loss: 11.516803389001286
12 norm of gradients: 0.08282064407063493
12 norm of weights: 2.2071104779400526
-----
Iteration Number: 6748
Loss: 11.516227810560379
12 norm of gradients: 0.08280311903235439
12 norm of weights: 2.20711303128794
-----
Iteration Number: 6749
Loss: 11.51565258041345
12 norm of gradients: 0.08278560200680177
12 norm of weights: 2.207115617003116
-----
Iteration Number: 6750
Loss: 11.51507769831858
12 norm of gradients: 0.08276809298866744
12 norm of weights: 2.207118235060455
-----
Iteration Number: 6751
Loss: 11.514503164034053
12 norm of gradients: 0.08275059197264652
12 norm of weights: 2.20712088543485
-----
Iteration Number: 6752
Loss: 11.513928977318335
12 norm of gradients: 0.08273309895343864
12 norm of weights: 2.2071235681012102
-----
Iteration Number: 6753
Loss: 11.513355137930079
12 norm of gradients: 0.08271561392574811
12 norm of weights: 2.2071262830344636
-----
Iteration Number: 6754
Loss: 11.512781645628124
12 norm of gradients: 0.08269813688428382
12 norm of weights: 2.207129030209554
-----
Iteration Number: 6755
Loss: 11.512208500171498
12 norm of gradients: 0.08268066782375924
12 norm of weights: 2.2071318096014445
-----
Iteration Number: 6756

```

Loss: 11.511635701319426
l2 norm of gradients: 0.08266320673889242
l2 norm of weights: 2.2071346211851144

Iteration Number: 6757
Loss: 11.511063248831304
l2 norm of gradients: 0.08264575362440607
l2 norm of weights: 2.207137464935561

Iteration Number: 6758
Loss: 11.51049114246673
l2 norm of gradients: 0.08262830847502745
l2 norm of weights: 2.2071403408277988

Iteration Number: 6759
Loss: 11.509919381985481
l2 norm of gradients: 0.08261087128548833
l2 norm of weights: 2.2071432488368594

Iteration Number: 6760
Loss: 11.509347967147516
l2 norm of gradients: 0.08259344205052513
l2 norm of weights: 2.207146188937793

Iteration Number: 6761
Loss: 11.508776897712997
l2 norm of gradients: 0.08257602076487883
l2 norm of weights: 2.207149161105666

Iteration Number: 6762
Loss: 11.508206173442256
l2 norm of gradients: 0.082558607423295
l2 norm of weights: 2.207152165315563

Iteration Number: 6763
Loss: 11.507635794095824
l2 norm of gradients: 0.08254120202052379
l2 norm of weights: 2.2071552015425846

Iteration Number: 6764
Loss: 11.507065759434406
l2 norm of gradients: 0.08252380455131982
l2 norm of weights: 2.207158269761851

Iteration Number: 6765
Loss: 11.506496069218901
l2 norm of gradients: 0.08250641501044235
l2 norm of weights: 2.2071613699484978


```

-----
Iteration Number: 6766
Loss: 11.505926723210395
l2 norm of gradients: 0.08248903339265515
l2 norm of weights: 2.2071645020776787
-----
Iteration Number: 6767
Loss: 11.505357721170157
l2 norm of gradients: 0.08247165969272661
l2 norm of weights: 2.207167666124565
-----
Iteration Number: 6768
Loss: 11.504789062859636
l2 norm of gradients: 0.0824542939054296
l2 norm of weights: 2.2071708620643444
-----
Iteration Number: 6769
Loss: 11.504220748040478
l2 norm of gradients: 0.08243693602554153
l2 norm of weights: 2.2071740898722236
-----
Iteration Number: 6770
Loss: 11.5036527764745
l2 norm of gradients: 0.08241958604784438
l2 norm of weights: 2.207177349523425
-----
Iteration Number: 6771
Loss: 11.50308514792372
l2 norm of gradients: 0.08240224396712466
l2 norm of weights: 2.2071806409931884
-----
Iteration Number: 6772
Loss: 11.502517862150329
l2 norm of gradients: 0.08238490977817339
l2 norm of weights: 2.2071839642567723
-----
Iteration Number: 6773
Loss: 11.501950918916707
l2 norm of gradients: 0.08236758347578611
l2 norm of weights: 2.2071873192894502
-----
Iteration Number: 6774
Loss: 11.501384317985417
l2 norm of gradients: 0.08235026505476298
l2 norm of weights: 2.2071907060665157
-----
Iteration Number: 6775
Loss: 11.500818059119204

```

l2 norm of gradients: 0.08233295450990853
l2 norm of weights: 2.207194124563277

Iteration Number: 6776
Loss: 11.500252142081004
l2 norm of gradients: 0.0823156518360319
l2 norm of weights: 2.2071975747550603

Iteration Number: 6777
Loss: 11.499686566633935
l2 norm of gradients: 0.08229835702794672
l2 norm of weights: 2.2072010566172104

Iteration Number: 6778
Loss: 11.49912133254129
l2 norm of gradients: 0.08228107008047111
l2 norm of weights: 2.207204570125088

Iteration Number: 6779
Loss: 11.498556439566556
l2 norm of gradients: 0.08226379098842776
l2 norm of weights: 2.207208115254071

Iteration Number: 6780
Loss: 11.4979918874734
l2 norm of gradients: 0.08224651974664375
l2 norm of weights: 2.207211691979554

Iteration Number: 6781
Loss: 11.497427676025673
l2 norm of gradients: 0.08222925634995072
l2 norm of weights: 2.2072153002769515

Iteration Number: 6782
Loss: 11.496863804987406
l2 norm of gradients: 0.08221200079318479
l2 norm of weights: 2.2072189401216917

Iteration Number: 6783
Loss: 11.496300274122818
l2 norm of gradients: 0.08219475307118658
l2 norm of weights: 2.2072226114892217

Iteration Number: 6784
Loss: 11.495737083196307
l2 norm of gradients: 0.08217751317880122
l2 norm of weights: 2.2072263143550064

Iteration Number: 6785
Loss: 11.495174231972449
l2 norm of gradients: 0.08216028111087822
l2 norm of weights: 2.2072300486945253

Iteration Number: 6786
Loss: 11.494611720216012
l2 norm of gradients: 0.08214305686227168
l2 norm of weights: 2.2072338144832786

Iteration Number: 6787
Loss: 11.494049547691946
l2 norm of gradients: 0.0821258404278401
l2 norm of weights: 2.20723761169678

Iteration Number: 6788
Loss: 11.493487714165376
l2 norm of gradients: 0.08210863180244649
l2 norm of weights: 2.207241440310563

Iteration Number: 6789
Loss: 11.492926219401607
l2 norm of gradients: 0.08209143098095822
l2 norm of weights: 2.2072453003001766

Iteration Number: 6790
Loss: 11.49236506316614
l2 norm of gradients: 0.08207423795824734
l2 norm of weights: 2.2072491916411883

Iteration Number: 6791
Loss: 11.491804245224643
l2 norm of gradients: 0.0820570527291901
l2 norm of weights: 2.207253114309181

Iteration Number: 6792
Loss: 11.491243765342968
l2 norm of gradients: 0.08203987528866735
l2 norm of weights: 2.207257068279756

Iteration Number: 6793
Loss: 11.49068362328716
l2 norm of gradients: 0.08202270563156443
l2 norm of weights: 2.207261053528531

Iteration Number: 6794
Loss: 11.49012381882343
l2 norm of gradients: 0.08200554375277096

```

12 norm of weights: 2.207265070031141
-----
Iteration Number: 6795
Loss: 11.489564351718176
12 norm of gradients: 0.08198838964718115
12 norm of weights: 2.2072691177632375
-----
Iteration Number: 6796
Loss: 11.489005221737978
12 norm of gradients: 0.08197124330969355
12 norm of weights: 2.2072731967004904
-----
Iteration Number: 6797
Loss: 11.488446428649597
12 norm of gradients: 0.08195410473521121
12 norm of weights: 2.2072773068185847
-----
Iteration Number: 6798
Loss: 11.487887972219971
12 norm of gradients: 0.08193697391864158
12 norm of weights: 2.207281448093223
-----
Iteration Number: 6799
Loss: 11.48732985221622
12 norm of gradients: 0.08191985085489654
12 norm of weights: 2.2072856205001266
-----
Iteration Number: 6800
Loss: 11.486772068405644
12 norm of gradients: 0.0819027355388924
12 norm of weights: 2.2072898240150316
-----
Iteration Number: 6801
Loss: 11.486214620555725
12 norm of gradients: 0.08188562796554981
12 norm of weights: 2.2072940586136918
-----
Iteration Number: 6802
Loss: 11.48565750843412
12 norm of gradients: 0.08186852812979394
12 norm of weights: 2.2072983242718784
-----
Iteration Number: 6803
Loss: 11.48510073180867
12 norm of gradients: 0.08185143602655434
12 norm of weights: 2.2073026209653785
-----
Iteration Number: 6804

```

Loss: 11.484544290447396
l2 norm of gradients: 0.08183435165076498
l2 norm of weights: 2.207306948669997

Iteration Number: 6805
Loss: 11.48398818411849
l2 norm of gradients: 0.08181727499736413
l2 norm of weights: 2.2073113073615565

Iteration Number: 6806
Loss: 11.483432412590336
l2 norm of gradients: 0.08180020606129454
l2 norm of weights: 2.207315697015894

Iteration Number: 6807
Loss: 11.482876975631484
l2 norm of gradients: 0.08178314483750339
l2 norm of weights: 2.2073201176088655

Iteration Number: 6808
Loss: 11.482321873010672
l2 norm of gradients: 0.08176609132094216
l2 norm of weights: 2.2073245691163437

Iteration Number: 6809
Loss: 11.481767104496814
l2 norm of gradients: 0.0817490455065668
l2 norm of weights: 2.2073290515142165

Iteration Number: 6810
Loss: 11.481212669859
l2 norm of gradients: 0.08173200738933754
l2 norm of weights: 2.207333564778392

Iteration Number: 6811
Loss: 11.480658568866502
l2 norm of gradients: 0.08171497696421912
l2 norm of weights: 2.207338108884791

Iteration Number: 6812
Loss: 11.480104801288768
l2 norm of gradients: 0.08169795422618056
l2 norm of weights: 2.2073426838093546

Iteration Number: 6813
Loss: 11.47955136689542
l2 norm of gradients: 0.08168093917019523
l2 norm of weights: 2.207347289528039

```

-----
Iteration Number: 6814
Loss: 11.478998265456267
l2 norm of gradients: 0.08166393179124097
l2 norm of weights: 2.207351926016817
-----
Iteration Number: 6815
Loss: 11.478445496741287
l2 norm of gradients: 0.08164693208429985
l2 norm of weights: 2.2073565932516797
-----
Iteration Number: 6816
Loss: 11.477893060520639
l2 norm of gradients: 0.08162994004435845
l2 norm of weights: 2.2073612912086333
-----
Iteration Number: 6817
Loss: 11.477340956564664
l2 norm of gradients: 0.08161295566640756
l2 norm of weights: 2.207366019863702
-----
Iteration Number: 6818
Loss: 11.47678918464387
l2 norm of gradients: 0.08159597894544242
l2 norm of weights: 2.2073707791929262
-----
Iteration Number: 6819
Loss: 11.47623774452895
l2 norm of gradients: 0.08157900987646258
l2 norm of weights: 2.2073755691723633
-----
Iteration Number: 6820
Loss: 11.475686635990769
l2 norm of gradients: 0.0815620484544719
l2 norm of weights: 2.2073803897780873
-----
Iteration Number: 6821
Loss: 11.475135858800373
l2 norm of gradients: 0.08154509467447868
l2 norm of weights: 2.207385240986189
-----
Iteration Number: 6822
Loss: 11.47458541272898
l2 norm of gradients: 0.08152814853149543
l2 norm of weights: 2.207390122772776
-----
Iteration Number: 6823
Loss: 11.474035297547985

```

```

12 norm of gradients: 0.08151121002053902
12 norm of weights: 2.207395035113973
-----
Iteration Number: 6824
Loss: 11.473485513028969
12 norm of gradients: 0.08149427913663077
12 norm of weights: 2.2073999779859204
-----
Iteration Number: 6825
Loss: 11.47293605894367
12 norm of gradients: 0.0814773558747962
12 norm of weights: 2.2074049513647758
-----
Iteration Number: 6826
Loss: 11.47238693506402
12 norm of gradients: 0.08146044023006517
12 norm of weights: 2.2074099552267143
-----
Iteration Number: 6827
Loss: 11.471838141162113
12 norm of gradients: 0.0814435321974718
12 norm of weights: 2.2074149895479263
-----
Iteration Number: 6828
Loss: 11.471289677010228
12 norm of gradients: 0.08142663177205467
12 norm of weights: 2.20742005430462
-----
Iteration Number: 6829
Loss: 11.470741542380818
12 norm of gradients: 0.08140973894885656
12 norm of weights: 2.2074251494730195
-----
Iteration Number: 6830
Loss: 11.470193737046504
12 norm of gradients: 0.08139285372292458
12 norm of weights: 2.2074302750293655
-----
Iteration Number: 6831
Loss: 11.46964626078009
12 norm of gradients: 0.08137597608931015
12 norm of weights: 2.2074354309499165
-----
Iteration Number: 6832
Loss: 11.46909911335455
12 norm of gradients: 0.08135910604306895
12 norm of weights: 2.2074406172109464
-----

```

Iteration Number: 6833
Loss: 11.46855229454304
l2 norm of gradients: 0.08134224357926098
l2 norm of weights: 2.2074458337887464

Iteration Number: 6834
Loss: 11.468005804118878
l2 norm of gradients: 0.08132538869295057
l2 norm of weights: 2.2074510806596237

Iteration Number: 6835
Loss: 11.467459641855568
l2 norm of gradients: 0.0813085413792062
l2 norm of weights: 2.2074563577999027

Iteration Number: 6836
Loss: 11.466913807526781
l2 norm of gradients: 0.08129170163310084
l2 norm of weights: 2.2074616651859236

Iteration Number: 6837
Loss: 11.466368300906367
l2 norm of gradients: 0.08127486944971156
l2 norm of weights: 2.207467002794045

Iteration Number: 6838
Loss: 11.465823121768347
l2 norm of gradients: 0.0812580448241197
l2 norm of weights: 2.2074723706006387

Iteration Number: 6839
Loss: 11.465278269886918
l2 norm of gradients: 0.08124122775141106
l2 norm of weights: 2.207477768582097

Iteration Number: 6840
Loss: 11.464733745036446
l2 norm of gradients: 0.08122441822667543
l2 norm of weights: 2.207483196714826

Iteration Number: 6841
Loss: 11.464189546991477
l2 norm of gradients: 0.08120761624500712
l2 norm of weights: 2.2074886549752493

Iteration Number: 6842
Loss: 11.463645675526722
l2 norm of gradients: 0.08119082180150454


```
l2 norm of weights: 2.207494143339807
-----
Iteration Number: 6843
Loss: 11.463102130417074
l2 norm of gradients: 0.08117403489127041
l2 norm of weights: 2.2074996617849556
-----
Iteration Number: 6844
Loss: 11.46255891143759
l2 norm of gradients: 0.0811572555094117
l2 norm of weights: 2.207505210287168
-----
Iteration Number: 6845
Loss: 11.46201601836351
l2 norm of gradients: 0.0811404836510396
l2 norm of weights: 2.207510788822934
-----
Iteration Number: 6846
Loss: 11.461473450970239
l2 norm of gradients: 0.0811237193112695
l2 norm of weights: 2.2075163973687597
-----
Iteration Number: 6847
Loss: 11.460931209033358
l2 norm of gradients: 0.08110696248522117
l2 norm of weights: 2.207522035901167
-----
Iteration Number: 6848
Loss: 11.460389292328616
l2 norm of gradients: 0.0810902131680185
l2 norm of weights: 2.207527704396695
-----
Iteration Number: 6849
Loss: 11.459847700631936
l2 norm of gradients: 0.08107347135478961
l2 norm of weights: 2.207533402831899
-----
Iteration Number: 6850
Loss: 11.459306433719417
l2 norm of gradients: 0.0810567370406669
l2 norm of weights: 2.2075391311833514
-----
Iteration Number: 6851
Loss: 11.458765491367327
l2 norm of gradients: 0.08104001022078698
l2 norm of weights: 2.2075448894276404
-----
Iteration Number: 6852
```

Loss: 11.458224873352101
l2 norm of gradients: 0.08102329089029064
l2 norm of weights: 2.2075506775413696

Iteration Number: 6853
Loss: 11.45768457945036
l2 norm of gradients: 0.08100657904432296
l2 norm of weights: 2.2075564955011613

Iteration Number: 6854
Loss: 11.457144609438881
l2 norm of gradients: 0.08098987467803312
l2 norm of weights: 2.2075623432836524

Iteration Number: 6855
Loss: 11.456604963094614
l2 norm of gradients: 0.0809731777865746
l2 norm of weights: 2.2075682208654968

Iteration Number: 6856
Loss: 11.456065640194689
l2 norm of gradients: 0.08095648836510505
l2 norm of weights: 2.207574128223364

Iteration Number: 6857
Loss: 11.455526640516396
l2 norm of gradients: 0.08093980640878631
l2 norm of weights: 2.207580065333942

Iteration Number: 6858
Loss: 11.454987963837207
l2 norm of gradients: 0.08092313191278445
l2 norm of weights: 2.207586032173933

Iteration Number: 6859
Loss: 11.454449609934759
l2 norm of gradients: 0.08090646487226973
l2 norm of weights: 2.2075920287200557

Iteration Number: 6860
Loss: 11.453911578586855
l2 norm of gradients: 0.08088980528241657
l2 norm of weights: 2.2075980549490466

Iteration Number: 6861
Loss: 11.453373869571475
l2 norm of gradients: 0.08087315313840357
l2 norm of weights: 2.2076041108376576

```
-----  
Iteration Number: 6862  
Loss: 11.45283648266677  
l2 norm of gradients: 0.08085650843541352  
l2 norm of weights: 2.207610196362656  
-----  
Iteration Number: 6863  
Loss: 11.452299417651055  
l2 norm of gradients: 0.0808398711686334  
l2 norm of weights: 2.207616311500828  
-----  
Iteration Number: 6864  
Loss: 11.451762674302813  
l2 norm of gradients: 0.08082324133325434  
l2 norm of weights: 2.2076224562289726  
-----  
Iteration Number: 6865  
Loss: 11.451226252400708  
l2 norm of gradients: 0.08080661892447166  
l2 norm of weights: 2.2076286305239083  
-----  
Iteration Number: 6866  
Loss: 11.450690151723563  
l2 norm of gradients: 0.08079000393748485  
l2 norm of weights: 2.2076348343624677  
-----  
Iteration Number: 6867  
Loss: 11.45015437205038  
l2 norm of gradients: 0.08077339636749756  
l2 norm of weights: 2.2076410677215  
-----  
Iteration Number: 6868  
Loss: 11.449618913160318  
l2 norm of gradients: 0.08075679620971753  
l2 norm of weights: 2.2076473305778728  
-----  
Iteration Number: 6869  
Loss: 11.449083774832712  
l2 norm of gradients: 0.08074020345935674  
l2 norm of weights: 2.207653622908467  
-----  
Iteration Number: 6870  
Loss: 11.448548956847068  
l2 norm of gradients: 0.08072361811163135  
l2 norm of weights: 2.2076599446901812  
-----  
Iteration Number: 6871  
Loss: 11.448014458983058
```

```
12 norm of gradients: 0.08070704016176154
12 norm of weights: 2.207666295899929
-----
Iteration Number: 6872
Loss: 11.44748028102052
12 norm of gradients: 0.08069046960497167
12 norm of weights: 2.207672676514643
-----
Iteration Number: 6873
Loss: 11.446946422739462
12 norm of gradients: 0.08067390643649035
12 norm of weights: 2.2076790865112694
-----
Iteration Number: 6874
Loss: 11.446412883920067
12 norm of gradients: 0.08065735065155015
12 norm of weights: 2.207685525866771
-----
Iteration Number: 6875
Loss: 11.445879664342675
12 norm of gradients: 0.08064080224538796
12 norm of weights: 2.207691994558128
-----
Iteration Number: 6876
Loss: 11.4453467637878
12 norm of gradients: 0.0806242612132446
12 norm of weights: 2.2076984925623346
-----
Iteration Number: 6877
Loss: 11.444814182036126
12 norm of gradients: 0.08060772755036519
12 norm of weights: 2.207705019856403
-----
Iteration Number: 6878
Loss: 11.444281918868505
12 norm of gradients: 0.08059120125199883
12 norm of weights: 2.2077115764173616
-----
Iteration Number: 6879
Loss: 11.443749974065945
12 norm of gradients: 0.08057468231339883
12 norm of weights: 2.2077181622222546
-----
Iteration Number: 6880
Loss: 11.443218347409633
12 norm of gradients: 0.08055817072982255
12 norm of weights: 2.2077247772481408
-----
```

Iteration Number: 6881
Loss: 11.442687038680921
l2 norm of gradients: 0.08054166649653155
l2 norm of weights: 2.2077314214720962

Iteration Number: 6882
Loss: 11.442156047661326
l2 norm of gradients: 0.08052516960879133
l2 norm of weights: 2.2077380948712153

Iteration Number: 6883
Loss: 11.441625374132535
l2 norm of gradients: 0.08050868006187167
l2 norm of weights: 2.2077447974226043

Iteration Number: 6884
Loss: 11.4410950178764
l2 norm of gradients: 0.08049219785104636
l2 norm of weights: 2.2077515291033887

Iteration Number: 6885
Loss: 11.440564978674933
l2 norm of gradients: 0.08047572297159321
l2 norm of weights: 2.2077582898907084

Iteration Number: 6886
Loss: 11.44003525631033
l2 norm of gradients: 0.08045925541879428
l2 norm of weights: 2.2077650797617205

Iteration Number: 6887
Loss: 11.439505850564927
l2 norm of gradients: 0.08044279518793558
l2 norm of weights: 2.207771898693598

Iteration Number: 6888
Loss: 11.438976761221255
l2 norm of gradients: 0.08042634227430728
l2 norm of weights: 2.2077787466635286

Iteration Number: 6889
Loss: 11.438447988061991
l2 norm of gradients: 0.08040989667320358
l2 norm of weights: 2.207785623648718

Iteration Number: 6890
Loss: 11.437919530869983
l2 norm of gradients: 0.08039345837992279

```
l2 norm of weights: 2.207792529626386
-----
Iteration Number: 6891
Loss: 11.437391389428251
l2 norm of gradients: 0.08037702738976728
l2 norm of weights: 2.2077994645737697
-----
Iteration Number: 6892
Loss: 11.436863563519967
l2 norm of gradients: 0.08036060369804345
l2 norm of weights: 2.2078064284681225
-----
Iteration Number: 6893
Loss: 11.436336052928485
l2 norm of gradients: 0.08034418730006178
l2 norm of weights: 2.207813421286713
-----
Iteration Number: 6894
Loss: 11.43580885743731
l2 norm of gradients: 0.08032777819113684
l2 norm of weights: 2.2078204430068245
-----
Iteration Number: 6895
Loss: 11.435281976830124
l2 norm of gradients: 0.08031137636658725
l2 norm of weights: 2.2078274936057594
-----
Iteration Number: 6896
Loss: 11.43475541089076
l2 norm of gradients: 0.08029498182173563
l2 norm of weights: 2.207834573060834
-----
Iteration Number: 6897
Loss: 11.434229159403234
l2 norm of gradients: 0.08027859455190874
l2 norm of weights: 2.20784168134938
-----
Iteration Number: 6898
Loss: 11.433703222151706
l2 norm of gradients: 0.08026221455243723
l2 norm of weights: 2.207848818448747
-----
Iteration Number: 6899
Loss: 11.433177598920519
l2 norm of gradients: 0.08024584181865595
l2 norm of weights: 2.207855984336299
-----
Iteration Number: 6900
```

Loss: 11.432652289494166
l2 norm of gradients: 0.08022947634590369
l2 norm of weights: 2.2078631789894168

Iteration Number: 6901
Loss: 11.432127293657317
l2 norm of gradients: 0.08021311812952332
l2 norm of weights: 2.207870402385496

Iteration Number: 6902
Loss: 11.43160261119479
l2 norm of gradients: 0.08019676716486171
l2 norm of weights: 2.207877654501949

Iteration Number: 6903
Loss: 11.43107824189159
l2 norm of gradients: 0.08018042344726976
l2 norm of weights: 2.2078849353162044

Iteration Number: 6904
Loss: 11.430554185532861
l2 norm of gradients: 0.0801640869721024
l2 norm of weights: 2.2078922448057057

Iteration Number: 6905
Loss: 11.430030441903927
l2 norm of gradients: 0.08014775773471855
l2 norm of weights: 2.207899582947913

Iteration Number: 6906
Loss: 11.429507010790267
l2 norm of gradients: 0.08013143573048115
l2 norm of weights: 2.207906949720302

Iteration Number: 6907
Loss: 11.42898389197753
l2 norm of gradients: 0.08011512095475723
l2 norm of weights: 2.2079143451003636

Iteration Number: 6908
Loss: 11.428461085251527
l2 norm of gradients: 0.08009881340291769
l2 norm of weights: 2.2079217690656057

Iteration Number: 6909
Loss: 11.427938590398224
l2 norm of gradients: 0.08008251307033755
l2 norm of weights: 2.207929221593552

```
-----  
Iteration Number: 6910  
Loss: 11.427416407203756  
l2 norm of gradients: 0.08006621995239571  
l2 norm of weights: 2.207936702661741  
-----  
Iteration Number: 6911  
Loss: 11.426894535454425  
l2 norm of gradients: 0.0800499340444752  
l2 norm of weights: 2.2079442122477273  
-----  
Iteration Number: 6912  
Loss: 11.42637297493669  
l2 norm of gradients: 0.0800336553419629  
l2 norm of weights: 2.207951750329082  
-----  
Iteration Number: 6913  
Loss: 11.42585172543717  
l2 norm of gradients: 0.0800173838402498  
l2 norm of weights: 2.2079593168833913  
-----  
Iteration Number: 6914  
Loss: 11.425330786742652  
l2 norm of gradients: 0.08000111953473081  
l2 norm of weights: 2.2079669118882572  
-----  
Iteration Number: 6915  
Loss: 11.424810158640085  
l2 norm of gradients: 0.07998486242080477  
l2 norm of weights: 2.207974535321298  
-----  
Iteration Number: 6916  
Loss: 11.424289840916574  
l2 norm of gradients: 0.07996861249387466  
l2 norm of weights: 2.2079821871601473  
-----  
Iteration Number: 6917  
Loss: 11.423769833359392  
l2 norm of gradients: 0.07995236974934725  
l2 norm of weights: 2.2079898673824547  
-----  
Iteration Number: 6918  
Loss: 11.423250135755971  
l2 norm of gradients: 0.07993613418263336  
l2 norm of weights: 2.207997575965885  
-----  
Iteration Number: 6919  
Loss: 11.422730747893908
```



```
12 norm of gradients: 0.07991990578914777
12 norm of weights: 2.208005312888119
-----
Iteration Number: 6920
Loss: 11.422211669560951
12 norm of gradients: 0.07990368456430927
12 norm of weights: 2.208013078126854
-----
Iteration Number: 6921
Loss: 11.421692900545025
12 norm of gradients: 0.07988747050354045
12 norm of weights: 2.2080208716598024
-----
Iteration Number: 6922
Loss: 11.421174440634202
12 norm of gradients: 0.07987126360226804
12 norm of weights: 2.2080286934646915
-----
Iteration Number: 6923
Loss: 11.420656289616728
12 norm of gradients: 0.07985506385592261
12 norm of weights: 2.208036543519266
-----
Iteration Number: 6924
Loss: 11.420138447280996
12 norm of gradients: 0.07983887125993866
12 norm of weights: 2.208044421801284
-----
Iteration Number: 6925
Loss: 11.419620913415566
12 norm of gradients: 0.07982268580975474
12 norm of weights: 2.2080523282885216
-----
Iteration Number: 6926
Loss: 11.419103687809162
12 norm of gradients: 0.07980650750081322
12 norm of weights: 2.208060262958769
-----
Iteration Number: 6927
Loss: 11.418586770250666
12 norm of gradients: 0.07979033632856052
12 norm of weights: 2.208068225789833
-----
Iteration Number: 6928
Loss: 11.418070160529117
12 norm of gradients: 0.07977417228844681
12 norm of weights: 2.2080762167595354
-----
```

Iteration Number: 6929
Loss: 11.417553858433717
l2 norm of gradients: 0.07975801537592637
l2 norm of weights: 2.208084235845714

Iteration Number: 6930
Loss: 11.417037863753832
l2 norm of gradients: 0.07974186558645738
l2 norm of weights: 2.2080922830262217

Iteration Number: 6931
Loss: 11.416522176278976
l2 norm of gradients: 0.0797257229155018
l2 norm of weights: 2.2081003582789274

Iteration Number: 6932
Loss: 11.41600679579884
l2 norm of gradients: 0.07970958735852562
l2 norm of weights: 2.2081084615817157

Iteration Number: 6933
Loss: 11.415491722103253
l2 norm of gradients: 0.07969345891099876
l2 norm of weights: 2.208116592912487

Iteration Number: 6934
Loss: 11.414976954982222
l2 norm of gradients: 0.07967733756839494
l2 norm of weights: 2.208124752249156

Iteration Number: 6935
Loss: 11.414462494225909
l2 norm of gradients: 0.07966122332619192
l2 norm of weights: 2.208132939569655

Iteration Number: 6936
Loss: 11.413948339624625
l2 norm of gradients: 0.07964511617987126
l2 norm of weights: 2.2081411548519294

Iteration Number: 6937
Loss: 11.413434490968855
l2 norm of gradients: 0.07962901612491843
l2 norm of weights: 2.208149398073943

Iteration Number: 6938
Loss: 11.412920948049232
l2 norm of gradients: 0.07961292315682289

```
l2 norm of weights: 2.208157669213672
-----
Iteration Number: 6939
Loss: 11.412407710656549
l2 norm of gradients: 0.07959683727107776
l2 norm of weights: 2.2081659682491113
-----
Iteration Number: 6940
Loss: 11.411894778581763
l2 norm of gradients: 0.07958075846318034
l2 norm of weights: 2.208174295158268
-----
Iteration Number: 6941
Loss: 11.411382151615982
l2 norm of gradients: 0.07956468672863162
l2 norm of weights: 2.2081826499191686
-----
Iteration Number: 6942
Loss: 11.410869829550478
l2 norm of gradients: 0.07954862206293649
l2 norm of weights: 2.2081910325098515
-----
Iteration Number: 6943
Loss: 11.41035781217668
l2 norm of gradients: 0.07953256446160378
l2 norm of weights: 2.208199442908372
-----
Iteration Number: 6944
Loss: 11.409846099286177
l2 norm of gradients: 0.07951651392014611
l2 norm of weights: 2.2082078810928016
-----
Iteration Number: 6945
Loss: 11.409334690670711
l2 norm of gradients: 0.07950047043408
l2 norm of weights: 2.2082163470412257
-----
Iteration Number: 6946
Loss: 11.40882358612218
l2 norm of gradients: 0.07948443399892591
l2 norm of weights: 2.2082248407317473
-----
Iteration Number: 6947
Loss: 11.408312785432647
l2 norm of gradients: 0.07946840461020803
l2 norm of weights: 2.2082333621424826
-----
Iteration Number: 6948
```

```
Loss: 11.40780228839433
l2 norm of gradients: 0.07945238226345451
l2 norm of weights: 2.2082419112515645
-----
Iteration Number: 6949
Loss: 11.407292094799597
l2 norm of gradients: 0.07943636695419726
l2 norm of weights: 2.208250488037141
-----
Iteration Number: 6950
Loss: 11.406782204440988
l2 norm of gradients: 0.07942035867797209
l2 norm of weights: 2.2082590924773755
-----
Iteration Number: 6951
Loss: 11.406272617111185
l2 norm of gradients: 0.07940435743031873
l2 norm of weights: 2.2082677245504474
-----
Iteration Number: 6952
Loss: 11.405763332603035
l2 norm of gradients: 0.07938836320678058
l2 norm of weights: 2.20827638423455
-----
Iteration Number: 6953
Loss: 11.405254350709535
l2 norm of gradients: 0.07937237600290498
l2 norm of weights: 2.208285071507894
-----
Iteration Number: 6954
Loss: 11.40474567122385
l2 norm of gradients: 0.07935639581424317
l2 norm of weights: 2.2082937863487038
-----
Iteration Number: 6955
Loss: 11.404237293939294
l2 norm of gradients: 0.07934042263635005
l2 norm of weights: 2.2083025287352194
-----
Iteration Number: 6956
Loss: 11.403729218649335
l2 norm of gradients: 0.07932445646478453
l2 norm of weights: 2.2083112986456976
-----
Iteration Number: 6957
Loss: 11.4032214451476
l2 norm of gradients: 0.07930849729510922
l2 norm of weights: 2.208320096058409
```

```
-----  
Iteration Number: 6958  
Loss: 11.402713973227875  
l2 norm of gradients: 0.07929254512289055  
l2 norm of weights: 2.2083289209516397  
-----  
Iteration Number: 6959  
Loss: 11.4022068026841  
l2 norm of gradients: 0.07927659994369882  
l2 norm of weights: 2.208337773303692  
-----  
Iteration Number: 6960  
Loss: 11.401699933310363  
l2 norm of gradients: 0.07926066175310813  
l2 norm of weights: 2.2083466530928826  
-----  
Iteration Number: 6961  
Loss: 11.401193364900921  
l2 norm of gradients: 0.07924473054669637  
l2 norm of weights: 2.208355560297544  
-----  
Iteration Number: 6962  
Loss: 11.400687097250177  
l2 norm of gradients: 0.07922880632004524  
l2 norm of weights: 2.208364494896024  
-----  
Iteration Number: 6963  
Loss: 11.40018113015269  
l2 norm of gradients: 0.0792128890687402  
l2 norm of weights: 2.2083734568666857  
-----  
Iteration Number: 6964  
Loss: 11.399675463403183  
l2 norm of gradients: 0.07919697878837065  
l2 norm of weights: 2.2083824461879074  
-----  
Iteration Number: 6965  
Loss: 11.399170096796519  
l2 norm of gradients: 0.07918107547452961  
l2 norm of weights: 2.208391462838082  
-----  
Iteration Number: 6966  
Loss: 11.398665030127733  
l2 norm of gradients: 0.07916517912281396  
l2 norm of weights: 2.2084005067956185  
-----  
Iteration Number: 6967  
Loss: 11.398160263191997
```

```

12 norm of gradients: 0.0791492897288244
12 norm of weights: 2.2084095780389412
-----
Iteration Number: 6968
Loss: 11.397655795784653
12 norm of gradients: 0.0791334072881654
12 norm of weights: 2.2084186765464895
-----
Iteration Number: 6969
Loss: 11.39715162770119
12 norm of gradients: 0.0791175317964451
12 norm of weights: 2.2084278022967174
-----
Iteration Number: 6970
Loss: 11.39664775873725
12 norm of gradients: 0.07910166324927562
12 norm of weights: 2.208436955268095
-----
Iteration Number: 6971
Loss: 11.396144188688632
12 norm of gradients: 0.07908580164227266
12 norm of weights: 2.2084461354391074
-----
Iteration Number: 6972
Loss: 11.395640917351294
12 norm of gradients: 0.07906994697105581
12 norm of weights: 2.208455342788254
-----
Iteration Number: 6973
Loss: 11.395137944521334
12 norm of gradients: 0.07905409923124831
12 norm of weights: 2.208464577294051
-----
Iteration Number: 6974
Loss: 11.39463526999502
12 norm of gradients: 0.0790382584184773
12 norm of weights: 2.2084738389350282
-----
Iteration Number: 6975
Loss: 11.394132893568761
12 norm of gradients: 0.07902242452837357
12 norm of weights: 2.2084831276897314
-----
Iteration Number: 6976
Loss: 11.393630815039131
12 norm of gradients: 0.07900659755657173
12 norm of weights: 2.2084924435367217
-----

```

Iteration Number: 6977
Loss: 11.39312903420284
l2 norm of gradients: 0.07899077749871004
l2 norm of weights: 2.2085017864545753

Iteration Number: 6978
Loss: 11.392627550856773
l2 norm of gradients: 0.07897496435043064
l2 norm of weights: 2.2085111564218827

Iteration Number: 6979
Loss: 11.392126364797953
l2 norm of gradients: 0.0789591581073793
l2 norm of weights: 2.2085205534172503

Iteration Number: 6980
Loss: 11.391625475823561
l2 norm of gradients: 0.0789433587652056
l2 norm of weights: 2.2085299774192997

Iteration Number: 6981
Loss: 11.391124883730932
l2 norm of gradients: 0.07892756631956285
l2 norm of weights: 2.2085394284066675

Iteration Number: 6982
Loss: 11.39062458831755
l2 norm of gradients: 0.07891178076610805
l2 norm of weights: 2.208548906358005

Iteration Number: 6983
Loss: 11.390124589381049
l2 norm of gradients: 0.0788960021005019
l2 norm of weights: 2.208558411251979

Iteration Number: 6984
Loss: 11.38962488671923
l2 norm of gradients: 0.07888023031840898
l2 norm of weights: 2.208567943067271

Iteration Number: 6985
Loss: 11.389125480130028
l2 norm of gradients: 0.07886446541549741
l2 norm of weights: 2.208577501782578

Iteration Number: 6986
Loss: 11.388626369411543
l2 norm of gradients: 0.07884870738743908

```

12 norm of weights: 2.2085870873766127
-----
Iteration Number: 6987
Loss: 11.388127554362022
12 norm of gradients: 0.07883295622990967
12 norm of weights: 2.2085966998281013
-----
Iteration Number: 6988
Loss: 11.387629034779861
12 norm of gradients: 0.0788172119385885
12 norm of weights: 2.2086063391157853
-----
Iteration Number: 6989
Loss: 11.38713081046361
12 norm of gradients: 0.07880147450915857
12 norm of weights: 2.2086160052184223
-----
Iteration Number: 6990
Loss: 11.38663288121198
12 norm of gradients: 0.07878574393730665
12 norm of weights: 2.2086256981147847
-----
Iteration Number: 6991
Loss: 11.38613524682382
12 norm of gradients: 0.07877002021872324
12 norm of weights: 2.2086354177836593
-----
Iteration Number: 6992
Loss: 11.385637907098133
12 norm of gradients: 0.07875430334910233
12 norm of weights: 2.208645164203848
-----
Iteration Number: 6993
Loss: 11.38514086183408
12 norm of gradients: 0.07873859332414185
12 norm of weights: 2.208654937354168
-----
Iteration Number: 6994
Loss: 11.384644110830966
12 norm of gradients: 0.07872289013954334
12 norm of weights: 2.208664737213452
-----
Iteration Number: 6995
Loss: 11.384147653888252
12 norm of gradients: 0.07870719379101192
12 norm of weights: 2.2086745637605456
-----
Iteration Number: 6996

```


Loss: 11.383651490805548
l2 norm of gradients: 0.07869150427425649
l2 norm of weights: 2.2086844169743123

Iteration Number: 6997
Loss: 11.383155621382615
l2 norm of gradients: 0.07867582158498963
l2 norm of weights: 2.208694296833628

Iteration Number: 6998
Loss: 11.38266004541936
l2 norm of gradients: 0.07866014571892756
l2 norm of weights: 2.208704203317385

Iteration Number: 6999
Loss: 11.382164762715849
l2 norm of gradients: 0.0786444766717902
l2 norm of weights: 2.2087141364044904

Iteration Number: 7000
Loss: 11.381669773072291
l2 norm of gradients: 0.07862881443930111
l2 norm of weights: 2.2087240960738663

Iteration Number: 7001
Loss: 11.381175076289049
l2 norm of gradients: 0.07861315901718748
l2 norm of weights: 2.2087340823044483

Iteration Number: 7002
Loss: 11.38068067216663
l2 norm of gradients: 0.07859751040118021
l2 norm of weights: 2.2087440950751884

Iteration Number: 7003
Loss: 11.380186560505706
l2 norm of gradients: 0.07858186858701391
l2 norm of weights: 2.208754134365054

Iteration Number: 7004
Loss: 11.379692741107077
l2 norm of gradients: 0.07856623357042669
l2 norm of weights: 2.2087642001530257

Iteration Number: 7005
Loss: 11.37919921377171
l2 norm of gradients: 0.07855060534716045
l2 norm of weights: 2.2087742924180995

```

-----
Iteration Number: 7006
Loss: 11.378705978300717
l2 norm of gradients: 0.07853498391296063
l2 norm of weights: 2.2087844111392876
-----
Iteration Number: 7007
Loss: 11.378213034495358
l2 norm of gradients: 0.07851936926357643
l2 norm of weights: 2.208794556295615
-----
Iteration Number: 7008
Loss: 11.37772038215704
l2 norm of gradients: 0.07850376139476055
l2 norm of weights: 2.208804727866123
-----
Iteration Number: 7009
Loss: 11.377228021087317
l2 norm of gradients: 0.07848816030226943
l2 norm of weights: 2.2088149258298677
-----
Iteration Number: 7010
Loss: 11.376735951087905
l2 norm of gradients: 0.07847256598186311
l2 norm of weights: 2.208825150165919
-----
Iteration Number: 7011
Loss: 11.376244171960655
l2 norm of gradients: 0.07845697842930523
l2 norm of weights: 2.2088354008533626
-----
Iteration Number: 7012
Loss: 11.375752683507574
l2 norm of gradients: 0.07844139764036306
l2 norm of weights: 2.2088456778712984
-----
Iteration Number: 7013
Loss: 11.375261485530812
l2 norm of gradients: 0.07842582361080755
l2 norm of weights: 2.208855981198842
-----
Iteration Number: 7014
Loss: 11.374770577832676
l2 norm of gradients: 0.07841025633641317
l2 norm of weights: 2.2088663108151225
-----
Iteration Number: 7015
Loss: 11.374279960215613

```

```

12 norm of gradients: 0.07839469581295809
12 norm of weights: 2.208876666699285
-----
Iteration Number: 7016
Loss: 11.373789632482222
12 norm of gradients: 0.07837914203622404
12 norm of weights: 2.208887048830489
-----
Iteration Number: 7017
Loss: 11.373299594435245
12 norm of gradients: 0.07836359500199636
12 norm of weights: 2.2088974571879074
-----
Iteration Number: 7018
Loss: 11.372809845877583
12 norm of gradients: 0.078348054706064
12 norm of weights: 2.20890789175073
-----
Iteration Number: 7019
Loss: 11.372320386612277
12 norm of gradients: 0.07833252114421954
12 norm of weights: 2.208918352498161
-----
Iteration Number: 7020
Loss: 11.371831216442509
12 norm of gradients: 0.07831699431225904
12 norm of weights: 2.2089288394094173
-----
Iteration Number: 7021
Loss: 11.37134233517163
12 norm of gradients: 0.07830147420598234
12 norm of weights: 2.2089393524637333
-----
Iteration Number: 7022
Loss: 11.370853742603114
12 norm of gradients: 0.07828596082119266
12 norm of weights: 2.2089498916403563
-----
Iteration Number: 7023
Loss: 11.370365438540595
12 norm of gradients: 0.078270454153697
12 norm of weights: 2.2089604569185486
-----
Iteration Number: 7024
Loss: 11.369877422787853
12 norm of gradients: 0.07825495419930578
12 norm of weights: 2.2089710482775877
-----

```

Iteration Number: 7025
Loss: 11.369389695148817
l2 norm of gradients: 0.0782394609538331
l2 norm of weights: 2.2089816656967654

Iteration Number: 7026
Loss: 11.368902255427557
l2 norm of gradients: 0.07822397441309657
l2 norm of weights: 2.2089923091553882

Iteration Number: 7027
Loss: 11.36841510342829
l2 norm of gradients: 0.07820849457291744
l2 norm of weights: 2.2090029786327774

Iteration Number: 7028
Loss: 11.367928238955386
l2 norm of gradients: 0.07819302142912048
l2 norm of weights: 2.2090136741082693

Iteration Number: 7029
Loss: 11.367441661813357
l2 norm of gradients: 0.07817755497753398
l2 norm of weights: 2.209024395561214

Iteration Number: 7030
Loss: 11.366955371806863
l2 norm of gradients: 0.07816209521398988
l2 norm of weights: 2.209035142970977

Iteration Number: 7031
Loss: 11.366469368740713
l2 norm of gradients: 0.07814664213432361
l2 norm of weights: 2.209045916316938

Iteration Number: 7032
Loss: 11.36598365241985
l2 norm of gradients: 0.0781311957343742
l2 norm of weights: 2.2090567155784915

Iteration Number: 7033
Loss: 11.365498222649373
l2 norm of gradients: 0.07811575600998417
l2 norm of weights: 2.2090675407350466

Iteration Number: 7034
Loss: 11.365013079234533
l2 norm of gradients: 0.0781003229569997

```
l2 norm of weights: 2.2090783917660266
-----
Iteration Number: 7035
Loss: 11.364528221980713
l2 norm of gradients: 0.07808489657127034
l2 norm of weights: 2.2090892686508705
-----
Iteration Number: 7036
Loss: 11.364043650693448
l2 norm of gradients: 0.07806947684864932
l2 norm of weights: 2.2091001713690313
-----
Iteration Number: 7037
Loss: 11.363559365178423
l2 norm of gradients: 0.07805406378499335
l2 norm of weights: 2.2091110998999755
-----
Iteration Number: 7038
Loss: 11.363075365241457
l2 norm of gradients: 0.07803865737616268
l2 norm of weights: 2.2091220542231853
-----
Iteration Number: 7039
Loss: 11.362591650688522
l2 norm of gradients: 0.07802325761802109
l2 norm of weights: 2.209133034318158
-----
Iteration Number: 7040
Loss: 11.362108221325737
l2 norm of gradients: 0.07800786450643585
l2 norm of weights: 2.209144040164405
-----
Iteration Number: 7041
Loss: 11.36162507695936
l2 norm of gradients: 0.07799247803727781
l2 norm of weights: 2.2091550717414505
-----
Iteration Number: 7042
Loss: 11.3611422173958
l2 norm of gradients: 0.07797709820642133
l2 norm of weights: 2.209166129028836
-----
Iteration Number: 7043
Loss: 11.360659642441606
l2 norm of gradients: 0.07796172500974424
l2 norm of weights: 2.2091772120061153
-----
Iteration Number: 7044
```

Loss: 11.360177351903468
l2 norm of gradients: 0.07794635844312787
l2 norm of weights: 2.2091883206528586

Iteration Number: 7045
Loss: 11.359695345588229
l2 norm of gradients: 0.07793099850245713
l2 norm of weights: 2.209199454948649

Iteration Number: 7046
Loss: 11.359213623302875
l2 norm of gradients: 0.07791564518362042
l2 norm of weights: 2.209210614873085

Iteration Number: 7047
Loss: 11.358732184854532
l2 norm of gradients: 0.07790029848250958
l2 norm of weights: 2.209221800405779

Iteration Number: 7048
Loss: 11.358251030050472
l2 norm of gradients: 0.07788495839501998
l2 norm of weights: 2.209233011526358

Iteration Number: 7049
Loss: 11.357770158698107
l2 norm of gradients: 0.07786962491705046
l2 norm of weights: 2.2092442482144645

Iteration Number: 7050
Loss: 11.357289570605005
l2 norm of gradients: 0.07785429804450346
l2 norm of weights: 2.209255510449754

Iteration Number: 7051
Loss: 11.356809265578859
l2 norm of gradients: 0.07783897777328473
l2 norm of weights: 2.2092667982118974

Iteration Number: 7052
Loss: 11.356329243427526
l2 norm of gradients: 0.07782366409930365
l2 norm of weights: 2.2092781114805793

Iteration Number: 7053
Loss: 11.35584950395899
l2 norm of gradients: 0.07780835701847297
l2 norm of weights: 2.2092894502354996

```
-----
Iteration Number: 7054
Loss: 11.355370046981387
l2 norm of gradients: 0.07779305652670902
l2 norm of weights: 2.209300814456372
-----
Iteration Number: 7055
Loss: 11.354890872302994
l2 norm of gradients: 0.07777776261993154
l2 norm of weights: 2.209312204122924
-----
Iteration Number: 7056
Loss: 11.354411979732228
l2 norm of gradients: 0.0777624752940637
l2 norm of weights: 2.2093236192148997
-----
Iteration Number: 7057
Loss: 11.353933369077657
l2 norm of gradients: 0.07774719454503225
l2 norm of weights: 2.2093350597120547
-----
Iteration Number: 7058
Loss: 11.353455040147981
l2 norm of gradients: 0.0777319203687673
l2 norm of weights: 2.2093465255941616
-----
Iteration Number: 7059
Loss: 11.352976992752057
l2 norm of gradients: 0.07771665276120243
l2 norm of weights: 2.209358016841005
-----
Iteration Number: 7060
Loss: 11.352499226698866
l2 norm of gradients: 0.0777013917182748
l2 norm of weights: 2.2093695334323864
-----
Iteration Number: 7061
Loss: 11.352021741797548
l2 norm of gradients: 0.07768613723592478
l2 norm of weights: 2.209381075348119
-----
Iteration Number: 7062
Loss: 11.351544537857375
l2 norm of gradients: 0.07767088931009646
l2 norm of weights: 2.209392642568033
-----
Iteration Number: 7063
Loss: 11.351067614687764
```

l2 norm of gradients: 0.07765564793673714
l2 norm of weights: 2.2094042350719705

Iteration Number: 7064
Loss: 11.350590972098283
l2 norm of gradients: 0.07764041311179776
l2 norm of weights: 2.2094158528397894

Iteration Number: 7065
Loss: 11.350114609898625
l2 norm of gradients: 0.07762518483123251
l2 norm of weights: 2.2094274958513616

Iteration Number: 7066
Loss: 11.349638527898637
l2 norm of gradients: 0.07760996309099917
l2 norm of weights: 2.2094391640865734

Iteration Number: 7067
Loss: 11.349162725908307
l2 norm of gradients: 0.07759474788705881
l2 norm of weights: 2.209450857525325

Iteration Number: 7068
Loss: 11.348687203737757
l2 norm of gradients: 0.07757953921537614
l2 norm of weights: 2.209462576147531

Iteration Number: 7069
Loss: 11.348211961197261
l2 norm of gradients: 0.07756433707191907
l2 norm of weights: 2.2094743199331206

Iteration Number: 7070
Loss: 11.347736998097222
l2 norm of gradients: 0.077549141452659
l2 norm of weights: 2.209486088862037

Iteration Number: 7071
Loss: 11.347262314248198
l2 norm of gradients: 0.07753395235357081
l2 norm of weights: 2.209497882914238

Iteration Number: 7072
Loss: 11.346787909460877
l2 norm of gradients: 0.07751876977063274
l2 norm of weights: 2.209509702069695

Iteration Number: 7073
Loss: 11.346313783546096
l2 norm of gradients: 0.07750359369982643
l2 norm of weights: 2.209521546308394

Iteration Number: 7074
Loss: 11.345839936314821
l2 norm of gradients: 0.077488424137137
l2 norm of weights: 2.209533415610336

Iteration Number: 7075
Loss: 11.345366367578173
l2 norm of gradients: 0.07747326107855287
l2 norm of weights: 2.209545309955535

Iteration Number: 7076
Loss: 11.34489307714741
l2 norm of gradients: 0.07745810452006593
l2 norm of weights: 2.2095572293240195

Iteration Number: 7077
Loss: 11.344420064833919
l2 norm of gradients: 0.07744295445767145
l2 norm of weights: 2.209569173695833

Iteration Number: 7078
Loss: 11.343947330449245
l2 norm of gradients: 0.0774278108873681
l2 norm of weights: 2.209581143051032

Iteration Number: 7079
Loss: 11.343474873805055
l2 norm of gradients: 0.07741267380515789
l2 norm of weights: 2.209593137369688

Iteration Number: 7080
Loss: 11.343002694713173
l2 norm of gradients: 0.07739754320704627
l2 norm of weights: 2.209605156631887

Iteration Number: 7081
Loss: 11.34253079298555
l2 norm of gradients: 0.07738241908904211
l2 norm of weights: 2.209617200817728

Iteration Number: 7082
Loss: 11.342059168434286
l2 norm of gradients: 0.07736730144715757

```
l2 norm of weights: 2.209629269907325
-----
Iteration Number: 7083
Loss: 11.341587820871617
l2 norm of gradients: 0.07735219027740822
l2 norm of weights: 2.209641363880806
-----
Iteration Number: 7084
Loss: 11.341116750109913
l2 norm of gradients: 0.07733708557581305
l2 norm of weights: 2.209653482718313
-----
Iteration Number: 7085
Loss: 11.340645955961698
l2 norm of gradients: 0.07732198733839432
l2 norm of weights: 2.2096656264000027
-----
Iteration Number: 7086
Loss: 11.34017543823962
l2 norm of gradients: 0.07730689556117776
l2 norm of weights: 2.2096777949060447
-----
Iteration Number: 7087
Loss: 11.339705196756471
l2 norm of gradients: 0.0772918102401924
l2 norm of weights: 2.209689988216624
-----
Iteration Number: 7088
Loss: 11.339235231325187
l2 norm of gradients: 0.07727673137147063
l2 norm of weights: 2.2097022063119396
-----
Iteration Number: 7089
Loss: 11.338765541758843
l2 norm of gradients: 0.07726165895104822
l2 norm of weights: 2.209714449172203
-----
Iteration Number: 7090
Loss: 11.338296127870644
l2 norm of gradients: 0.07724659297496429
l2 norm of weights: 2.209726716777642
-----
Iteration Number: 7091
Loss: 11.337826989473943
l2 norm of gradients: 0.07723153343926133
l2 norm of weights: 2.209739009108497
-----
Iteration Number: 7092
```

Loss: 11.33735812638223
l2 norm of gradients: 0.07721648033998511
l2 norm of weights: 2.2097513261450232

Iteration Number: 7093
Loss: 11.336889538409125
l2 norm of gradients: 0.07720143367318481
l2 norm of weights: 2.209763667867489

Iteration Number: 7094
Loss: 11.336421225368396
l2 norm of gradients: 0.07718639343491288
l2 norm of weights: 2.2097760342561776

Iteration Number: 7095
Loss: 11.33595318707395
l2 norm of gradients: 0.07717135962122519
l2 norm of weights: 2.209788425291387

Iteration Number: 7096
Loss: 11.335485423339822
l2 norm of gradients: 0.07715633222818084
l2 norm of weights: 2.2098008409534273

Iteration Number: 7097
Loss: 11.335017933980195
l2 norm of gradients: 0.07714131125184236
l2 norm of weights: 2.209813281222624

Iteration Number: 7098
Loss: 11.334550718809393
l2 norm of gradients: 0.07712629668827559
l2 norm of weights: 2.209825746079316

Iteration Number: 7099
Loss: 11.334083777641856
l2 norm of gradients: 0.07711128853354961
l2 norm of weights: 2.209838235503857

Iteration Number: 7100
Loss: 11.333617110292188
l2 norm of gradients: 0.07709628678373691
l2 norm of weights: 2.2098507494766135

Iteration Number: 7101
Loss: 11.33315071657512
l2 norm of gradients: 0.07708129143491323
l2 norm of weights: 2.2098632879779667

```
-----
Iteration Number: 7102
Loss: 11.332684596305517
l2 norm of gradients: 0.07706630248315767
l2 norm of weights: 2.209875850988312
-----
Iteration Number: 7103
Loss: 11.332218749298384
l2 norm of gradients: 0.07705131992455261
l2 norm of weights: 2.209888438488059
-----
Iteration Number: 7104
Loss: 11.331753175368867
l2 norm of gradients: 0.07703634375518376
l2 norm of weights: 2.2099010504576295
-----
Iteration Number: 7105
Loss: 11.331287874332236
l2 norm of gradients: 0.0770213739711401
l2 norm of weights: 2.2099136868774614
-----
Iteration Number: 7106
Loss: 11.330822846003919
l2 norm of gradients: 0.07700641056851391
l2 norm of weights: 2.209926347728005
-----
Iteration Number: 7107
Loss: 11.330358090199466
l2 norm of gradients: 0.07699145354340084
l2 norm of weights: 2.209939032989726
-----
Iteration Number: 7108
Loss: 11.329893606734565
l2 norm of gradients: 0.0769765028918997
l2 norm of weights: 2.209951742643103
-----
Iteration Number: 7109
Loss: 11.329429395425043
l2 norm of gradients: 0.07696155861011275
l2 norm of weights: 2.2099644766686275
-----
Iteration Number: 7110
Loss: 11.328965456086868
l2 norm of gradients: 0.07694662069414533
l2 norm of weights: 2.209977235046808
-----
Iteration Number: 7111
Loss: 11.32850178853613
```

```

12 norm of gradients: 0.0769316891401063
12 norm of weights: 2.2099900177581637
-----
Iteration Number: 7112
Loss: 11.328038392589079
12 norm of gradients: 0.07691676394410758
12 norm of weights: 2.2100028247832295
-----
Iteration Number: 7113
Loss: 11.327575268062075
12 norm of gradients: 0.07690184510226454
12 norm of weights: 2.2100156561025535
-----
Iteration Number: 7114
Loss: 11.327112414771632
12 norm of gradients: 0.07688693261069565
12 norm of weights: 2.2100285116966982
-----
Iteration Number: 7115
Loss: 11.32664983253439
12 norm of gradients: 0.07687202646552284
12 norm of weights: 2.210041391546239
-----
Iteration Number: 7116
Loss: 11.326187521167135
12 norm of gradients: 0.07685712666287117
12 norm of weights: 2.2100542956317666
-----
Iteration Number: 7117
Loss: 11.325725480486776
12 norm of gradients: 0.07684223319886901
12 norm of weights: 2.210067223933884
-----
Iteration Number: 7118
Loss: 11.32526371031037
12 norm of gradients: 0.07682734606964797
12 norm of weights: 2.210080176433209
-----
Iteration Number: 7119
Loss: 11.324802210455102
12 norm of gradients: 0.07681246527134292
12 norm of weights: 2.2100931531103725
-----
Iteration Number: 7120
Loss: 11.324340980738294
12 norm of gradients: 0.07679759080009199
12 norm of weights: 2.210106153946021
-----

```

Iteration Number: 7121
Loss: 11.323880020977398
l2 norm of gradients: 0.07678272265203659
l2 norm of weights: 2.210119178920812

Iteration Number: 7122
Loss: 11.323419330990014
l2 norm of gradients: 0.0767678608233213
l2 norm of weights: 2.2101322280154196

Iteration Number: 7123
Loss: 11.322958910593862
l2 norm of gradients: 0.07675300531009399
l2 norm of weights: 2.210145301210529

Iteration Number: 7124
Loss: 11.32249875960681
l2 norm of gradients: 0.07673815610850578
l2 norm of weights: 2.210158398486842

Iteration Number: 7125
Loss: 11.322038877846854
l2 norm of gradients: 0.076723313214711
l2 norm of weights: 2.2101715198250718

Iteration Number: 7126
Loss: 11.321579265132122
l2 norm of gradients: 0.07670847662486721
l2 norm of weights: 2.2101846652059467

Iteration Number: 7127
Loss: 11.321119921280886
l2 norm of gradients: 0.07669364633513526
l2 norm of weights: 2.210197834610208

Iteration Number: 7128
Loss: 11.32066084611154
l2 norm of gradients: 0.07667882234167911
l2 norm of weights: 2.2102110280186116

Iteration Number: 7129
Loss: 11.320202039442622
l2 norm of gradients: 0.07666400464066604
l2 norm of weights: 2.2102242454119265

Iteration Number: 7130
Loss: 11.3197435010928
l2 norm of gradients: 0.07664919322826655

```
l2 norm of weights: 2.2102374867709353
-----
Iteration Number: 7131
Loss: 11.319285230880876
l2 norm of gradients: 0.07663438810065426
l2 norm of weights: 2.2102507520764347
-----
Iteration Number: 7132
Loss: 11.31882722862579
l2 norm of gradients: 0.0766195892540061
l2 norm of weights: 2.210264041309235
-----
Iteration Number: 7133
Loss: 11.31836949414661
l2 norm of gradients: 0.07660479668450218
l2 norm of weights: 2.2102773544501604
-----
Iteration Number: 7134
Loss: 11.31791202726254
l2 norm of gradients: 0.07659001038832584
l2 norm of weights: 2.2102906914800484
-----
Iteration Number: 7135
Loss: 11.317454827792925
l2 norm of gradients: 0.07657523036166353
l2 norm of weights: 2.21030405237975
-----
Iteration Number: 7136
Loss: 11.316997895557227
l2 norm of gradients: 0.07656045660070497
l2 norm of weights: 2.2103174371301315
-----
Iteration Number: 7137
Loss: 11.316541230375051
l2 norm of gradients: 0.07654568910164314
l2 norm of weights: 2.21033084571207
-----
Iteration Number: 7138
Loss: 11.316084832066144
l2 norm of gradients: 0.07653092786067409
l2 norm of weights: 2.2103442781064593
-----
Iteration Number: 7139
Loss: 11.315628700450372
l2 norm of gradients: 0.07651617287399713
l2 norm of weights: 2.2103577342942047
-----
Iteration Number: 7140
```

Loss: 11.315172835347738
l2 norm of gradients: 0.07650142413781473
l2 norm of weights: 2.2103712142562255

Iteration Number: 7141
Loss: 11.314717236578378
l2 norm of gradients: 0.07648668164833256
l2 norm of weights: 2.2103847179734557

Iteration Number: 7142
Loss: 11.314261903962567
l2 norm of gradients: 0.07647194540175949
l2 norm of weights: 2.210398245426842

Iteration Number: 7143
Loss: 11.313806837320705
l2 norm of gradients: 0.07645721539430747
l2 norm of weights: 2.210411796597345

Iteration Number: 7144
Loss: 11.31335203647333
l2 norm of gradients: 0.07644249162219176
l2 norm of weights: 2.2104253714659388

Iteration Number: 7145
Loss: 11.312897501241105
l2 norm of gradients: 0.07642777408163068
l2 norm of weights: 2.210438970013611

Iteration Number: 7146
Loss: 11.31244323144483
l2 norm of gradients: 0.07641306276884581
l2 norm of weights: 2.2104525922213627

Iteration Number: 7147
Loss: 11.311989226905439
l2 norm of gradients: 0.07639835768006184
l2 norm of weights: 2.210466238070209

Iteration Number: 7148
Loss: 11.311535487443999
l2 norm of gradients: 0.07638365881150656
l2 norm of weights: 2.210479907541179

Iteration Number: 7149
Loss: 11.311082012881705
l2 norm of gradients: 0.07636896615941106
l2 norm of weights: 2.2104936006153135


```
-----  
Iteration Number: 7150  
Loss: 11.310628803039885  
l2 norm of gradients: 0.07635427972000948  
l2 norm of weights: 2.210507317273669  
-----  
Iteration Number: 7151  
Loss: 11.310175857739997  
l2 norm of gradients: 0.07633959948953915  
l2 norm of weights: 2.210521057497314  
-----  
Iteration Number: 7152  
Loss: 11.309723176803633  
l2 norm of gradients: 0.07632492546424051  
l2 norm of weights: 2.210534821267332  
-----  
Iteration Number: 7153  
Loss: 11.309270760052518  
l2 norm of gradients: 0.07631025764035718  
l2 norm of weights: 2.2105486085648183  
-----  
Iteration Number: 7154  
Loss: 11.308818607308508  
l2 norm of gradients: 0.07629559601413594  
l2 norm of weights: 2.210562419370883  
-----  
Iteration Number: 7155  
Loss: 11.308366718393582  
l2 norm of gradients: 0.07628094058182663  
l2 norm of weights: 2.210576253666649  
-----  
Iteration Number: 7156  
Loss: 11.307915093129868  
l2 norm of gradients: 0.07626629133968231  
l2 norm of weights: 2.210590111433253  
-----  
Iteration Number: 7157  
Loss: 11.307463731339606  
l2 norm of gradients: 0.07625164828395914  
l2 norm of weights: 2.2106039926518455  
-----  
Iteration Number: 7158  
Loss: 11.307012632845176  
l2 norm of gradients: 0.07623701141091636  
l2 norm of weights: 2.21061789730359  
-----  
Iteration Number: 7159  
Loss: 11.306561797469094
```

```
l2 norm of gradients: 0.07622238071681646
l2 norm of weights: 2.2106318253696644
-----
Iteration Number: 7160
Loss: 11.306111225033993
l2 norm of gradients: 0.0762077561979249
l2 norm of weights: 2.210645776831258
-----
Iteration Number: 7161
Loss: 11.305660915362653
l2 norm of gradients: 0.07619313785051038
l2 norm of weights: 2.210659751669575
-----
Iteration Number: 7162
Loss: 11.30521086827797
l2 norm of gradients: 0.07617852567084464
l2 norm of weights: 2.210673749865834
-----
Iteration Number: 7163
Loss: 11.30476108360298
l2 norm of gradients: 0.07616391965520256
l2 norm of weights: 2.2106877714012647
-----
Iteration Number: 7164
Loss: 11.30431156116084
l2 norm of gradients: 0.07614931979986213
l2 norm of weights: 2.2107018162571124
-----
Iteration Number: 7165
Loss: 11.303862300774847
l2 norm of gradients: 0.07613472610110444
l2 norm of weights: 2.210715884414634
-----
Iteration Number: 7166
Loss: 11.303413302268428
l2 norm of gradients: 0.07612013855521371
l2 norm of weights: 2.210729975855102
-----
Iteration Number: 7167
Loss: 11.302964565465134
l2 norm of gradients: 0.07610555715847721
l2 norm of weights: 2.2107440905597997
-----
Iteration Number: 7168
Loss: 11.302516090188643
l2 norm of gradients: 0.07609098190718534
l2 norm of weights: 2.2107582285100253
-----
```

Iteration Number: 7169
Loss: 11.302067876262775
l2 norm of gradients: 0.07607641279763162
l2 norm of weights: 2.210772389687091

Iteration Number: 7170
Loss: 11.301619923511467
l2 norm of gradients: 0.07606184982611253
l2 norm of weights: 2.2107865740723205

Iteration Number: 7171
Loss: 11.301172231758793
l2 norm of gradients: 0.07604729298892786
l2 norm of weights: 2.2108007816470527

Iteration Number: 7172
Loss: 11.300724800828954
l2 norm of gradients: 0.0760327422823803
l2 norm of weights: 2.210815012392639

Iteration Number: 7173
Loss: 11.300277630546283
l2 norm of gradients: 0.07601819770277565
l2 norm of weights: 2.210829266290444

Iteration Number: 7174
Loss: 11.29983072073524
l2 norm of gradients: 0.07600365924642286
l2 norm of weights: 2.2108435433218454

Iteration Number: 7175
Loss: 11.29938407122041
l2 norm of gradients: 0.07598912690963387
l2 norm of weights: 2.2108578434682356

Iteration Number: 7176
Loss: 11.298937681826514
l2 norm of gradients: 0.07597460068872375
l2 norm of weights: 2.210872166711019

Iteration Number: 7177
Loss: 11.298491552378394
l2 norm of gradients: 0.07596008058001064
l2 norm of weights: 2.210886513031614

Iteration Number: 7178
Loss: 11.298045682701034
l2 norm of gradients: 0.07594556657981567

```

12 norm of weights: 2.2109008824114516
-----
Iteration Number: 7179
Loss: 11.297600072619533
12 norm of gradients: 0.07593105868446316
12 norm of weights: 2.2109152748319767
-----
Iteration Number: 7180
Loss: 11.297154721959128
12 norm of gradients: 0.07591655689028033
12 norm of weights: 2.210929690274648
-----
Iteration Number: 7181
Loss: 11.296709630545175
12 norm of gradients: 0.07590206119359762
12 norm of weights: 2.2109441287209353
-----
Iteration Number: 7182
Loss: 11.296264798203167
12 norm of gradients: 0.0758875715907484
12 norm of weights: 2.2109585901523245
-----
Iteration Number: 7183
Loss: 11.295820224758724
12 norm of gradients: 0.07587308807806913
12 norm of weights: 2.2109730745503136
-----
Iteration Number: 7184
Loss: 11.295375910037588
12 norm of gradients: 0.0758586106518993
12 norm of weights: 2.2109875818964126
-----
Iteration Number: 7185
Loss: 11.294931853865634
12 norm of gradients: 0.07584413930858147
12 norm of weights: 2.211002112172147
-----
Iteration Number: 7186
Loss: 11.294488056068863
12 norm of gradients: 0.07582967404446127
12 norm of weights: 2.2110166653590535
-----
Iteration Number: 7187
Loss: 11.29404451647341
12 norm of gradients: 0.07581521485588727
12 norm of weights: 2.2110312414386835
-----
Iteration Number: 7188

```

Loss: 11.293601234905527
l2 norm of gradients: 0.07580076173921113
l2 norm of weights: 2.2110458403926008

Iteration Number: 7189
Loss: 11.2931582111916
l2 norm of gradients: 0.07578631469078756
l2 norm of weights: 2.2110604622023824

Iteration Number: 7190
Loss: 11.292715445158144
l2 norm of gradients: 0.07577187370697427
l2 norm of weights: 2.211075106849619

Iteration Number: 7191
Loss: 11.292272936631795
l2 norm of gradients: 0.07575743878413203
l2 norm of weights: 2.211089774315915

Iteration Number: 7192
Loss: 11.291830685439322
l2 norm of gradients: 0.07574300991862457
l2 norm of weights: 2.2111044645828857

Iteration Number: 7193
Loss: 11.29138869140762
l2 norm of gradients: 0.07572858710681862
l2 norm of weights: 2.211119177632163

Iteration Number: 7194
Loss: 11.290946954363708
l2 norm of gradients: 0.07571417034508406
l2 norm of weights: 2.2111339134453885

Iteration Number: 7195
Loss: 11.29050547413474
l2 norm of gradients: 0.07569975962979364
l2 norm of weights: 2.211148672004219

Iteration Number: 7196
Loss: 11.290064250547983
l2 norm of gradients: 0.07568535495732318
l2 norm of weights: 2.2111634532903244

Iteration Number: 7197
Loss: 11.289623283430844
l2 norm of gradients: 0.07567095632405152
l2 norm of weights: 2.2111782572853866

```

-----
Iteration Number: 7198
Loss: 11.289182572610851
l2 norm of gradients: 0.07565656372636047
l2 norm of weights: 2.2111930839711027
-----
Iteration Number: 7199
Loss: 11.288742117915655
l2 norm of gradients: 0.07564217716063486
l2 norm of weights: 2.2112079333291805
-----
Iteration Number: 7200
Loss: 11.288301919173044
l2 norm of gradients: 0.07562779662326245
l2 norm of weights: 2.211222805341342
-----
Iteration Number: 7201
Loss: 11.287861976210918
l2 norm of gradients: 0.07561342211063413
l2 norm of weights: 2.211237699989323
-----
Iteration Number: 7202
Loss: 11.28742228885731
l2 norm of gradients: 0.07559905361914361
l2 norm of weights: 2.211252617254871
-----
Iteration Number: 7203
Loss: 11.28698285694039
l2 norm of gradients: 0.07558469114518775
l2 norm of weights: 2.211267557119748
-----
Iteration Number: 7204
Loss: 11.286543680288437
l2 norm of gradients: 0.07557033468516633
l2 norm of weights: 2.2112825195657284
-----
Iteration Number: 7205
Loss: 11.286104758729861
l2 norm of gradients: 0.075555984235482
l2 norm of weights: 2.2112975045745995
-----
Iteration Number: 7206
Loss: 11.2856660920932
l2 norm of gradients: 0.07554163979254055
l2 norm of weights: 2.211312512128161
-----
Iteration Number: 7207
Loss: 11.28522768020712

```

```
12 norm of gradients: 0.07552730135275068
12 norm of weights: 2.211327542208228
-----
Iteration Number: 7208
Loss: 11.284789522900407
12 norm of gradients: 0.07551296891252406
12 norm of weights: 2.2113425947966263
-----
Iteration Number: 7209
Loss: 11.284351620001976
12 norm of gradients: 0.07549864246827534
12 norm of weights: 2.2113576698751958
-----
Iteration Number: 7210
Loss: 11.283913971340866
12 norm of gradients: 0.07548432201642208
12 norm of weights: 2.2113727674257895
-----
Iteration Number: 7211
Loss: 11.283476576746239
12 norm of gradients: 0.07547000755338489
12 norm of weights: 2.2113878874302726
-----
Iteration Number: 7212
Loss: 11.283039436047389
12 norm of gradients: 0.07545569907558729
12 norm of weights: 2.211403029870524
-----
Iteration Number: 7213
Loss: 11.282602549073726
12 norm of gradients: 0.07544139657945574
12 norm of weights: 2.211418194728436
-----
Iteration Number: 7214
Loss: 11.282165915654794
12 norm of gradients: 0.07542710006141971
12 norm of weights: 2.211433381985913
-----
Iteration Number: 7215
Loss: 11.281729535620254
12 norm of gradients: 0.07541280951791156
12 norm of weights: 2.211448591624873
-----
Iteration Number: 7216
Loss: 11.281293408799893
12 norm of gradients: 0.07539852494536663
12 norm of weights: 2.2114638236272466
-----
```

Iteration Number: 7217
Loss: 11.28085753502363
l2 norm of gradients: 0.07538424634022317
l2 norm of weights: 2.2114790779749773

Iteration Number: 7218
Loss: 11.2804219141215
l2 norm of gradients: 0.07536997369892244
l2 norm of weights: 2.2114943546500223

Iteration Number: 7219
Loss: 11.279986545923668
l2 norm of gradients: 0.07535570701790859
l2 norm of weights: 2.2115096536343506

Iteration Number: 7220
Loss: 11.279551430260414
l2 norm of gradients: 0.07534144629362864
l2 norm of weights: 2.2115249749099455

Iteration Number: 7221
Loss: 11.279116566962157
l2 norm of gradients: 0.0753271915225327
l2 norm of weights: 2.2115403184588023

Iteration Number: 7222
Loss: 11.278681955859428
l2 norm of gradients: 0.07531294270107368
l2 norm of weights: 2.2115556842629296

Iteration Number: 7223
Loss: 11.278247596782883
l2 norm of gradients: 0.07529869982570743
l2 norm of weights: 2.211571072304348

Iteration Number: 7224
Loss: 11.277813489563307
l2 norm of gradients: 0.07528446289289278
l2 norm of weights: 2.211586482565093

Iteration Number: 7225
Loss: 11.277379634031611
l2 norm of gradients: 0.07527023189909145
l2 norm of weights: 2.2116019150272113

Iteration Number: 7226
Loss: 11.276946030018818
l2 norm of gradients: 0.07525600684076801


```

12 norm of weights: 2.2116173696727626
-----
Iteration Number: 7227
Loss: 11.276512677356086
12 norm of gradients: 0.07524178771439008
12 norm of weights: 2.21163284648382
-----
Iteration Number: 7228
Loss: 11.276079575874691
12 norm of gradients: 0.07522757451642809
12 norm of weights: 2.2116483454424705
-----
Iteration Number: 7229
Loss: 11.275646725406036
12 norm of gradients: 0.0752133672433554
12 norm of weights: 2.2116638665308117
-----
Iteration Number: 7230
Loss: 11.275214125781636
12 norm of gradients: 0.07519916589164827
12 norm of weights: 2.2116794097309556
-----
Iteration Number: 7231
Loss: 11.274781776833146
12 norm of gradients: 0.07518497045778587
12 norm of weights: 2.2116949750250265
-----
Iteration Number: 7232
Loss: 11.274349678392333
12 norm of gradients: 0.07517078093825028
12 norm of weights: 2.211710562395162
-----
Iteration Number: 7233
Loss: 11.27391783029109
12 norm of gradients: 0.07515659732952644
12 norm of weights: 2.2117261718235124
-----
Iteration Number: 7234
Loss: 11.273486232361432
12 norm of gradients: 0.07514241962810222
12 norm of weights: 2.2117418032922402
-----
Iteration Number: 7235
Loss: 11.273054884435496
12 norm of gradients: 0.07512824783046836
12 norm of weights: 2.211757456783521
-----
Iteration Number: 7236

```

Loss: 11.272623786345552
l2 norm of gradients: 0.07511408193311848
l2 norm of weights: 2.211773132279545

Iteration Number: 7237
Loss: 11.272192937923965
l2 norm of gradients: 0.07509992193254908
l2 norm of weights: 2.211788829762512

Iteration Number: 7238
Loss: 11.271762339003255
l2 norm of gradients: 0.0750857678252596
l2 norm of weights: 2.2118045492146368

Iteration Number: 7239
Loss: 11.271331989416048
l2 norm of gradients: 0.07507161960775227
l2 norm of weights: 2.2118202906181463

Iteration Number: 7240
Loss: 11.270901888995095
l2 norm of gradients: 0.07505747727653223
l2 norm of weights: 2.211836053955281

Iteration Number: 7241
Loss: 11.270472037573262
l2 norm of gradients: 0.07504334082810751
l2 norm of weights: 2.211851839208292

Iteration Number: 7242
Loss: 11.270042434983548
l2 norm of gradients: 0.07502921025898902
l2 norm of weights: 2.211867646359446

Iteration Number: 7243
Loss: 11.269613081059067
l2 norm of gradients: 0.07501508556569046
l2 norm of weights: 2.2118834753910206

Iteration Number: 7244
Loss: 11.269183975633059
l2 norm of gradients: 0.07500096674472846
l2 norm of weights: 2.211899326285307

Iteration Number: 7245
Loss: 11.268755118538884
l2 norm of gradients: 0.07498685379262256
l2 norm of weights: 2.211915199024608

```

-----
Iteration Number: 7246
Loss: 11.268326509610022
l2 norm of gradients: 0.07497274670589493
l2 norm of weights: 2.211931093591241
-----
Iteration Number: 7247
Loss: 11.267898148680077
l2 norm of gradients: 0.0749586454810709
l2 norm of weights: 2.2119470099675342
-----
Iteration Number: 7248
Loss: 11.26747003558277
l2 norm of gradients: 0.0749445501146784
l2 norm of weights: 2.2119629481358296
-----
Iteration Number: 7249
Loss: 11.267042170151948
l2 norm of gradients: 0.07493046060324839
l2 norm of weights: 2.2119789080784815
-----
Iteration Number: 7250
Loss: 11.266614552221576
l2 norm of gradients: 0.07491637694331447
l2 norm of weights: 2.2119948897778574
-----
Iteration Number: 7251
Loss: 11.266187181625748
l2 norm of gradients: 0.07490229913141337
l2 norm of weights: 2.2120108932163367
-----
Iteration Number: 7252
Loss: 11.265760058198662
l2 norm of gradients: 0.07488822716408437
l2 norm of weights: 2.2120269183763126
-----
Iteration Number: 7253
Loss: 11.265333181774658
l2 norm of gradients: 0.07487416103786972
l2 norm of weights: 2.21204296524019
-----
Iteration Number: 7254
Loss: 11.264906552188181
l2 norm of gradients: 0.07486010074931454
l2 norm of weights: 2.212059033790386
-----
Iteration Number: 7255
Loss: 11.264480169273803

```

```

12 norm of gradients: 0.07484604629496669
12 norm of weights: 2.212075124009332
-----
Iteration Number: 7256
Loss: 11.26405403286621
12 norm of gradients: 0.07483199767137691
12 norm of weights: 2.212091235879471
-----
Iteration Number: 7257
Loss: 11.263628142800224
12 norm of gradients: 0.07481795487509875
12 norm of weights: 2.212107369383259
-----
Iteration Number: 7258
Loss: 11.26320249891077
12 norm of gradients: 0.07480391790268857
12 norm of weights: 2.212123524503164
-----
Iteration Number: 7259
Loss: 11.262777101032908
12 norm of gradients: 0.0747898867507056
12 norm of weights: 2.212139701221667
-----
Iteration Number: 7260
Loss: 11.262351949001799
12 norm of gradients: 0.0747758614157118
12 norm of weights: 2.2121558995212625
-----
Iteration Number: 7261
Loss: 11.261927042652744
12 norm of gradients: 0.07476184189427197
12 norm of weights: 2.212172119384456
-----
Iteration Number: 7262
Loss: 11.261502381821154
12 norm of gradients: 0.07474782818295377
12 norm of weights: 2.2121883607937662
-----
Iteration Number: 7263
Loss: 11.261077966342564
12 norm of gradients: 0.07473382027832762
12 norm of weights: 2.212204623731725
-----
Iteration Number: 7264
Loss: 11.260653796052619
12 norm of gradients: 0.07471981817696678
12 norm of weights: 2.2122209081808766
-----

```

Iteration Number: 7265
Loss: 11.260229870787102
l2 norm of gradients: 0.07470582187544726
l2 norm of weights: 2.2122372141237774

Iteration Number: 7266
Loss: 11.259806190381891
l2 norm of gradients: 0.07469183137034785
l2 norm of weights: 2.212253541542997

Iteration Number: 7267
Loss: 11.259382754673013
l2 norm of gradients: 0.07467784665825025
l2 norm of weights: 2.2122698904211164

Iteration Number: 7268
Loss: 11.258959563496584
l2 norm of gradients: 0.07466386773573883
l2 norm of weights: 2.21228626074073

Iteration Number: 7269
Loss: 11.258536616688865
l2 norm of gradients: 0.07464989459940079
l2 norm of weights: 2.2123026524844454

Iteration Number: 7270
Loss: 11.258113914086216
l2 norm of gradients: 0.07463592724582618
l2 norm of weights: 2.2123190656348815

Iteration Number: 7271
Loss: 11.257691455525128
l2 norm of gradients: 0.07462196567160771
l2 norm of weights: 2.21233550017467

Iteration Number: 7272
Loss: 11.25726924084221
l2 norm of gradients: 0.07460800987334099
l2 norm of weights: 2.2123519560864553

Iteration Number: 7273
Loss: 11.25684726987419
l2 norm of gradients: 0.07459405984762431
l2 norm of weights: 2.212368433352895

Iteration Number: 7274
Loss: 11.256425542457901
l2 norm of gradients: 0.07458011559105876

```

12 norm of weights: 2.2123849319566578
-----
Iteration Number: 7275
Loss: 11.25600405843032
12 norm of gradients: 0.07456617710024827
12 norm of weights: 2.212401451880426
-----
Iteration Number: 7276
Loss: 11.255582817628522
12 norm of gradients: 0.07455224437179947
12 norm of weights: 2.212417993106893
-----
Iteration Number: 7277
Loss: 11.255161819889706
12 norm of gradients: 0.07453831740232172
12 norm of weights: 2.212434555618767
-----
Iteration Number: 7278
Loss: 11.254741065051194
12 norm of gradients: 0.07452439618842725
12 norm of weights: 2.2124511393987674
-----
Iteration Number: 7279
Loss: 11.254320552950421
12 norm of gradients: 0.07451048072673098
12 norm of weights: 2.212467744429625
-----
Iteration Number: 7280
Loss: 11.253900283424944
12 norm of gradients: 0.0744965710138506
12 norm of weights: 2.2124843706940838
-----
Iteration Number: 7281
Loss: 11.253480256312432
12 norm of gradients: 0.0744826670464065
12 norm of weights: 2.212501018174902
-----
Iteration Number: 7282
Loss: 11.25306047145068
12 norm of gradients: 0.07446876882102196
12 norm of weights: 2.2125176868548477
-----
Iteration Number: 7283
Loss: 11.2526409286776
12 norm of gradients: 0.07445487633432286
12 norm of weights: 2.2125343767167025
-----
Iteration Number: 7284

```

Loss: 11.25222162783121
l2 norm of gradients: 0.07444098958293792
l2 norm of weights: 2.2125510877432606

Iteration Number: 7285
Loss: 11.25180256874966
l2 norm of gradients: 0.07442710856349852
l2 norm of weights: 2.212567819917328

Iteration Number: 7286
Loss: 11.251383751271213
l2 norm of gradients: 0.07441323327263887
l2 norm of weights: 2.212584573221724

Iteration Number: 7287
Loss: 11.250965175234246
l2 norm of gradients: 0.07439936370699587
l2 norm of weights: 2.2126013476392794

Iteration Number: 7288
Loss: 11.250546840477258
l2 norm of gradients: 0.07438549986320915
l2 norm of weights: 2.2126181431528384

Iteration Number: 7289
Loss: 11.250128746838863
l2 norm of gradients: 0.07437164173792112
l2 norm of weights: 2.2126349597452557

Iteration Number: 7290
Loss: 11.24971089415779
l2 norm of gradients: 0.0743577893277768
l2 norm of weights: 2.212651797399401

Iteration Number: 7291
Loss: 11.249293282272887
l2 norm of gradients: 0.07434394262942408
l2 norm of weights: 2.212668656098154

Iteration Number: 7292
Loss: 11.248875911023124
l2 norm of gradients: 0.07433010163951347
l2 norm of weights: 2.2126855358244075

Iteration Number: 7293
Loss: 11.248458780247582
l2 norm of gradients: 0.07431626635469822
l2 norm of weights: 2.212702436561068

```

-----
Iteration Number: 7294
Loss: 11.248041889785458
l2 norm of gradients: 0.07430243677163435
l2 norm of weights: 2.2127193582910527
-----
Iteration Number: 7295
Loss: 11.24762523947607
l2 norm of gradients: 0.07428861288698056
l2 norm of weights: 2.212736300997291
-----
Iteration Number: 7296
Loss: 11.247208829158849
l2 norm of gradients: 0.07427479469739819
l2 norm of weights: 2.2127532646627257
-----
Iteration Number: 7297
Loss: 11.246792658673348
l2 norm of gradients: 0.07426098219955142
l2 norm of weights: 2.2127702492703123
-----
Iteration Number: 7298
Loss: 11.246376727859229
l2 norm of gradients: 0.07424717539010706
l2 norm of weights: 2.2127872548030165
-----
Iteration Number: 7299
Loss: 11.245961036556274
l2 norm of gradients: 0.07423337426573459
l2 norm of weights: 2.2128042812438182
-----
Iteration Number: 7300
Loss: 11.245545584604379
l2 norm of gradients: 0.07421957882310624
l2 norm of weights: 2.212821328575709
-----
Iteration Number: 7301
Loss: 11.245130371843564
l2 norm of gradients: 0.07420578905889695
l2 norm of weights: 2.212838396781692
-----
Iteration Number: 7302
Loss: 11.244715398113955
l2 norm of gradients: 0.07419200496978434
l2 norm of weights: 2.2128554858447838
-----
Iteration Number: 7303
Loss: 11.244300663255801

```



```

12 norm of gradients: 0.0741782265524487
12 norm of weights: 2.212872595748013
-----
Iteration Number: 7304
Loss: 11.243886167109462
12 norm of gradients: 0.07416445380357303
12 norm of weights: 2.2128897264744207
-----
Iteration Number: 7305
Loss: 11.243471909515415
12 norm of gradients: 0.07415068671984293
12 norm of weights: 2.212906878007059
-----
Iteration Number: 7306
Loss: 11.243057890314255
12 norm of gradients: 0.07413692529794685
12 norm of weights: 2.2129240503289926
-----
Iteration Number: 7307
Loss: 11.242644109346692
12 norm of gradients: 0.0741231695345758
12 norm of weights: 2.2129412434232996
-----
Iteration Number: 7308
Loss: 11.24223056645355
12 norm of gradients: 0.07410941942642352
12 norm of weights: 2.2129584572730696
-----
Iteration Number: 7309
Loss: 11.241817261475767
12 norm of gradients: 0.07409567497018633
12 norm of weights: 2.212975691861404
-----
Iteration Number: 7310
Loss: 11.241404194254399
12 norm of gradients: 0.07408193616256332
12 norm of weights: 2.212992947171417
-----
Iteration Number: 7311
Loss: 11.240991364630617
12 norm of gradients: 0.07406820300025628
12 norm of weights: 2.213010223186235
-----
Iteration Number: 7312
Loss: 11.240578772445707
12 norm of gradients: 0.0740544754799695
12 norm of weights: 2.2130275198889957
-----

```

Iteration Number: 7313
Loss: 11.240166417541065
l2 norm of gradients: 0.07404075359841014
l2 norm of weights: 2.2130448372628506

Iteration Number: 7314
Loss: 11.239754299758214
l2 norm of gradients: 0.07402703735228786
l2 norm of weights: 2.213062175290962

Iteration Number: 7315
Loss: 11.23934241893878
l2 norm of gradients: 0.07401332673831502
l2 norm of weights: 2.213079533956505

Iteration Number: 7316
Loss: 11.238930774924503
l2 norm of gradients: 0.07399962175320668
l2 norm of weights: 2.213096913242666

Iteration Number: 7317
Loss: 11.23851936755725
l2 norm of gradients: 0.0739859223936805
l2 norm of weights: 2.2131143131326447

Iteration Number: 7318
Loss: 11.23810819667899
l2 norm of gradients: 0.07397222865645685
l2 norm of weights: 2.213131733609653

Iteration Number: 7319
Loss: 11.237697262131812
l2 norm of gradients: 0.07395854053825868
l2 norm of weights: 2.2131491746569134

Iteration Number: 7320
Loss: 11.237286563757923
l2 norm of gradients: 0.07394485803581162
l2 norm of weights: 2.2131666362576623

Iteration Number: 7321
Loss: 11.236876101399634
l2 norm of gradients: 0.07393118114584393
l2 norm of weights: 2.2131841183951475

Iteration Number: 7322
Loss: 11.236465874899382
l2 norm of gradients: 0.07391750986508648

l2 norm of weights: 2.213201621052628

Iteration Number: 7323
Loss: 11.236055884099704
l2 norm of gradients: 0.07390384419027285
l2 norm of weights: 2.213219144213377

Iteration Number: 7324
Loss: 11.235646128843266
l2 norm of gradients: 0.07389018411813918
l2 norm of weights: 2.213236687860678

Iteration Number: 7325
Loss: 11.235236608972835
l2 norm of gradients: 0.07387652964542424
l2 norm of weights: 2.2132542519778267

Iteration Number: 7326
Loss: 11.234827324331304
l2 norm of gradients: 0.0738628807688695
l2 norm of weights: 2.2132718365481323

Iteration Number: 7327
Loss: 11.23441827476167
l2 norm of gradients: 0.07384923748521899
l2 norm of weights: 2.213289441554914

Iteration Number: 7328
Loss: 11.234009460107046
l2 norm of gradients: 0.07383559979121937
l2 norm of weights: 2.213307066981505

Iteration Number: 7329
Loss: 11.233600880210664
l2 norm of gradients: 0.07382196768361994
l2 norm of weights: 2.213324712811249

Iteration Number: 7330
Loss: 11.233192534915856
l2 norm of gradients: 0.07380834115917259
l2 norm of weights: 2.213342379027504

Iteration Number: 7331
Loss: 11.232784424066082
l2 norm of gradients: 0.07379472021463183
l2 norm of weights: 2.213360065613637

Iteration Number: 7332

Loss: 11.23237654750491
l2 norm of gradients: 0.07378110484675478
l2 norm of weights: 2.213377772553029

Iteration Number: 7333
Loss: 11.231968905076014
l2 norm of gradients: 0.0737674950523012
l2 norm of weights: 2.213395499829073

Iteration Number: 7334
Loss: 11.231561496623193
l2 norm of gradients: 0.07375389082803341
l2 norm of weights: 2.2134132474251733

Iteration Number: 7335
Loss: 11.231154321990358
l2 norm of gradients: 0.07374029217071634
l2 norm of weights: 2.2134310153247463

Iteration Number: 7336
Loss: 11.230747381021514
l2 norm of gradients: 0.07372669907711758
l2 norm of weights: 2.213448803511221

Iteration Number: 7337
Loss: 11.230340673560807
l2 norm of gradients: 0.07371311154400718
l2 norm of weights: 2.213466611968038

Iteration Number: 7338
Loss: 11.22993419945247
l2 norm of gradients: 0.07369952956815798
l2 norm of weights: 2.2134844406786494

Iteration Number: 7339
Loss: 11.229527958540867
l2 norm of gradients: 0.0736859531463452
l2 norm of weights: 2.2135022896265197

Iteration Number: 7340
Loss: 11.229121950670466
l2 norm of gradients: 0.07367238227534678
l2 norm of weights: 2.2135201587951263

Iteration Number: 7341
Loss: 11.22871617568585
l2 norm of gradients: 0.07365881695194326
l2 norm of weights: 2.2135380481679574

```

-----
Iteration Number: 7342
Loss: 11.228310633431706
l2 norm of gradients: 0.07364525717291766
l2 norm of weights: 2.2135559577285133
-----
Iteration Number: 7343
Loss: 11.227905323752848
l2 norm of gradients: 0.07363170293505569
l2 norm of weights: 2.2135738874603064
-----
Iteration Number: 7344
Loss: 11.227500246494195
l2 norm of gradients: 0.07361815423514553
l2 norm of weights: 2.213591837346861
-----
Iteration Number: 7345
Loss: 11.227095401500772
l2 norm of gradients: 0.07360461106997801
l2 norm of weights: 2.213609807371713
-----
Iteration Number: 7346
Loss: 11.22669078861772
l2 norm of gradients: 0.07359107343634652
l2 norm of weights: 2.2136277975184115
-----
Iteration Number: 7347
Loss: 11.2262864076903
l2 norm of gradients: 0.07357754133104702
l2 norm of weights: 2.213645807770516
-----
Iteration Number: 7348
Loss: 11.225882258563875
l2 norm of gradients: 0.07356401475087802
l2 norm of weights: 2.2136638381115987
-----
Iteration Number: 7349
Loss: 11.22547834108392
l2 norm of gradients: 0.07355049369264059
l2 norm of weights: 2.2136818885252434
-----
Iteration Number: 7350
Loss: 11.225074655096023
l2 norm of gradients: 0.0735369781531384
l2 norm of weights: 2.213699958995046
-----
Iteration Number: 7351
Loss: 11.224671200445888

```

```

12 norm of gradients: 0.07352346812917765
12 norm of weights: 2.2137180495046143
-----
Iteration Number: 7352
Loss: 11.22426797697933
12 norm of gradients: 0.07350996361756709
12 norm of weights: 2.213736160037568
-----
Iteration Number: 7353
Loss: 11.223864984542262
12 norm of gradients: 0.07349646461511801
12 norm of weights: 2.213754290577538
-----
Iteration Number: 7354
Loss: 11.223462222980727
12 norm of gradients: 0.0734829711186443
12 norm of weights: 2.213772441108168
-----
Iteration Number: 7355
Loss: 11.223059692140867
12 norm of gradients: 0.07346948312496235
12 norm of weights: 2.2137906116131125
-----
Iteration Number: 7356
Loss: 11.222657391868935
12 norm of gradients: 0.07345600063089115
12 norm of weights: 2.2138088020760396
-----
Iteration Number: 7357
Loss: 11.2222553220113
12 norm of gradients: 0.07344252363325217
12 norm of weights: 2.2138270124806274
-----
Iteration Number: 7358
Loss: 11.221853482414444
12 norm of gradients: 0.07342905212886944
12 norm of weights: 2.2138452428105664
-----
Iteration Number: 7359
Loss: 11.221451872924954
12 norm of gradients: 0.0734155861145696
12 norm of weights: 2.2138634930495598
-----
Iteration Number: 7360
Loss: 11.221050493389523
12 norm of gradients: 0.07340212558718162
12 norm of weights: 2.2138817631813215
-----

```

Iteration Number: 7361
Loss: 11.220649343654971
l2 norm of gradients: 0.07338867054353727
l2 norm of weights: 2.2139000531895774

Iteration Number: 7362
Loss: 11.220248423568211
l2 norm of gradients: 0.0733752209804707
l2 norm of weights: 2.2139183630580654

Iteration Number: 7363
Loss: 11.219847732976277
l2 norm of gradients: 0.07336177689481854
l2 norm of weights: 2.213936692770536

Iteration Number: 7364
Loss: 11.219447271726308
l2 norm of gradients: 0.07334833828342013
l2 norm of weights: 2.2139550423107495

Iteration Number: 7365
Loss: 11.219047039665554
l2 norm of gradients: 0.07333490514311704
l2 norm of weights: 2.21397341166248

Iteration Number: 7366
Loss: 11.218647036641379
l2 norm of gradients: 0.07332147747075368
l2 norm of weights: 2.213991800809512

Iteration Number: 7367
Loss: 11.218247262501256
l2 norm of gradients: 0.07330805526317674
l2 norm of weights: 2.2140102097356427

Iteration Number: 7368
Loss: 11.217847717092758
l2 norm of gradients: 0.07329463851723557
l2 norm of weights: 2.21402863842468

Iteration Number: 7369
Loss: 11.217448400263581
l2 norm of gradients: 0.07328122722978193
l2 norm of weights: 2.214047086860445

Iteration Number: 7370
Loss: 11.21704931186153
l2 norm of gradients: 0.07326782139767016

```
l2 norm of weights: 2.214065555026769
-----
Iteration Number: 7371
Loss: 11.216650451734502
l2 norm of gradients: 0.07325442101775699
l2 norm of weights: 2.214084042907496
-----
Iteration Number: 7372
Loss: 11.216251819730534
l2 norm of gradients: 0.07324102608690185
l2 norm of weights: 2.2141025504864813
-----
Iteration Number: 7373
Loss: 11.21585341569774
l2 norm of gradients: 0.07322763660196649
l2 norm of weights: 2.2141210777475924
-----
Iteration Number: 7374
Loss: 11.215455239484369
l2 norm of gradients: 0.07321425255981522
l2 norm of weights: 2.214139624674708
-----
Iteration Number: 7375
Loss: 11.215057290938764
l2 norm of gradients: 0.07320087395731488
l2 norm of weights: 2.2141581912517183
-----
Iteration Number: 7376
Loss: 11.214659569909385
l2 norm of gradients: 0.07318750079133472
l2 norm of weights: 2.214176777462526
-----
Iteration Number: 7377
Loss: 11.214262076244793
l2 norm of gradients: 0.07317413305874658
l2 norm of weights: 2.2141953832910444
-----
Iteration Number: 7378
Loss: 11.213864809793666
l2 norm of gradients: 0.0731607707564247
l2 norm of weights: 2.2142140087212003
-----
Iteration Number: 7379
Loss: 11.21346777040479
l2 norm of gradients: 0.07314741388124586
l2 norm of weights: 2.21423265373693
-----
Iteration Number: 7380
```


Loss: 11.213070957927059
l2 norm of gradients: 0.07313406243008926
l2 norm of weights: 2.2142513183221824

Iteration Number: 7381
Loss: 11.21267437220947
l2 norm of gradients: 0.07312071639983671
l2 norm of weights: 2.214270002460919

Iteration Number: 7382
Loss: 11.212278013101132
l2 norm of gradients: 0.07310737578737231
l2 norm of weights: 2.2142887061371113

Iteration Number: 7383
Loss: 11.211881880451273
l2 norm of gradients: 0.07309404058958277
l2 norm of weights: 2.2143074293347427

Iteration Number: 7384
Loss: 11.211485974109213
l2 norm of gradients: 0.07308071080335725
l2 norm of weights: 2.2143261720378096

Iteration Number: 7385
Loss: 11.211090293924386
l2 norm of gradients: 0.07306738642558731
l2 norm of weights: 2.2143449342303185

Iteration Number: 7386
Loss: 11.210694839746346
l2 norm of gradients: 0.07305406745316707
l2 norm of weights: 2.2143637158962886

Iteration Number: 7387
Loss: 11.210299611424732
l2 norm of gradients: 0.07304075388299305
l2 norm of weights: 2.2143825170197498

Iteration Number: 7388
Loss: 11.209904608809314
l2 norm of gradients: 0.07302744571196426
l2 norm of weights: 2.2144013375847442

Iteration Number: 7389
Loss: 11.209509831749958
l2 norm of gradients: 0.07301414293698215
l2 norm of weights: 2.2144201775753256

```

-----
Iteration Number: 7390
Loss: 11.20911528009664
l2 norm of gradients: 0.07300084555495062
l2 norm of weights: 2.214439036975558
-----
Iteration Number: 7391
Loss: 11.208720953699444
l2 norm of gradients: 0.07298755356277607
l2 norm of weights: 2.2144579157695197
-----
Iteration Number: 7392
Loss: 11.208326852408565
l2 norm of gradients: 0.07297426695736725
l2 norm of weights: 2.2144768139412982
-----
Iteration Number: 7393
Loss: 11.207932976074302
l2 norm of gradients: 0.0729609857356355
l2 norm of weights: 2.214495731474993
-----
Iteration Number: 7394
Loss: 11.207539324547056
l2 norm of gradients: 0.07294770989449453
l2 norm of weights: 2.214514668354715
-----
Iteration Number: 7395
Loss: 11.207145897677348
l2 norm of gradients: 0.07293443943086042
l2 norm of weights: 2.2145336245645884
-----
Iteration Number: 7396
Loss: 11.206752695315796
l2 norm of gradients: 0.07292117434165181
l2 norm of weights: 2.2145526000887465
-----
Iteration Number: 7397
Loss: 11.20635971731314
l2 norm of gradients: 0.07290791462378972
l2 norm of weights: 2.214571594911336
-----
Iteration Number: 7398
Loss: 11.205966963520202
l2 norm of gradients: 0.07289466027419762
l2 norm of weights: 2.214590609016514
-----
Iteration Number: 7399
Loss: 11.205574433787937

```

```

12 norm of gradients: 0.0728814112898014
12 norm of weights: 2.2146096423884494
-----
Iteration Number: 7400
Loss: 11.205182127967394
12 norm of gradients: 0.07286816766752938
12 norm of weights: 2.2146286950113225
-----
Iteration Number: 7401
Loss: 11.204790045909727
12 norm of gradients: 0.07285492940431233
12 norm of weights: 2.214647766869326
-----
Iteration Number: 7402
Loss: 11.204398187466204
12 norm of gradients: 0.0728416964970834
12 norm of weights: 2.214666857946663
-----
Iteration Number: 7403
Loss: 11.204006552488199
12 norm of gradients: 0.07282846894277824
12 norm of weights: 2.214685968227548
-----
Iteration Number: 7404
Loss: 11.20361514082719
12 norm of gradients: 0.0728152467383348
12 norm of weights: 2.2147050976962084
-----
Iteration Number: 7405
Loss: 11.203223952334756
12 norm of gradients: 0.07280202988069356
12 norm of weights: 2.214724246336881
-----
Iteration Number: 7406
Loss: 11.202832986862596
12 norm of gradients: 0.07278881836679736
12 norm of weights: 2.214743414133816
-----
Iteration Number: 7407
Loss: 11.202442244262505
12 norm of gradients: 0.0727756121935915
12 norm of weights: 2.214762601071274
-----
Iteration Number: 7408
Loss: 11.202051724386385
12 norm of gradients: 0.07276241135802358
12 norm of weights: 2.2147818071335275
-----

```

Iteration Number: 7409
Loss: 11.201661427086256
l2 norm of gradients: 0.07274921585704373
l2 norm of weights: 2.2148010323048593

Iteration Number: 7410
Loss: 11.201271352214226
l2 norm of gradients: 0.07273602568760441
l2 norm of weights: 2.214820276569565

Iteration Number: 7411
Loss: 11.200881499622524
l2 norm of gradients: 0.0727228408466605
l2 norm of weights: 2.2148395399119516

Iteration Number: 7412
Loss: 11.200491869163475
l2 norm of gradients: 0.07270966133116932
l2 norm of weights: 2.2148588223163364

Iteration Number: 7413
Loss: 11.200102460689518
l2 norm of gradients: 0.07269648713809052
l2 norm of weights: 2.214878123767049

Iteration Number: 7414
Loss: 11.199713274053192
l2 norm of gradients: 0.07268331826438616
l2 norm of weights: 2.2148974442484306

Iteration Number: 7415
Loss: 11.199324309107146
l2 norm of gradients: 0.07267015470702069
l2 norm of weights: 2.2149167837448327

Iteration Number: 7416
Loss: 11.19893556570413
l2 norm of gradients: 0.07265699646296105
l2 norm of weights: 2.214936142240619

Iteration Number: 7417
Loss: 11.198547043697003
l2 norm of gradients: 0.07264384352917638
l2 norm of weights: 2.2149555197201645

Iteration Number: 7418
Loss: 11.198158742938729
l2 norm of gradients: 0.07263069590263839

```
l2 norm of weights: 2.2149749161678556
-----
Iteration Number: 7419
Loss: 11.197770663282382
l2 norm of gradients: 0.07261755358032103
l2 norm of weights: 2.21499433156809
-----
Iteration Number: 7420
Loss: 11.19738280458113
l2 norm of gradients: 0.07260441655920068
l2 norm of weights: 2.215013765905276
-----
Iteration Number: 7421
Loss: 11.196995166688255
l2 norm of gradients: 0.07259128483625614
l2 norm of weights: 2.2150332191638347
-----
Iteration Number: 7422
Loss: 11.19660774945714
l2 norm of gradients: 0.07257815840846855
l2 norm of weights: 2.215052691328198
-----
Iteration Number: 7423
Loss: 11.19622055274128
l2 norm of gradients: 0.07256503727282136
l2 norm of weights: 2.2150721823828077
-----
Iteration Number: 7424
Loss: 11.195833576394268
l2 norm of gradients: 0.07255192142630051
l2 norm of weights: 2.2150916923121193
-----
Iteration Number: 7425
Loss: 11.195446820269808
l2 norm of gradients: 0.07253881086589418
l2 norm of weights: 2.215111221100598
-----
Iteration Number: 7426
Loss: 11.195060284221695
l2 norm of gradients: 0.07252570558859306
l2 norm of weights: 2.2151307687327213
-----
Iteration Number: 7427
Loss: 11.194673968103844
l2 norm of gradients: 0.07251260559138999
l2 norm of weights: 2.215150335192977
-----
Iteration Number: 7428
```

Loss: 11.194287871770275
l2 norm of gradients: 0.07249951087128041
l2 norm of weights: 2.215169920465864

Iteration Number: 7429
Loss: 11.193901995075102
l2 norm of gradients: 0.07248642142526197
l2 norm of weights: 2.215189524535894

Iteration Number: 7430
Loss: 11.193516337872547
l2 norm of gradients: 0.07247333725033465
l2 norm of weights: 2.2152091473875886

Iteration Number: 7431
Loss: 11.193130900016937
l2 norm of gradients: 0.07246025834350094
l2 norm of weights: 2.215228789005482

Iteration Number: 7432
Loss: 11.192745681362709
l2 norm of gradients: 0.07244718470176548
l2 norm of weights: 2.2152484493741182

Iteration Number: 7433
Loss: 11.192360681764402
l2 norm of gradients: 0.07243411632213541
l2 norm of weights: 2.215268128478053

Iteration Number: 7434
Loss: 11.191975901076647
l2 norm of gradients: 0.07242105320162014
l2 norm of weights: 2.2152878263018545

Iteration Number: 7435
Loss: 11.191591339154199
l2 norm of gradients: 0.07240799533723141
l2 norm of weights: 2.2153075428301

Iteration Number: 7436
Loss: 11.191206995851902
l2 norm of gradients: 0.07239494272598342
l2 norm of weights: 2.2153272780473796

Iteration Number: 7437
Loss: 11.190822871024706
l2 norm of gradients: 0.07238189536489248
l2 norm of weights: 2.215347031938294

```

-----
Iteration Number: 7438
Loss: 11.190438964527672
l2 norm of gradients: 0.07236885325097747
l2 norm of weights: 2.2153668044874557
-----
Iteration Number: 7439
Loss: 11.190055276215958
l2 norm of gradients: 0.07235581638125947
l2 norm of weights: 2.2153865956794876
-----
Iteration Number: 7440
Loss: 11.18967180594483
l2 norm of gradients: 0.0723427847527619
l2 norm of weights: 2.2154064054990243
-----
Iteration Number: 7441
Loss: 11.189288553569655
l2 norm of gradients: 0.07232975836251056
l2 norm of weights: 2.2154262339307116
-----
Iteration Number: 7442
Loss: 11.18890551894591
l2 norm of gradients: 0.07231673720753351
l2 norm of weights: 2.215446080959206
-----
Iteration Number: 7443
Loss: 11.188522701929159
l2 norm of gradients: 0.07230372128486115
l2 norm of weights: 2.2154659465691764
-----
Iteration Number: 7444
Loss: 11.188140102375083
l2 norm of gradients: 0.07229071059152625
l2 norm of weights: 2.215485830745301
-----
Iteration Number: 7445
Loss: 11.187757720139466
l2 norm of gradients: 0.07227770512456384
l2 norm of weights: 2.215505733472271
-----
Iteration Number: 7446
Loss: 11.187375555078193
l2 norm of gradients: 0.0722647048810113
l2 norm of weights: 2.2155256547347877
-----
Iteration Number: 7447
Loss: 11.186993607047249

```

```

12 norm of gradients: 0.07225170985790827
12 norm of weights: 2.2155455945175637
-----
Iteration Number: 7448
Loss: 11.186611875902727
12 norm of gradients: 0.07223872005229673
12 norm of weights: 2.215565552805323
-----
Iteration Number: 7449
Loss: 11.186230361500815
12 norm of gradients: 0.072225735461221
12 norm of weights: 2.2155855295828006
-----
Iteration Number: 7450
Loss: 11.185849063697816
12 norm of gradients: 0.07221275608172766
12 norm of weights: 2.2156055248347424
-----
Iteration Number: 7451
Loss: 11.185467982350126
12 norm of gradients: 0.07219978191086565
12 norm of weights: 2.215625538545906
-----
Iteration Number: 7452
Loss: 11.185087117314245
12 norm of gradients: 0.07218681294568607
12 norm of weights: 2.2156455707010596
-----
Iteration Number: 7453
Loss: 11.184706468446782
12 norm of gradients: 0.0721738491832425
12 norm of weights: 2.2156656212849826
-----
Iteration Number: 7454
Loss: 11.184326035604439
12 norm of gradients: 0.07216089062059071
12 norm of weights: 2.215685690282466
-----
Iteration Number: 7455
Loss: 11.183945818644027
12 norm of gradients: 0.07214793725478877
12 norm of weights: 2.215705777678311
-----
Iteration Number: 7456
Loss: 11.183565817422457
12 norm of gradients: 0.07213498908289705
12 norm of weights: 2.21572588345733
-----

```


Iteration Number: 7457
Loss: 11.183186031796746
l2 norm of gradients: 0.0721220461019782
l2 norm of weights: 2.215746007604348

Iteration Number: 7458
Loss: 11.182806461624004
l2 norm of gradients: 0.07210910830909718
l2 norm of weights: 2.2157661501041988

Iteration Number: 7459
Loss: 11.182427106761454
l2 norm of gradients: 0.07209617570132124
l2 norm of weights: 2.2157863109417293

Iteration Number: 7460
Loss: 11.182047967066415
l2 norm of gradients: 0.07208324827571987
l2 norm of weights: 2.2158064901017958

Iteration Number: 7461
Loss: 11.18166904239631
l2 norm of gradients: 0.07207032602936482
l2 norm of weights: 2.2158266875692663

Iteration Number: 7462
Loss: 11.181290332608663
l2 norm of gradients: 0.0720574089593302
l2 norm of weights: 2.2158469033290205

Iteration Number: 7463
Loss: 11.1809118375611
l2 norm of gradients: 0.07204449706269231
l2 norm of weights: 2.215867137365948

Iteration Number: 7464
Loss: 11.180533557111346
l2 norm of gradients: 0.0720315903365298
l2 norm of weights: 2.215887389664951

Iteration Number: 7465
Loss: 11.18015549111723
l2 norm of gradients: 0.07201868877792346
l2 norm of weights: 2.21590766021094

Iteration Number: 7466
Loss: 11.179777639436686
l2 norm of gradients: 0.07200579238395657

```

12 norm of weights: 2.2159279489888397
-----
Iteration Number: 7467
Loss: 11.179400001927743
12 norm of gradients: 0.07199290115171442
12 norm of weights: 2.2159482559835837
-----
Iteration Number: 7468
Loss: 11.17902257844854
12 norm of gradients: 0.07198001507828468
12 norm of weights: 2.215968581180117
-----
Iteration Number: 7469
Loss: 11.178645368857302
12 norm of gradients: 0.0719671341607573
12 norm of weights: 2.2159889245633964
-----
Iteration Number: 7470
Loss: 11.178268373012372
12 norm of gradients: 0.07195425839622445
12 norm of weights: 2.2160092861183887
-----
Iteration Number: 7471
Loss: 11.177891590772186
12 norm of gradients: 0.07194138778178061
12 norm of weights: 2.216029665830072
-----
Iteration Number: 7472
Loss: 11.177515021995282
12 norm of gradients: 0.0719285223145224
12 norm of weights: 2.2160500636834355
-----
Iteration Number: 7473
Loss: 11.177138666540296
12 norm of gradients: 0.07191566199154875
12 norm of weights: 2.2160704796634794
-----
Iteration Number: 7474
Loss: 11.176762524265975
12 norm of gradients: 0.07190280680996095
12 norm of weights: 2.216090913755215
-----
Iteration Number: 7475
Loss: 11.176386595031149
12 norm of gradients: 0.0718899567668623
12 norm of weights: 2.216111365943663
-----
Iteration Number: 7476

```

Loss: 11.176010878694772
l2 norm of gradients: 0.07187711185935852
l2 norm of weights: 2.216131836213858

Iteration Number: 7477
Loss: 11.17563537511588
l2 norm of gradients: 0.07186427208455756
l2 norm of weights: 2.2161523245508428

Iteration Number: 7478
Loss: 11.175260084153614
l2 norm of gradients: 0.07185143743956948
l2 norm of weights: 2.216172830939673

Iteration Number: 7479
Loss: 11.174885005667218
l2 norm of gradients: 0.0718386079215067
l2 norm of weights: 2.216193355365413

Iteration Number: 7480
Loss: 11.174510139516036
l2 norm of gradients: 0.07182578352748385
l2 norm of weights: 2.2162138978131405

Iteration Number: 7481
Loss: 11.174135485559514
l2 norm of gradients: 0.07181296425461779
l2 norm of weights: 2.216234458267943

Iteration Number: 7482
Loss: 11.173761043657196
l2 norm of gradients: 0.07180015010002751
l2 norm of weights: 2.2162550367149185

Iteration Number: 7483
Loss: 11.173386813668722
l2 norm of gradients: 0.07178734106083438
l2 norm of weights: 2.216275633139176

Iteration Number: 7484
Loss: 11.173012795453843
l2 norm of gradients: 0.07177453713416193
l2 norm of weights: 2.2162962475258365

Iteration Number: 7485
Loss: 11.172638988872398
l2 norm of gradients: 0.07176173831713585
l2 norm of weights: 2.2163168798600306

```

-----
Iteration Number: 7486
Loss: 11.172265393784334
l2 norm of gradients: 0.07174894460688412
l2 norm of weights: 2.2163375301269
-----
Iteration Number: 7487
Loss: 11.171892010049696
l2 norm of gradients: 0.07173615600053687
l2 norm of weights: 2.216358198311598
-----
Iteration Number: 7488
Loss: 11.171518837528627
l2 norm of gradients: 0.07172337249522663
l2 norm of weights: 2.2163788843992878
-----
Iteration Number: 7489
Loss: 11.171145876081365
l2 norm of gradients: 0.07171059408808783
l2 norm of weights: 2.2163995883751437
-----
Iteration Number: 7490
Loss: 11.170773125568262
l2 norm of gradients: 0.07169782077625743
l2 norm of weights: 2.216420310224352
-----
Iteration Number: 7491
Loss: 11.170400585849759
l2 norm of gradients: 0.07168505255687431
l2 norm of weights: 2.2164410499321074
-----
Iteration Number: 7492
Loss: 11.170028256786395
l2 norm of gradients: 0.07167228942707982
l2 norm of weights: 2.216461807483618
-----
Iteration Number: 7493
Loss: 11.16965613823881
l2 norm of gradients: 0.07165953138401729
l2 norm of weights: 2.216482582864101
-----
Iteration Number: 7494
Loss: 11.169284230067754
l2 norm of gradients: 0.0716467784248324
l2 norm of weights: 2.2165033760587853
-----
Iteration Number: 7495
Loss: 11.16891253213406

```

```

12 norm of gradients: 0.07163403054667293
12 norm of weights: 2.21652418705291
-----
Iteration Number: 7496
Loss: 11.168541044298667
12 norm of gradients: 0.07162128774668895
12 norm of weights: 2.2165450158317253
-----
Iteration Number: 7497
Loss: 11.168169766422613
12 norm of gradients: 0.07160855002203263
12 norm of weights: 2.2165658623804925
-----
Iteration Number: 7498
Loss: 11.167798698367038
12 norm of gradients: 0.07159581736985837
12 norm of weights: 2.216586726684483
-----
Iteration Number: 7499
Loss: 11.16742783999318
12 norm of gradients: 0.07158308978732278
12 norm of weights: 2.21660760872898
-----
Iteration Number: 7500
Loss: 11.167057191162371
12 norm of gradients: 0.07157036727158464
12 norm of weights: 2.216628508499275
-----
Iteration Number: 7501
Loss: 11.16668675173604
12 norm of gradients: 0.0715576498198049
12 norm of weights: 2.2166494259806746
-----
Iteration Number: 7502
Loss: 11.16631652157573
12 norm of gradients: 0.07154493742914668
12 norm of weights: 2.2166703611584913
-----
Iteration Number: 7503
Loss: 11.165946500543066
12 norm of gradients: 0.0715322300967753
12 norm of weights: 2.216691314018052
-----
Iteration Number: 7504
Loss: 11.165576688499778
12 norm of gradients: 0.0715195278198583
12 norm of weights: 2.2167122845446925
-----

```

Iteration Number: 7505
Loss: 11.165207085307692
l2 norm of gradients: 0.07150683059556527
l2 norm of weights: 2.2167332727237596

Iteration Number: 7506
Loss: 11.164837690828739
l2 norm of gradients: 0.07149413842106817
l2 norm of weights: 2.2167542785406114

Iteration Number: 7507
Loss: 11.164468504924937
l2 norm of gradients: 0.07148145129354092
l2 norm of weights: 2.216775301980616

Iteration Number: 7508
Loss: 11.164099527458415
l2 norm of gradients: 0.07146876921015975
l2 norm of weights: 2.2167963430291526

Iteration Number: 7509
Loss: 11.16373075829139
l2 norm of gradients: 0.07145609216810296
l2 norm of weights: 2.2168174016716113

Iteration Number: 7510
Loss: 11.163362197286185
l2 norm of gradients: 0.07144342016455107
l2 norm of weights: 2.216838477893393

Iteration Number: 7511
Loss: 11.16299384430521
l2 norm of gradients: 0.07143075319668679
l2 norm of weights: 2.2168595716799078

Iteration Number: 7512
Loss: 11.162625699210988
l2 norm of gradients: 0.0714180912616949
l2 norm of weights: 2.2168806830165786

Iteration Number: 7513
Loss: 11.162257761866124
l2 norm of gradients: 0.07140543435676243
l2 norm of weights: 2.2169018118888375

Iteration Number: 7514
Loss: 11.161890032133336
l2 norm of gradients: 0.07139278247907851

```
l2 norm of weights: 2.2169229582821277
-----
Iteration Number: 7515
Loss: 11.161522509875423
l2 norm of gradients: 0.07138013562583437
l2 norm of weights: 2.2169441221819035
-----
Iteration Number: 7516
Loss: 11.161155194955295
l2 norm of gradients: 0.07136749379422351
l2 norm of weights: 2.216965303573629
-----
Iteration Number: 7517
Loss: 11.160788087235957
l2 norm of gradients: 0.07135485698144148
l2 norm of weights: 2.2169865024427793
-----
Iteration Number: 7518
Loss: 11.160421186580509
l2 norm of gradients: 0.07134222518468604
l2 norm of weights: 2.2170077187748407
-----
Iteration Number: 7519
Loss: 11.160054492852142
l2 norm of gradients: 0.07132959840115703
l2 norm of weights: 2.2170289525553097
-----
Iteration Number: 7520
Loss: 11.159688005914159
l2 norm of gradients: 0.0713169766280565
l2 norm of weights: 2.217050203769693
-----
Iteration Number: 7521
Loss: 11.159321725629951
l2 norm of gradients: 0.07130435986258858
l2 norm of weights: 2.217071472403508
-----
Iteration Number: 7522
Loss: 11.158955651863003
l2 norm of gradients: 0.07129174810195958
l2 norm of weights: 2.217092758442284
-----
Iteration Number: 7523
Loss: 11.158589784476906
l2 norm of gradients: 0.0712791413433779
l2 norm of weights: 2.217114061871559
-----
Iteration Number: 7524
```

Loss: 11.158224123335343
l2 norm of gradients: 0.0712665395840541
l2 norm of weights: 2.217135382676883

Iteration Number: 7525
Loss: 11.15785866830209
l2 norm of gradients: 0.07125394282120086
l2 norm of weights: 2.2171567208438154

Iteration Number: 7526
Loss: 11.157493419241026
l2 norm of gradients: 0.07124135105203296
l2 norm of weights: 2.2171780763579276

Iteration Number: 7527
Loss: 11.157128376016127
l2 norm of gradients: 0.07122876427376737
l2 norm of weights: 2.217199449204801

Iteration Number: 7528
Loss: 11.156763538491457
l2 norm of gradients: 0.07121618248362314
l2 norm of weights: 2.2172208393700266

Iteration Number: 7529
Loss: 11.15639890653119
l2 norm of gradients: 0.07120360567882143
l2 norm of weights: 2.217242246839207

Iteration Number: 7530
Loss: 11.15603447999959
l2 norm of gradients: 0.0711910338565856
l2 norm of weights: 2.2172636715979555

Iteration Number: 7531
Loss: 11.155670258761011
l2 norm of gradients: 0.07117846701414095
l2 norm of weights: 2.217285113631895

Iteration Number: 7532
Loss: 11.15530624267991
l2 norm of gradients: 0.07116590514871506
l2 norm of weights: 2.2173065729266597

Iteration Number: 7533
Loss: 11.154942431620844
l2 norm of gradients: 0.07115334825753752
l2 norm of weights: 2.217328049467895


```
-----  
Iteration Number: 7534  
Loss: 11.154578825448459  
l2 norm of gradients: 0.07114079633784015  
l2 norm of weights: 2.217349543241254  
-----  
Iteration Number: 7535  
Loss: 11.1542154240275  
l2 norm of gradients: 0.07112824938685677  
l2 norm of weights: 2.217371054232404  
-----  
Iteration Number: 7536  
Loss: 11.153852227222805  
l2 norm of gradients: 0.07111570740182327  
l2 norm of weights: 2.21739258242702  
-----  
Iteration Number: 7537  
Loss: 11.153489234899315  
l2 norm of gradients: 0.07110317037997779  
l2 norm of weights: 2.2174141278107893  
-----  
Iteration Number: 7538  
Loss: 11.153126446922059  
l2 norm of gradients: 0.07109063831856044  
l2 norm of weights: 2.217435690369409  
-----  
Iteration Number: 7539  
Loss: 11.15276386315617  
l2 norm of gradients: 0.07107811121481349  
l2 norm of weights: 2.217457270088586  
-----  
Iteration Number: 7540  
Loss: 11.15240148346687  
l2 norm of gradients: 0.07106558906598126  
l2 norm of weights: 2.2174788669540386  
-----  
Iteration Number: 7541  
Loss: 11.152039307719475  
l2 norm of gradients: 0.07105307186931023  
l2 norm of weights: 2.2175004809514953  
-----  
Iteration Number: 7542  
Loss: 11.151677335779407  
l2 norm of gradients: 0.07104055962204889  
l2 norm of weights: 2.2175221120666953  
-----  
Iteration Number: 7543  
Loss: 11.151315567512174
```

```
l2 norm of gradients: 0.07102805232144793
l2 norm of weights: 2.2175437602853876
-----
Iteration Number: 7544
Loss: 11.150954002783381
l2 norm of gradients: 0.07101554996475998
l2 norm of weights: 2.2175654255933326
-----
Iteration Number: 7545
Loss: 11.150592641458731
l2 norm of gradients: 0.07100305254923989
l2 norm of weights: 2.2175871079763
-----
Iteration Number: 7546
Loss: 11.150231483404022
l2 norm of gradients: 0.07099056007214448
l2 norm of weights: 2.217608807420071
-----
Iteration Number: 7547
Loss: 11.149870528485147
l2 norm of gradients: 0.07097807253073278
l2 norm of weights: 2.217630523910437
-----
Iteration Number: 7548
Loss: 11.149509776568088
l2 norm of gradients: 0.07096558992226577
l2 norm of weights: 2.217652257433199
-----
Iteration Number: 7549
Loss: 11.149149227518937
l2 norm of gradients: 0.07095311224400656
l2 norm of weights: 2.2176740079741695
-----
Iteration Number: 7550
Loss: 11.148788881203862
l2 norm of gradients: 0.07094063949322038
l2 norm of weights: 2.217695775519171
-----
Iteration Number: 7551
Loss: 11.148428737489139
l2 norm of gradients: 0.07092817166717445
l2 norm of weights: 2.217717560054036
-----
Iteration Number: 7552
Loss: 11.148068796241136
l2 norm of gradients: 0.07091570876313807
l2 norm of weights: 2.217739361564608
-----
```

Iteration Number: 7553
Loss: 11.147709057326315
l2 norm of gradients: 0.07090325077838269
l2 norm of weights: 2.21776118003674

Iteration Number: 7554
Loss: 11.147349520611234
l2 norm of gradients: 0.07089079771018171
l2 norm of weights: 2.217783015456297

Iteration Number: 7555
Loss: 11.14699018596254
l2 norm of gradients: 0.07087834955581068
l2 norm of weights: 2.217804867809153

Iteration Number: 7556
Loss: 11.146631053246983
l2 norm of gradients: 0.07086590631254716
l2 norm of weights: 2.2178267370811926

Iteration Number: 7557
Loss: 11.146272122331398
l2 norm of gradients: 0.07085346797767085
l2 norm of weights: 2.2178486232583112

Iteration Number: 7558
Loss: 11.14591339308273
l2 norm of gradients: 0.07084103454846331
l2 norm of weights: 2.217870526326414

Iteration Number: 7559
Loss: 11.145554865367995
l2 norm of gradients: 0.07082860602220842
l2 norm of weights: 2.217892446271417

Iteration Number: 7560
Loss: 11.145196539054325
l2 norm of gradients: 0.07081618239619193
l2 norm of weights: 2.2179143830792465

Iteration Number: 7561
Loss: 11.144838414008934
l2 norm of gradients: 0.07080376366770172
l2 norm of weights: 2.2179363367358387

Iteration Number: 7562
Loss: 11.144480490099134
l2 norm of gradients: 0.07079134983402763

```
12 norm of weights: 2.217958307227141
-----
Iteration Number: 7563
Loss: 11.144122767192332
12 norm of gradients: 0.07077894089246159
12 norm of weights: 2.2179802945391094
-----
Iteration Number: 7564
Loss: 11.143765245156027
12 norm of gradients: 0.07076653684029767
12 norm of weights: 2.218002298657712
-----
Iteration Number: 7565
Loss: 11.143407923857815
12 norm of gradients: 0.07075413767483184
12 norm of weights: 2.2180243195689275
-----
Iteration Number: 7566
Loss: 11.143050803165373
12 norm of gradients: 0.07074174339336214
12 norm of weights: 2.2180463572587423
-----
Iteration Number: 7567
Loss: 11.142693882946494
12 norm of gradients: 0.07072935399318875
12 norm of weights: 2.2180684117131557
-----
Iteration Number: 7568
Loss: 11.142337163069044
12 norm of gradients: 0.07071696947161375
12 norm of weights: 2.2180904829181767
-----
Iteration Number: 7569
Loss: 11.141980643400997
12 norm of gradients: 0.07070458982594134
12 norm of weights: 2.2181125708598235
-----
Iteration Number: 7570
Loss: 11.141624323810412
12 norm of gradients: 0.07069221505347766
12 norm of weights: 2.218134675524125
-----
Iteration Number: 7571
Loss: 11.141268204165446
12 norm of gradients: 0.07067984515153101
12 norm of weights: 2.218156796897122
-----
Iteration Number: 7572
```

Loss: 11.140912284334341
l2 norm of gradients: 0.0706674801174116
l2 norm of weights: 2.218178934964863

Iteration Number: 7573
Loss: 11.140556564185445
l2 norm of gradients: 0.07065511994843175
l2 norm of weights: 2.2182010897134083

Iteration Number: 7574
Loss: 11.140201043587192
l2 norm of gradients: 0.0706427646419057
l2 norm of weights: 2.2182232611288284

Iteration Number: 7575
Loss: 11.13984572240811
l2 norm of gradients: 0.07063041419514983
l2 norm of weights: 2.218245449197203

Iteration Number: 7576
Loss: 11.139490600516819
l2 norm of gradients: 0.07061806860548243
l2 norm of weights: 2.2182676539046238

Iteration Number: 7577
Loss: 11.139135677782035
l2 norm of gradients: 0.07060572787022389
l2 norm of weights: 2.218289875237191

Iteration Number: 7578
Loss: 11.138780954072564
l2 norm of gradients: 0.07059339198669654
l2 norm of weights: 2.2183121131810166

Iteration Number: 7579
Loss: 11.138426429257303
l2 norm of gradients: 0.07058106095222476
l2 norm of weights: 2.2183343677222207

Iteration Number: 7580
Loss: 11.138072103205248
l2 norm of gradients: 0.07056873476413497
l2 norm of weights: 2.2183566388469353

Iteration Number: 7581
Loss: 11.137717975785485
l2 norm of gradients: 0.07055641341975552
l2 norm of weights: 2.2183789265413023

```
-----  
Iteration Number: 7582  
Loss: 11.137364046867193  
l2 norm of gradients: 0.07054409691641689  
l2 norm of weights: 2.2184012307914736  
-----  
Iteration Number: 7583  
Loss: 11.137010316319637  
l2 norm of gradients: 0.07053178525145137  
l2 norm of weights: 2.2184235515836113  
-----  
Iteration Number: 7584  
Loss: 11.136656784012182  
l2 norm of gradients: 0.07051947842219343  
l2 norm of weights: 2.218445888903887  
-----  
Iteration Number: 7585  
Loss: 11.136303449814289  
l2 norm of gradients: 0.07050717642597942  
l2 norm of weights: 2.218468242738484  
-----  
Iteration Number: 7586  
Loss: 11.1359503135955  
l2 norm of gradients: 0.07049487926014779  
l2 norm of weights: 2.218490613073595  
-----  
Iteration Number: 7587  
Loss: 11.135597375225458  
l2 norm of gradients: 0.07048258692203895  
l2 norm of weights: 2.218512999895422  
-----  
Iteration Number: 7588  
Loss: 11.13524463457389  
l2 norm of gradients: 0.07047029940899517  
l2 norm of weights: 2.2185354031901787  
-----  
Iteration Number: 7589  
Loss: 11.134892091510629  
l2 norm of gradients: 0.07045801671836094  
l2 norm of weights: 2.2185578229440877  
-----  
Iteration Number: 7590  
Loss: 11.134539745905588  
l2 norm of gradients: 0.07044573884748258  
l2 norm of weights: 2.2185802591433816  
-----  
Iteration Number: 7591  
Loss: 11.134187597628772
```

```

12 norm of gradients: 0.0704334657937084
12 norm of weights: 2.2186027117743046
-----
Iteration Number: 7592
Loss: 11.133835646550287
12 norm of gradients: 0.0704211975543888
12 norm of weights: 2.21862518082311
-----
Iteration Number: 7593
Loss: 11.133483892540319
12 norm of gradients: 0.07040893412687604
12 norm of weights: 2.2186476662760612
-----
Iteration Number: 7594
Loss: 11.133132335469154
12 norm of gradients: 0.07039667550852441
12 norm of weights: 2.2186701681194316
-----
Iteration Number: 7595
Loss: 11.13278097520717
12 norm of gradients: 0.07038442169669024
12 norm of weights: 2.218692686339505
-----
Iteration Number: 7596
Loss: 11.132429811624835
12 norm of gradients: 0.07037217268873171
12 norm of weights: 2.218715220922576
-----
Iteration Number: 7597
Loss: 11.132078844592707
12 norm of gradients: 0.07035992848200906
12 norm of weights: 2.218737771854947
-----
Iteration Number: 7598
Loss: 11.131728073981435
12 norm of gradients: 0.0703476890738845
12 norm of weights: 2.218760339122934
-----
Iteration Number: 7599
Loss: 11.13137749966176
12 norm of gradients: 0.07033545446172218
12 norm of weights: 2.218782922712859
-----
Iteration Number: 7600
Loss: 11.131027121504516
12 norm of gradients: 0.0703232246428882
12 norm of weights: 2.218805522611058
-----

```

Iteration Number: 7601
Loss: 11.130676939380631
l2 norm of gradients: 0.0703109996147507
l2 norm of weights: 2.2188281388038735

Iteration Number: 7602
Loss: 11.130326953161116
l2 norm of gradients: 0.07029877937467968
l2 norm of weights: 2.218850771277661

Iteration Number: 7603
Loss: 11.129977162717077
l2 norm of gradients: 0.0702865639200472
l2 norm of weights: 2.2188734200187845

Iteration Number: 7604
Loss: 11.129627567919714
l2 norm of gradients: 0.07027435324822723
l2 norm of weights: 2.218896085013618

Iteration Number: 7605
Loss: 11.129278168640315
l2 norm of gradients: 0.07026214735659568
l2 norm of weights: 2.2189187662485463

Iteration Number: 7606
Loss: 11.128928964750266
l2 norm of gradients: 0.07024994624253046
l2 norm of weights: 2.218941463709963

Iteration Number: 7607
Loss: 11.128579956121024
l2 norm of gradients: 0.07023774990341142
l2 norm of weights: 2.2189641773842737

Iteration Number: 7608
Loss: 11.128231142624164
l2 norm of gradients: 0.07022555833662035
l2 norm of weights: 2.218986907257892

Iteration Number: 7609
Loss: 11.12788252413133
l2 norm of gradients: 0.07021337153954095
l2 norm of weights: 2.2190096533172428

Iteration Number: 7610
Loss: 11.127534100514268
l2 norm of gradients: 0.07020118950955892


```
12 norm of weights: 2.2190324155487597
-----
Iteration Number: 7611
Loss: 11.127185871644807
12 norm of gradients: 0.07018901224406192
12 norm of weights: 2.219055193938888
-----
Iteration Number: 7612
Loss: 11.126837837394875
12 norm of gradients: 0.07017683974043949
12 norm of weights: 2.219077988474082
-----
Iteration Number: 7613
Loss: 11.126489997636481
12 norm of gradients: 0.07016467199608319
12 norm of weights: 2.2191007991408056
-----
Iteration Number: 7614
Loss: 11.12614235224174
12 norm of gradients: 0.07015250900838643
12 norm of weights: 2.2191236259255334
-----
Iteration Number: 7615
Loss: 11.125794901082836
12 norm of gradients: 0.07014035077474463
12 norm of weights: 2.2191464688147495
-----
Iteration Number: 7616
Loss: 11.125447644032057
12 norm of gradients: 0.07012819729255507
12 norm of weights: 2.219169327794949
-----
Iteration Number: 7617
Loss: 11.12510058096178
12 norm of gradients: 0.07011604855921706
12 norm of weights: 2.219192202852635
-----
Iteration Number: 7618
Loss: 11.124753711744468
12 norm of gradients: 0.07010390457213177
12 norm of weights: 2.2192150939743223
-----
Iteration Number: 7619
Loss: 11.124407036252679
12 norm of gradients: 0.07009176532870227
12 norm of weights: 2.219238001146535
-----
Iteration Number: 7620
```

Loss: 11.124060554359055
l2 norm of gradients: 0.07007963082633369
l2 norm of weights: 2.2192609243558072

Iteration Number: 7621
Loss: 11.123714265936332
l2 norm of gradients: 0.07006750106243291
l2 norm of weights: 2.2192838635886827

Iteration Number: 7622
Loss: 11.123368170857335
l2 norm of gradients: 0.07005537603440887
l2 norm of weights: 2.219306818831715

Iteration Number: 7623
Loss: 11.123022268994978
l2 norm of gradients: 0.07004325573967239
l2 norm of weights: 2.2193297900714684

Iteration Number: 7624
Loss: 11.122676560222267
l2 norm of gradients: 0.07003114017563616
l2 norm of weights: 2.219352777294516

Iteration Number: 7625
Loss: 11.12233104441229
l2 norm of gradients: 0.07001902933971484
l2 norm of weights: 2.2193757804874426

Iteration Number: 7626
Loss: 11.12198572143824
l2 norm of gradients: 0.07000692322932499
l2 norm of weights: 2.2193987996368403

Iteration Number: 7627
Loss: 11.121640591173378
l2 norm of gradients: 0.06999482184188506
l2 norm of weights: 2.2194218347293133

Iteration Number: 7628
Loss: 11.121295653491075
l2 norm of gradients: 0.06998272517481544
l2 norm of weights: 2.219444885751475

Iteration Number: 7629
Loss: 11.120950908264776
l2 norm of gradients: 0.06997063322553847
l2 norm of weights: 2.2194679526899477

```

-----
Iteration Number: 7630
Loss: 11.120606355368029
l2 norm of gradients: 0.06995854599147826
l2 norm of weights: 2.2194910355313646
-----
Iteration Number: 7631
Loss: 11.120261994674458
l2 norm of gradients: 0.06994646347006098
l2 norm of weights: 2.219514134262369
-----
Iteration Number: 7632
Loss: 11.119917826057783
l2 norm of gradients: 0.06993438565871457
l2 norm of weights: 2.2195372488696132
-----
Iteration Number: 7633
Loss: 11.119573849391811
l2 norm of gradients: 0.069922312554869
l2 norm of weights: 2.21956037933976
-----
Iteration Number: 7634
Loss: 11.11923006455044
l2 norm of gradients: 0.06991024415595604
l2 norm of weights: 2.2195835256594814
-----
Iteration Number: 7635
Loss: 11.11888647140766
l2 norm of gradients: 0.06989818045940939
l2 norm of weights: 2.21960668781546
-----
Iteration Number: 7636
Loss: 11.118543069837536
l2 norm of gradients: 0.06988612146266462
l2 norm of weights: 2.219629865794387
-----
Iteration Number: 7637
Loss: 11.118199859714238
l2 norm of gradients: 0.06987406716315923
l2 norm of weights: 2.2196530595829653
-----
Iteration Number: 7638
Loss: 11.117856840912015
l2 norm of gradients: 0.06986201755833257
l2 norm of weights: 2.219676269167906
-----
Iteration Number: 7639
Loss: 11.11751401330521

```

```

12 norm of gradients: 0.06984997264562598
12 norm of weights: 2.2196994945359307
-----
Iteration Number: 7640
Loss: 11.117171376768253
12 norm of gradients: 0.06983793242248254
12 norm of weights: 2.2197227356737703
-----
Iteration Number: 7641
Loss: 11.11682893117566
12 norm of gradients: 0.06982589688634731
12 norm of weights: 2.2197459925681664
-----
Iteration Number: 7642
Loss: 11.116486676402031
12 norm of gradients: 0.06981386603466717
12 norm of weights: 2.219769265205869
-----
Iteration Number: 7643
Loss: 11.11614461232207
12 norm of gradients: 0.06980183986489097
12 norm of weights: 2.21979255357364
-----
Iteration Number: 7644
Loss: 11.115802738810556
12 norm of gradients: 0.06978981837446936
12 norm of weights: 2.219815857658248
-----
Iteration Number: 7645
Loss: 11.115461055742358
12 norm of gradients: 0.0697778015608549
12 norm of weights: 2.2198391774464747
-----
Iteration Number: 7646
Loss: 11.115119562992435
12 norm of gradients: 0.069765789421502
12 norm of weights: 2.219862512925109
-----
Iteration Number: 7647
Loss: 11.11477826043584
12 norm of gradients: 0.06975378195386699
12 norm of weights: 2.2198858640809513
-----
Iteration Number: 7648
Loss: 11.114437147947703
12 norm of gradients: 0.06974177915540798
12 norm of weights: 2.219909230900811
-----

```

Iteration Number: 7649
Loss: 11.114096225403243
l2 norm of gradients: 0.06972978102358506
l2 norm of weights: 2.219932613371506

Iteration Number: 7650
Loss: 11.113755492677777
l2 norm of gradients: 0.06971778755586014
l2 norm of weights: 2.219956011479866

Iteration Number: 7651
Loss: 11.113414949646703
l2 norm of gradients: 0.06970579874969696
l2 norm of weights: 2.21997942521273

Iteration Number: 7652
Loss: 11.113074596185506
l2 norm of gradients: 0.06969381460256117
l2 norm of weights: 2.2200028545569457

Iteration Number: 7653
Loss: 11.112734432169757
l2 norm of gradients: 0.06968183511192028
l2 norm of weights: 2.220026299499371

Iteration Number: 7654
Loss: 11.112394457475126
l2 norm of gradients: 0.06966986027524359
l2 norm of weights: 2.220049760026874

Iteration Number: 7655
Loss: 11.112054671977356
l2 norm of gradients: 0.06965789009000241
l2 norm of weights: 2.2200732361263316

Iteration Number: 7656
Loss: 11.111715075552281
l2 norm of gradients: 0.06964592455366969
l2 norm of weights: 2.2200967277846315

Iteration Number: 7657
Loss: 11.111375668075832
l2 norm of gradients: 0.06963396366372042
l2 norm of weights: 2.2201202349886704

Iteration Number: 7658
Loss: 11.111036449424017
l2 norm of gradients: 0.06962200741763136

```

12 norm of weights: 2.220143757725354
-----
Iteration Number: 7659
Loss: 11.110697419472933
12 norm of gradients: 0.06961005581288111
12 norm of weights: 2.2201672959815997
-----
Iteration Number: 7660
Loss: 11.11035857809877
12 norm of gradients: 0.06959810884695017
12 norm of weights: 2.220190849744332
-----
Iteration Number: 7661
Loss: 11.110019925177795
12 norm of gradients: 0.06958616651732083
12 norm of weights: 2.220214419000487
-----
Iteration Number: 7662
Loss: 11.109681460586375
12 norm of gradients: 0.06957422882147726
12 norm of weights: 2.2202380037370104
-----
Iteration Number: 7663
Loss: 11.10934318420095
12 norm of gradients: 0.06956229575690544
12 norm of weights: 2.220261603940856
-----
Iteration Number: 7664
Loss: 11.109005095898059
12 norm of gradients: 0.0695503673210932
12 norm of weights: 2.220285219598988
-----
Iteration Number: 7665
Loss: 11.108667195554318
12 norm of gradients: 0.06953844351153027
12 norm of weights: 2.2203088506983812
-----
Iteration Number: 7666
Loss: 11.108329483046436
12 norm of gradients: 0.06952652432570805
12 norm of weights: 2.2203324972260194
-----
Iteration Number: 7667
Loss: 11.107991958251214
12 norm of gradients: 0.06951460976111999
12 norm of weights: 2.2203561591688947
-----
Iteration Number: 7668

```

Loss: 11.107654621045523
l2 norm of gradients: 0.06950269981526125
l2 norm of weights: 2.2203798365140113

Iteration Number: 7669
Loss: 11.107317471306338
l2 norm of gradients: 0.06949079448562884
l2 norm of weights: 2.220403529248381

Iteration Number: 7670
Loss: 11.10698050891071
l2 norm of gradients: 0.06947889376972155
l2 norm of weights: 2.2204272373590257

Iteration Number: 7671
Loss: 11.106643733735774
l2 norm of gradients: 0.06946699766504008
l2 norm of weights: 2.2204509608329777

Iteration Number: 7672
Loss: 11.106307145658768
l2 norm of gradients: 0.0694551061690869
l2 norm of weights: 2.2204746996572773

Iteration Number: 7673
Loss: 11.105970744556998
l2 norm of gradients: 0.06944321927936634
l2 norm of weights: 2.2204984538189767

Iteration Number: 7674
Loss: 11.10563453030786
l2 norm of gradients: 0.06943133699338452
l2 norm of weights: 2.220522223305135

Iteration Number: 7675
Loss: 11.10529850278885
l2 norm of gradients: 0.06941945930864936
l2 norm of weights: 2.2205460081028234

Iteration Number: 7676
Loss: 11.104962661877533
l2 norm of gradients: 0.06940758622267067
l2 norm of weights: 2.2205698081991203

Iteration Number: 7677
Loss: 11.104627007451567
l2 norm of gradients: 0.06939571773296004
l2 norm of weights: 2.2205936235811152

```
-----  
Iteration Number: 7678  
Loss: 11.104291539388695  
l2 norm of gradients: 0.0693838538370308  
l2 norm of weights: 2.2206174542359074  
-----  
Iteration Number: 7679  
Loss: 11.10395625756675  
l2 norm of gradients: 0.0693719945323982  
l2 norm of weights: 2.220641300150604  
-----  
Iteration Number: 7680  
Loss: 11.103621161863645  
l2 norm of gradients: 0.06936013981657925  
l2 norm of weights: 2.2206651613123234  
-----  
Iteration Number: 7681  
Loss: 11.103286252157384  
l2 norm of gradients: 0.06934828968709278  
l2 norm of weights: 2.2206890377081927  
-----  
Iteration Number: 7682  
Loss: 11.10295152832605  
l2 norm of gradients: 0.06933644414145938  
l2 norm of weights: 2.220712929325349  
-----  
Iteration Number: 7683  
Loss: 11.10261699024782  
l2 norm of gradients: 0.0693246031772015  
l2 norm of weights: 2.220736836150938  
-----  
Iteration Number: 7684  
Loss: 11.102282637800949  
l2 norm of gradients: 0.06931276679184344  
l2 norm of weights: 2.2207607581721156  
-----  
Iteration Number: 7685  
Loss: 11.101948470863784  
l2 norm of gradients: 0.0693009349829111  
l2 norm of weights: 2.2207846953760475  
-----  
Iteration Number: 7686  
Loss: 11.10161448931475  
l2 norm of gradients: 0.06928910774793241  
l2 norm of weights: 2.220808647749908  
-----  
Iteration Number: 7687  
Loss: 11.101280693032365
```


l2 norm of gradients: 0.06927728508443694
l2 norm of weights: 2.220832615280882

Iteration Number: 7688
Loss: 11.100947081895232
l2 norm of gradients: 0.06926546698995617
l2 norm of weights: 2.2208565979561627

Iteration Number: 7689
Loss: 11.100613655782027
l2 norm of gradients: 0.06925365346202327
l2 norm of weights: 2.2208805957629534

Iteration Number: 7690
Loss: 11.100280414571529
l2 norm of gradients: 0.06924184449817324
l2 norm of weights: 2.220904608688467

Iteration Number: 7691
Loss: 11.099947358142591
l2 norm of gradients: 0.06923004009594289
l2 norm of weights: 2.220928636719926

Iteration Number: 7692
Loss: 11.099614486374152
l2 norm of gradients: 0.06921824025287081
l2 norm of weights: 2.220952679844561

Iteration Number: 7693
Loss: 11.099281799145244
l2 norm of gradients: 0.06920644496649735
l2 norm of weights: 2.2209767380496146

Iteration Number: 7694
Loss: 11.09894929633497
l2 norm of gradients: 0.06919465423436463
l2 norm of weights: 2.221000811322336

Iteration Number: 7695
Loss: 11.098616977822529
l2 norm of gradients: 0.06918286805401663
l2 norm of weights: 2.2210248996499855

Iteration Number: 7696
Loss: 11.098284843487201
l2 norm of gradients: 0.06917108642299903
l2 norm of weights: 2.2210490030198327

Iteration Number: 7697
Loss: 11.09795289320835
l2 norm of gradients: 0.06915930933885933
l2 norm of weights: 2.221073121419156

Iteration Number: 7698
Loss: 11.097621126865429
l2 norm of gradients: 0.06914753679914679
l2 norm of weights: 2.221097254835244

Iteration Number: 7699
Loss: 11.097289544337972
l2 norm of gradients: 0.0691357688014124
l2 norm of weights: 2.2211214032553945

Iteration Number: 7700
Loss: 11.096958145505592
l2 norm of gradients: 0.06912400534320906
l2 norm of weights: 2.2211455666669147

Iteration Number: 7701
Loss: 11.096626930248002
l2 norm of gradients: 0.06911224642209124
l2 norm of weights: 2.22116974505712

Iteration Number: 7702
Loss: 11.096295898444978
l2 norm of gradients: 0.06910049203561537
l2 norm of weights: 2.2211939384133372

Iteration Number: 7703
Loss: 11.095965049976405
l2 norm of gradients: 0.06908874218133952
l2 norm of weights: 2.2212181467229017

Iteration Number: 7704
Loss: 11.095634384722231
l2 norm of gradients: 0.06907699685682356
l2 norm of weights: 2.221242369973157

Iteration Number: 7705
Loss: 11.095303902562504
l2 norm of gradients: 0.06906525605962917
l2 norm of weights: 2.2212666081514585

Iteration Number: 7706
Loss: 11.094973603377339
l2 norm of gradients: 0.06905351978731974

```
l2 norm of weights: 2.2212908612451683
-----
Iteration Number: 7707
Loss: 11.094643487046953
l2 norm of gradients: 0.06904178803746039
l2 norm of weights: 2.2213151292416597
-----
Iteration Number: 7708
Loss: 11.094313553451634
l2 norm of gradients: 0.0690300608076181
l2 norm of weights: 2.221339412128315
-----
Iteration Number: 7709
Loss: 11.093983802471765
l2 norm of gradients: 0.06901833809536152
l2 norm of weights: 2.221363709892525
-----
Iteration Number: 7710
Loss: 11.0936542339878
l2 norm of gradients: 0.06900661989826104
l2 norm of weights: 2.2213880225216913
-----
Iteration Number: 7711
Loss: 11.093324847880291
l2 norm of gradients: 0.06899490621388886
l2 norm of weights: 2.2214123500032223
-----
Iteration Number: 7712
Loss: 11.092995644029862
l2 norm of gradients: 0.06898319703981898
l2 norm of weights: 2.2214366923245397
-----
Iteration Number: 7713
Loss: 11.092666622317225
l2 norm of gradients: 0.06897149237362696
l2 norm of weights: 2.22146104947307
-----
Iteration Number: 7714
Loss: 11.092337782623176
l2 norm of gradients: 0.06895979221289032
l2 norm of weights: 2.2214854214362534
-----
Iteration Number: 7715
Loss: 11.092009124828598
l2 norm of gradients: 0.06894809655518812
l2 norm of weights: 2.221509808201536
-----
Iteration Number: 7716
```

Loss: 11.091680648814455
l2 norm of gradients: 0.0689364053981014
l2 norm of weights: 2.2215342097563746

Iteration Number: 7717
Loss: 11.091352354461785
l2 norm of gradients: 0.06892471873921274
l2 norm of weights: 2.221558626088235

Iteration Number: 7718
Loss: 11.091024241651725
l2 norm of gradients: 0.06891303657610652
l2 norm of weights: 2.2215830571845934

Iteration Number: 7719
Loss: 11.090696310265487
l2 norm of gradients: 0.0689013589063689
l2 norm of weights: 2.2216075030329328

Iteration Number: 7720
Loss: 11.090368560184368
l2 norm of gradients: 0.0688896857275878
l2 norm of weights: 2.2216319636207484

Iteration Number: 7721
Loss: 11.090040991289746
l2 norm of gradients: 0.06887801703735269
l2 norm of weights: 2.2216564389355424

Iteration Number: 7722
Loss: 11.089713603463084
l2 norm of gradients: 0.06886635283325497
l2 norm of weights: 2.2216809289648274

Iteration Number: 7723
Loss: 11.08938639658593
l2 norm of gradients: 0.06885469311288774
l2 norm of weights: 2.2217054336961257

Iteration Number: 7724
Loss: 11.089059370539912
l2 norm of gradients: 0.06884303787384573
l2 norm of weights: 2.2217299531169674

Iteration Number: 7725
Loss: 11.08873252520674
l2 norm of gradients: 0.06883138711372547
l2 norm of weights: 2.2217544872148927

```
-----
Iteration Number: 7726
Loss: 11.088405860468214
l2 norm of gradients: 0.06881974083012522
l2 norm of weights: 2.221779035977451
-----
Iteration Number: 7727
Loss: 11.088079376206206
l2 norm of gradients: 0.06880809902064493
l2 norm of weights: 2.221803599392201
-----
Iteration Number: 7728
Loss: 11.087753072302679
l2 norm of gradients: 0.0687964616828863
l2 norm of weights: 2.2218281774467106
-----
Iteration Number: 7729
Loss: 11.087426948639676
l2 norm of gradients: 0.0687848288144527
l2 norm of weights: 2.2218527701285566
-----
Iteration Number: 7730
Loss: 11.087101005099324
l2 norm of gradients: 0.0687732004129493
l2 norm of weights: 2.2218773774253253
-----
Iteration Number: 7731
Loss: 11.086775241563824
l2 norm of gradients: 0.0687615764759829
l2 norm of weights: 2.2219019993246123
-----
Iteration Number: 7732
Loss: 11.086449657915479
l2 norm of gradients: 0.0687499570011621
l2 norm of weights: 2.221926635814022
-----
Iteration Number: 7733
Loss: 11.086124254036653
l2 norm of gradients: 0.0687383419860971
l2 norm of weights: 2.2219512868811684
-----
Iteration Number: 7734
Loss: 11.085799029809806
l2 norm of gradients: 0.06872673142839997
l2 norm of weights: 2.2219759525136746
-----
Iteration Number: 7735
Loss: 11.085473985117474
```

```

12 norm of gradients: 0.0687151253256843
12 norm of weights: 2.2220006326991726
-----
Iteration Number: 7736
Loss: 11.085149119842278
12 norm of gradients: 0.06870352367556551
12 norm of weights: 2.222025327425304
-----
Iteration Number: 7737
Loss: 11.08482443386692
12 norm of gradients: 0.06869192647566076
12 norm of weights: 2.2220500366797196
-----
Iteration Number: 7738
Loss: 11.084499927074187
12 norm of gradients: 0.06868033372358878
12 norm of weights: 2.2220747604500786
-----
Iteration Number: 7739
Loss: 11.084175599346942
12 norm of gradients: 0.06866874541697013
12 norm of weights: 2.22209949872405
-----
Iteration Number: 7740
Loss: 11.083851450568137
12 norm of gradients: 0.06865716155342688
12 norm of weights: 2.222124251489312
-----
Iteration Number: 7741
Loss: 11.083527480620804
12 norm of gradients: 0.06864558213058314
12 norm of weights: 2.222149018733552
-----
Iteration Number: 7742
Loss: 11.083203689388046
12 norm of gradients: 0.06863400714606435
12 norm of weights: 2.2221738004444664
-----
Iteration Number: 7743
Loss: 11.082880076753067
12 norm of gradients: 0.06862243659749788
12 norm of weights: 2.2221985966097604
-----
Iteration Number: 7744
Loss: 11.082556642599144
12 norm of gradients: 0.0686108704825127
12 norm of weights: 2.2222234072171485
-----

```

```

Iteration Number: 7745
Loss: 11.082233386809628
l2 norm of gradients: 0.06859930879873948
l2 norm of weights: 2.222248232254354
-----
Iteration Number: 7746
Loss: 11.081910309267961
l2 norm of gradients: 0.06858775154381061
l2 norm of weights: 2.2222730717091106
-----
Iteration Number: 7747
Loss: 11.081587409857663
l2 norm of gradients: 0.06857619871536014
l2 norm of weights: 2.2222979255691597
-----
Iteration Number: 7748
Loss: 11.081264688462342
l2 norm of gradients: 0.06856465031102377
l2 norm of weights: 2.222322793822253
-----
Iteration Number: 7749
Loss: 11.080942144965679
l2 norm of gradients: 0.068553106328439
l2 norm of weights: 2.2223476764561494
-----
Iteration Number: 7750
Loss: 11.080619779251435
l2 norm of gradients: 0.06854156676524482
l2 norm of weights: 2.2223725734586193
-----
Iteration Number: 7751
Loss: 11.080297591203461
l2 norm of gradients: 0.06853003161908217
l2 norm of weights: 2.2223974848174404
-----
Iteration Number: 7752
Loss: 11.079975580705687
l2 norm of gradients: 0.06851850088759341
l2 norm of weights: 2.2224224105204002
-----
Iteration Number: 7753
Loss: 11.07965374764212
l2 norm of gradients: 0.06850697456842275
l2 norm of weights: 2.2224473505552953
-----
Iteration Number: 7754
Loss: 11.079332091896847
l2 norm of gradients: 0.06849545265921597

```

```

12 norm of weights: 2.222472304909931
-----
Iteration Number: 7755
Loss: 11.079010613354042
12 norm of gradients: 0.06848393515762059
12 norm of weights: 2.2224972735721216
-----
Iteration Number: 7756
Loss: 11.07868931189796
12 norm of gradients: 0.06847242206128579
12 norm of weights: 2.2225222565296914
-----
Iteration Number: 7757
Loss: 11.078368187412932
12 norm of gradients: 0.06846091336786243
12 norm of weights: 2.222547253770473
-----
Iteration Number: 7758
Loss: 11.07804723978337
12 norm of gradients: 0.06844940907500294
12 norm of weights: 2.2225722652823072
-----
Iteration Number: 7759
Loss: 11.077726468893776
12 norm of gradients: 0.06843790918036154
12 norm of weights: 2.222597291053046
-----
Iteration Number: 7760
Loss: 11.07740587462872
12 norm of gradients: 0.06842641368159408
12 norm of weights: 2.222622331070548
-----
Iteration Number: 7761
Loss: 11.077085456872862
12 norm of gradients: 0.06841492257635808
12 norm of weights: 2.2226473853226834
-----
Iteration Number: 7762
Loss: 11.076765215510937
12 norm of gradients: 0.06840343586231265
12 norm of weights: 2.2226724537973284
-----
Iteration Number: 7763
Loss: 11.076445150427762
12 norm of gradients: 0.06839195353711867
12 norm of weights: 2.2226975364823707
-----
Iteration Number: 7764

```


Loss: 11.076125261508238
l2 norm of gradients: 0.06838047559843864
l2 norm of weights: 2.222722633365706

Iteration Number: 7765
Loss: 11.075805548637344
l2 norm of gradients: 0.06836900204393662
l2 norm of weights: 2.2227477444352393

Iteration Number: 7766
Loss: 11.075486011700137
l2 norm of gradients: 0.06835753287127852
l2 norm of weights: 2.222772869678884

Iteration Number: 7767
Loss: 11.075166650581762
l2 norm of gradients: 0.06834606807813169
l2 norm of weights: 2.222798009084563

Iteration Number: 7768
Loss: 11.074847465167428
l2 norm of gradients: 0.06833460766216531
l2 norm of weights: 2.222823162640208

Iteration Number: 7769
Loss: 11.074528455342447
l2 norm of gradients: 0.06832315162105008
l2 norm of weights: 2.2228483303337603

Iteration Number: 7770
Loss: 11.074209620992194
l2 norm of gradients: 0.06831169995245848
l2 norm of weights: 2.2228735121531686

Iteration Number: 7771
Loss: 11.073890962002132
l2 norm of gradients: 0.06830025265406447
l2 norm of weights: 2.2228987080863924

Iteration Number: 7772
Loss: 11.073572478257798
l2 norm of gradients: 0.06828880972354377
l2 norm of weights: 2.222923918121399

Iteration Number: 7773
Loss: 11.073254169644814
l2 norm of gradients: 0.06827737115857374
l2 norm of weights: 2.222949142246165

```

-----
Iteration Number: 7774
Loss: 11.072936036048882
l2 norm of gradients: 0.06826593695683335
l2 norm of weights: 2.222974380448676
-----
Iteration Number: 7775
Loss: 11.072618077355777
l2 norm of gradients: 0.06825450711600321
l2 norm of weights: 2.222999632716926
-----
Iteration Number: 7776
Loss: 11.072300293451367
l2 norm of gradients: 0.06824308163376556
l2 norm of weights: 2.2230248990389185
-----
Iteration Number: 7777
Loss: 11.071982684221584
l2 norm of gradients: 0.06823166050780434
l2 norm of weights: 2.2230501794026667
-----
Iteration Number: 7778
Loss: 11.071665249552456
l2 norm of gradients: 0.0682202437358051
l2 norm of weights: 2.2230754737961904
-----
Iteration Number: 7779
Loss: 11.071347989330077
l2 norm of gradients: 0.06820883131545491
l2 norm of weights: 2.2231007822075206
-----
Iteration Number: 7780
Loss: 11.071030903440626
l2 norm of gradients: 0.06819742324444264
l2 norm of weights: 2.223126104624696
-----
Iteration Number: 7781
Loss: 11.070713991770358
l2 norm of gradients: 0.06818601952045868
l2 norm of weights: 2.223151441035765
-----
Iteration Number: 7782
Loss: 11.070397254205618
l2 norm of gradients: 0.06817462014119507
l2 norm of weights: 2.223176791428783
-----
Iteration Number: 7783
Loss: 11.070080690632816

```

```

12 norm of gradients: 0.06816322510434558
12 norm of weights: 2.2232021557918173
-----
Iteration Number: 7784
Loss: 11.069764300938454
12 norm of gradients: 0.06815183440760542
12 norm of weights: 2.2232275341129415
-----
Iteration Number: 7785
Loss: 11.069448085009105
12 norm of gradients: 0.0681404480486715
12 norm of weights: 2.2232529263802396
-----
Iteration Number: 7786
Loss: 11.069132042731425
12 norm of gradients: 0.06812906602524246
12 norm of weights: 2.2232783325818035
-----
Iteration Number: 7787
Loss: 11.068816173992149
12 norm of gradients: 0.06811768833501841
12 norm of weights: 2.2233037527057347
-----
Iteration Number: 7788
Loss: 11.068500478678086
12 norm of gradients: 0.06810631497570116
12 norm of weights: 2.223329186740143
-----
Iteration Number: 7789
Loss: 11.068184956676129
12 norm of gradients: 0.06809494594499406
12 norm of weights: 2.2233546346731474
-----
Iteration Number: 7790
Loss: 11.06786960787325
12 norm of gradients: 0.06808358124060222
12 norm of weights: 2.2233800964928756
-----
Iteration Number: 7791
Loss: 11.067554432156502
12 norm of gradients: 0.06807222086023218
12 norm of weights: 2.2234055721874637
-----
Iteration Number: 7792
Loss: 11.06723942941301
12 norm of gradients: 0.06806086480159224
12 norm of weights: 2.2234310617450577
-----

```

Iteration Number: 7793
Loss: 11.066924599529985
l2 norm of gradients: 0.06804951306239228
l2 norm of weights: 2.2234565651538114

Iteration Number: 7794
Loss: 11.066609942394711
l2 norm of gradients: 0.06803816564034369
l2 norm of weights: 2.223482082401888

Iteration Number: 7795
Loss: 11.066295457894551
l2 norm of gradients: 0.06802682253315957
l2 norm of weights: 2.2235076134774596

Iteration Number: 7796
Loss: 11.065981145916952
l2 norm of gradients: 0.06801548373855461
l2 norm of weights: 2.2235331583687064

Iteration Number: 7797
Loss: 11.065667006349436
l2 norm of gradients: 0.06800414925424508
l2 norm of weights: 2.2235587170638182

Iteration Number: 7798
Loss: 11.065353039079602
l2 norm of gradients: 0.06799281907794887
l2 norm of weights: 2.2235842895509927

Iteration Number: 7799
Loss: 11.065039243995132
l2 norm of gradients: 0.0679814932073854
l2 norm of weights: 2.223609875818437

Iteration Number: 7800
Loss: 11.06472562098378
l2 norm of gradients: 0.06797017164027583
l2 norm of weights: 2.223635475854368

Iteration Number: 7801
Loss: 11.064412169933384
l2 norm of gradients: 0.06795885437434279
l2 norm of weights: 2.223661089647009

Iteration Number: 7802
Loss: 11.064098890731858
l2 norm of gradients: 0.06794754140731052

```

12 norm of weights: 2.2236867171845938
-----
Iteration Number: 7803
Loss: 11.06378578326719
12 norm of gradients: 0.06793623273690498
12 norm of weights: 2.2237123584553644
-----
Iteration Number: 7804
Loss: 11.063472847427459
12 norm of gradients: 0.06792492836085355
12 norm of weights: 2.223738013447572
-----
Iteration Number: 7805
Loss: 11.063160083100803
12 norm of gradients: 0.06791362827688528
12 norm of weights: 2.2237636821494755
-----
Iteration Number: 7806
Loss: 11.062847490175459
12 norm of gradients: 0.06790233248273085
12 norm of weights: 2.223789364549344
-----
Iteration Number: 7807
Loss: 11.062535068539727
12 norm of gradients: 0.06789104097612246
12 norm of weights: 2.223815060635455
-----
Iteration Number: 7808
Loss: 11.062222818081985
12 norm of gradients: 0.0678797537547939
12 norm of weights: 2.223840770396093
-----
Iteration Number: 7809
Loss: 11.061910738690697
12 norm of gradients: 0.06786847081648056
12 norm of weights: 2.223866493819554
-----
Iteration Number: 7810
Loss: 11.061598830254406
12 norm of gradients: 0.06785719215891944
12 norm of weights: 2.2238922308941396
-----
Iteration Number: 7811
Loss: 11.061287092661717
12 norm of gradients: 0.06784591777984912
12 norm of weights: 2.2239179816081633
-----
Iteration Number: 7812

```

Loss: 11.060975525801332
l2 norm of gradients: 0.06783464767700967
l2 norm of weights: 2.223943745949945

Iteration Number: 7813
Loss: 11.060664129562019
l2 norm of gradients: 0.06782338184814284
l2 norm of weights: 2.223969523907815

Iteration Number: 7814
Loss: 11.060352903832626
l2 norm of gradients: 0.0678121202909919
l2 norm of weights: 2.223995315470111

Iteration Number: 7815
Loss: 11.06004184850208
l2 norm of gradients: 0.06780086300330171
l2 norm of weights: 2.224021120625179

Iteration Number: 7816
Loss: 11.059730963459385
l2 norm of gradients: 0.06778960998281876
l2 norm of weights: 2.2240469393613753

Iteration Number: 7817
Loss: 11.059420248593623
l2 norm of gradients: 0.06777836122729103
l2 norm of weights: 2.224072771667064

Iteration Number: 7818
Loss: 11.05910970379395
l2 norm of gradients: 0.06776711673446804
l2 norm of weights: 2.224098617530618

Iteration Number: 7819
Loss: 11.058799328949599
l2 norm of gradients: 0.06775587650210098
l2 norm of weights: 2.224124476940419

Iteration Number: 7820
Loss: 11.058489123949888
l2 norm of gradients: 0.06774464052794255
l2 norm of weights: 2.224150349884856

Iteration Number: 7821
Loss: 11.058179088684202
l2 norm of gradients: 0.06773340880974701
l2 norm of weights: 2.2241762363523296

```
-----
Iteration Number: 7822
Loss: 11.057869223042013
l2 norm of gradients: 0.06772218134527025
l2 norm of weights: 2.224202136331246
-----
Iteration Number: 7823
Loss: 11.057559526912865
l2 norm of gradients: 0.06771095813226963
l2 norm of weights: 2.224228049810022
-----
Iteration Number: 7824
Loss: 11.057250000186372
l2 norm of gradients: 0.06769973916850415
l2 norm of weights: 2.2242539767770815
-----
Iteration Number: 7825
Loss: 11.056940642752233
l2 norm of gradients: 0.06768852445173423
l2 norm of weights: 2.2242799172208585
-----
Iteration Number: 7826
Loss: 11.056631454500232
l2 norm of gradients: 0.06767731397972203
l2 norm of weights: 2.2243058711297956
-----
Iteration Number: 7827
Loss: 11.05632243532021
l2 norm of gradients: 0.06766610775023117
l2 norm of weights: 2.2243318384923425
-----
Iteration Number: 7828
Loss: 11.056013585102102
l2 norm of gradients: 0.06765490576102681
l2 norm of weights: 2.2243578192969586
-----
Iteration Number: 7829
Loss: 11.055704903735911
l2 norm of gradients: 0.0676437080098757
l2 norm of weights: 2.2243838135321123
-----
Iteration Number: 7830
Loss: 11.055396391111717
l2 norm of gradients: 0.06763251449454609
l2 norm of weights: 2.2244098211862795
-----
Iteration Number: 7831
Loss: 11.055088047119677
```

```

12 norm of gradients: 0.06762132521280785
12 norm of weights: 2.2244358422479458
-----
Iteration Number: 7832
Loss: 11.054779871650029
12 norm of gradients: 0.06761014016243241
12 norm of weights: 2.224461876705604
-----
Iteration Number: 7833
Loss: 11.054471864593086
12 norm of gradients: 0.06759895934119263
12 norm of weights: 2.2244879245477573
-----
Iteration Number: 7834
Loss: 11.054164025839228
12 norm of gradients: 0.06758778274686297
12 norm of weights: 2.2245139857629157
-----
Iteration Number: 7835
Loss: 11.053856355278924
12 norm of gradients: 0.0675766103772195
12 norm of weights: 2.224540060339599
-----
Iteration Number: 7836
Loss: 11.053548852802713
12 norm of gradients: 0.06756544223003973
12 norm of weights: 2.2245661482663355
-----
Iteration Number: 7837
Loss: 11.053241518301215
12 norm of gradients: 0.06755427830310276
12 norm of weights: 2.224592249531661
-----
Iteration Number: 7838
Loss: 11.052934351665117
12 norm of gradients: 0.06754311859418928
12 norm of weights: 2.224618364124121
-----
Iteration Number: 7839
Loss: 11.052627352785187
12 norm of gradients: 0.0675319631010814
12 norm of weights: 2.224644492032269
-----
Iteration Number: 7840
Loss: 11.052320521552275
12 norm of gradients: 0.06752081182156286
12 norm of weights: 2.2246706332446675
-----

```


Iteration Number: 7841
Loss: 11.0520138578573
l2 norm of gradients: 0.06750966475341884
l2 norm of weights: 2.2246967877498864

Iteration Number: 7842
Loss: 11.051707361591257
l2 norm of gradients: 0.06749852189443618
l2 norm of weights: 2.224722955536506

Iteration Number: 7843
Loss: 11.051401032645218
l2 norm of gradients: 0.06748738324240316
l2 norm of weights: 2.2247491365931125

Iteration Number: 7844
Loss: 11.051094870910337
l2 norm of gradients: 0.06747624879510958
l2 norm of weights: 2.2247753309083036

Iteration Number: 7845
Loss: 11.05078887627783
l2 norm of gradients: 0.06746511855034681
l2 norm of weights: 2.2248015384706834

Iteration Number: 7846
Loss: 11.050483048639002
l2 norm of gradients: 0.06745399250590775
l2 norm of weights: 2.2248277592688654

Iteration Number: 7847
Loss: 11.050177387885228
l2 norm of gradients: 0.06744287065958679
l2 norm of weights: 2.2248539932914717

Iteration Number: 7848
Loss: 11.049871893907955
l2 norm of gradients: 0.06743175300917988
l2 norm of weights: 2.224880240527132

Iteration Number: 7849
Loss: 11.049566566598717
l2 norm of gradients: 0.0674206395524844
l2 norm of weights: 2.2249065009644853

Iteration Number: 7850
Loss: 11.04926140584911
l2 norm of gradients: 0.06740953028729944

```

12 norm of weights: 2.2249327745921788
-----
Iteration Number: 7851
Loss: 11.048956411550812
12 norm of gradients: 0.06739842521142535
12 norm of weights: 2.2249590613988683
-----
Iteration Number: 7852
Loss: 11.04865158359558
12 norm of gradients: 0.0673873243226642
12 norm of weights: 2.2249853613732182
-----
Iteration Number: 7853
Loss: 11.04834692187524
12 norm of gradients: 0.06737622761881953
12 norm of weights: 2.225011674503901
-----
Iteration Number: 7854
Loss: 11.048042426281697
12 norm of gradients: 0.0673651350976963
12 norm of weights: 2.225038000779598
-----
Iteration Number: 7855
Loss: 11.047738096706926
12 norm of gradients: 0.06735404675710113
12 norm of weights: 2.225064340188998
-----
Iteration Number: 7856
Loss: 11.047433933042983
12 norm of gradients: 0.06734296259484206
12 norm of weights: 2.2250906927208005
-----
Iteration Number: 7857
Loss: 11.047129935182
12 norm of gradients: 0.06733188260872858
12 norm of weights: 2.2251170583637103
-----
Iteration Number: 7858
Loss: 11.046826103016175
12 norm of gradients: 0.06732080679657182
12 norm of weights: 2.2251434371064436
-----
Iteration Number: 7859
Loss: 11.046522436437792
12 norm of gradients: 0.06730973515618437
12 norm of weights: 2.225169828937723
-----
Iteration Number: 7860

```

Loss: 11.046218935339201
l2 norm of gradients: 0.06729866768538026
l2 norm of weights: 2.225196233846281

Iteration Number: 7861
Loss: 11.045915599612837
l2 norm of gradients: 0.06728760438197509
l2 norm of weights: 2.2252226518208573

Iteration Number: 7862
Loss: 11.045612429151198
l2 norm of gradients: 0.06727654524378593
l2 norm of weights: 2.2252490828502003

Iteration Number: 7863
Loss: 11.045309423846863
l2 norm of gradients: 0.06726549026863139
l2 norm of weights: 2.2252755269230673

Iteration Number: 7864
Loss: 11.04500658359249
l2 norm of gradients: 0.06725443945433152
l2 norm of weights: 2.2253019840282238

Iteration Number: 7865
Loss: 11.044703908280804
l2 norm of gradients: 0.06724339279870795
l2 norm of weights: 2.2253284541544436

Iteration Number: 7866
Loss: 11.0444013978046
l2 norm of gradients: 0.06723235029958365
l2 norm of weights: 2.2253549372905086

Iteration Number: 7867
Loss: 11.044099052056767
l2 norm of gradients: 0.06722131195478327
l2 norm of weights: 2.2253814334252096

Iteration Number: 7868
Loss: 11.04379687093025
l2 norm of gradients: 0.06721027776213287
l2 norm of weights: 2.2254079425473456

Iteration Number: 7869
Loss: 11.043494854318073
l2 norm of gradients: 0.06719924771945994
l2 norm of weights: 2.225434464645724

```
-----  
Iteration Number: 7870  
Loss: 11.043193002113343  
l2 norm of gradients: 0.06718822182459357  
l2 norm of weights: 2.22546099970916  
-----  
Iteration Number: 7871  
Loss: 11.042891314209228  
l2 norm of gradients: 0.06717720007536425  
l2 norm of weights: 2.225487547726478  
-----  
Iteration Number: 7872  
Loss: 11.042589790498981  
l2 norm of gradients: 0.067166182469604  
l2 norm of weights: 2.2255141086865113  
-----  
Iteration Number: 7873  
Loss: 11.042288430875924  
l2 norm of gradients: 0.06715516900514633  
l2 norm of weights: 2.225540682578099  
-----  
Iteration Number: 7874  
Loss: 11.041987235233451  
l2 norm of gradients: 0.0671441596798262  
l2 norm of weights: 2.225567269390091  
-----  
Iteration Number: 7875  
Loss: 11.041686203465042  
l2 norm of gradients: 0.06713315449148011  
l2 norm of weights: 2.225593869111345  
-----  
Iteration Number: 7876  
Loss: 11.041385335464229  
l2 norm of gradients: 0.06712215343794588  
l2 norm of weights: 2.2256204817307266  
-----  
Iteration Number: 7877  
Loss: 11.04108463112464  
l2 norm of gradients: 0.06711115651706309  
l2 norm of weights: 2.2256471072371102  
-----  
Iteration Number: 7878  
Loss: 11.040784090339967  
l2 norm of gradients: 0.06710016372667252  
l2 norm of weights: 2.2256737456193774  
-----  
Iteration Number: 7879  
Loss: 11.040483713003978
```

```

12 norm of gradients: 0.0670891750646166
12 norm of weights: 2.2257003968664195
-----
Iteration Number: 7880
Loss: 11.040183499010512
12 norm of gradients: 0.06707819052873915
12 norm of weights: 2.225727060967136
-----
Iteration Number: 7881
Loss: 11.039883448253484
12 norm of gradients: 0.0670672101168855
12 norm of weights: 2.225753737910433
-----
Iteration Number: 7882
Loss: 11.039583560626879
12 norm of gradients: 0.06705623382690241
12 norm of weights: 2.225780427685227
-----
Iteration Number: 7883
Loss: 11.039283836024769
12 norm of gradients: 0.0670452616566382
12 norm of weights: 2.225807130280442
-----
Iteration Number: 7884
Loss: 11.038984274341281
12 norm of gradients: 0.06703429360394256
12 norm of weights: 2.22583384568501
-----
Iteration Number: 7885
Loss: 11.03868487547063
12 norm of gradients: 0.06702332966666667
12 norm of weights: 2.225860573887872
-----
Iteration Number: 7886
Loss: 11.03838563930709
12 norm of gradients: 0.06701236984266322
12 norm of weights: 2.2258873148779754
-----
Iteration Number: 7887
Loss: 11.038086565745026
12 norm of gradients: 0.06700141412978632
12 norm of weights: 2.2259140686442787
-----
Iteration Number: 7888
Loss: 11.037787654678862
12 norm of gradients: 0.06699046252589155
12 norm of weights: 2.225940835175747
-----

```

Iteration Number: 7889
Loss: 11.037488906003102
l2 norm of gradients: 0.06697951502883594
l2 norm of weights: 2.2259676144613527

Iteration Number: 7890
Loss: 11.037190319612325
l2 norm of gradients: 0.06696857163647803
l2 norm of weights: 2.225994406490079

Iteration Number: 7891
Loss: 11.036891895401178
l2 norm of gradients: 0.06695763234667777
l2 norm of weights: 2.2260212112509157

Iteration Number: 7892
Loss: 11.036593633264387
l2 norm of gradients: 0.0669466971572966
l2 norm of weights: 2.2260480287328606

Iteration Number: 7893
Loss: 11.036295533096741
l2 norm of gradients: 0.06693576606619737
l2 norm of weights: 2.2260748589249206

Iteration Number: 7894
Loss: 11.03599759479311
l2 norm of gradients: 0.06692483907124443
l2 norm of weights: 2.22610170181611

Iteration Number: 7895
Loss: 11.035699818248444
l2 norm of gradients: 0.06691391617030355
l2 norm of weights: 2.2261285573954526

Iteration Number: 7896
Loss: 11.035402203357748
l2 norm of gradients: 0.06690299736124196
l2 norm of weights: 2.226155425651979

Iteration Number: 7897
Loss: 11.035104750016114
l2 norm of gradients: 0.06689208264192835
l2 norm of weights: 2.2261823065747284

Iteration Number: 7898
Loss: 11.034807458118703
l2 norm of gradients: 0.06688117201023284

```

12 norm of weights: 2.226209200152749
-----
Iteration Number: 7899
Loss: 11.034510327560746
12 norm of gradients: 0.06687026546402705
12 norm of weights: 2.226236106375097
-----
Iteration Number: 7900
Loss: 11.03421335823755
12 norm of gradients: 0.06685936300118396
12 norm of weights: 2.226263025230835
-----
Iteration Number: 7901
Loss: 11.033916550044493
12 norm of gradients: 0.06684846461957802
12 norm of weights: 2.226289956709036
-----
Iteration Number: 7902
Loss: 11.033619902877028
12 norm of gradients: 0.0668375703170852
12 norm of weights: 2.226316900798781
-----
Iteration Number: 7903
Loss: 11.033323416630676
12 norm of gradients: 0.06682668009158278
12 norm of weights: 2.226343857489158
-----
Iteration Number: 7904
Loss: 11.033027091201033
12 norm of gradients: 0.0668157939409496
12 norm of weights: 2.226370826769263
-----
Iteration Number: 7905
Loss: 11.032730926483772
12 norm of gradients: 0.06680491186306585
12 norm of weights: 2.226397808628202
-----
Iteration Number: 7906
Loss: 11.032434922374632
12 norm of gradients: 0.0667940338558132
12 norm of weights: 2.226424803055088
-----
Iteration Number: 7907
Loss: 11.032139078769422
12 norm of gradients: 0.0667831599170747
12 norm of weights: 2.2264518100390416
-----
Iteration Number: 7908

```

Loss: 11.031843395564039
l2 norm of gradients: 0.06677229004473498
l2 norm of weights: 2.226478829569192

Iteration Number: 7909
Loss: 11.03154787265443
l2 norm of gradients: 0.06676142423667994
l2 norm of weights: 2.2265058616346782

Iteration Number: 7910
Loss: 11.03125250993663
l2 norm of gradients: 0.06675056249079699
l2 norm of weights: 2.2265329062246444

Iteration Number: 7911
Loss: 11.030957307306743
l2 norm of gradients: 0.06673970480497494
l2 norm of weights: 2.2265599633282442

Iteration Number: 7912
Loss: 11.030662264660942
l2 norm of gradients: 0.06672885117710405
l2 norm of weights: 2.2265870329346407

Iteration Number: 7913
Loss: 11.03036738189547
l2 norm of gradients: 0.066718001605076
l2 norm of weights: 2.2266141150330028

Iteration Number: 7914
Loss: 11.030072658906649
l2 norm of gradients: 0.06670715608678382
l2 norm of weights: 2.2266412096125094

Iteration Number: 7915
Loss: 11.029778095590872
l2 norm of gradients: 0.06669631462012215
l2 norm of weights: 2.2266683166623467

Iteration Number: 7916
Loss: 11.029483691844598
l2 norm of gradients: 0.06668547720298687
l2 norm of weights: 2.2266954361717084

Iteration Number: 7917
Loss: 11.029189447564361
l2 norm of gradients: 0.06667464383327538
l2 norm of weights: 2.2267225681297975


```

-----
Iteration Number: 7918
Loss: 11.028895362646766
l2 norm of gradients: 0.06666381450888642
l2 norm of weights: 2.2267497125258244
-----
Iteration Number: 7919
Loss: 11.028601436988499
l2 norm of gradients: 0.06665298922772023
l2 norm of weights: 2.226776869349008
-----
Iteration Number: 7920
Loss: 11.028307670486297
l2 norm of gradients: 0.0666421679876784
l2 norm of weights: 2.226804038588574
-----
Iteration Number: 7921
Loss: 11.02801406303699
l2 norm of gradients: 0.06663135078666403
l2 norm of weights: 2.2268312202337586
-----
Iteration Number: 7922
Loss: 11.027720614537465
l2 norm of gradients: 0.0666205376225815
l2 norm of weights: 2.2268584142738037
-----
Iteration Number: 7923
Loss: 11.027427324884691
l2 norm of gradients: 0.0666097284933367
l2 norm of weights: 2.22688562069796
-----
Iteration Number: 7924
Loss: 11.0271341939757
l2 norm of gradients: 0.0665989233968369
l2 norm of weights: 2.2269128394954874
-----
Iteration Number: 7925
Loss: 11.0268412217076
l2 norm of gradients: 0.0665881223309908
l2 norm of weights: 2.2269400706556524
-----
Iteration Number: 7926
Loss: 11.02654840797757
l2 norm of gradients: 0.06657732529370844
l2 norm of weights: 2.22696731416773
-----
Iteration Number: 7927
Loss: 11.026255752682856

```

```

12 norm of gradients: 0.06656653228290137
12 norm of weights: 2.226994570021003
-----
Iteration Number: 7928
Loss: 11.025963255720786
12 norm of gradients: 0.06655574329648249
12 norm of weights: 2.2270218382047635
-----
Iteration Number: 7929
Loss: 11.025670916988744
12 norm of gradients: 0.06654495833236608
12 norm of weights: 2.22704911870831
-----
Iteration Number: 7930
Loss: 11.025378736384198
12 norm of gradients: 0.06653417738846786
12 norm of weights: 2.22707641152095
-----
Iteration Number: 7931
Loss: 11.025086713804683
12 norm of gradients: 0.06652340046270494
12 norm of weights: 2.2271037166319982
-----
Iteration Number: 7932
Loss: 11.0247948491478
12 norm of gradients: 0.06651262755299583
12 norm of weights: 2.227131034030778
-----
Iteration Number: 7933
Loss: 11.024503142311227
12 norm of gradients: 0.06650185865726044
12 norm of weights: 2.2271583637066206
-----
Iteration Number: 7934
Loss: 11.024211593192717
12 norm of gradients: 0.06649109377342008
12 norm of weights: 2.227185705648866
-----
Iteration Number: 7935
Loss: 11.023920201690075
12 norm of gradients: 0.06648033289939746
12 norm of weights: 2.2272130598468602
-----
Iteration Number: 7936
Loss: 11.023628967701203
12 norm of gradients: 0.06646957603311664
12 norm of weights: 2.2272404262899594
-----

```

Iteration Number: 7937
Loss: 11.02333789112405
l2 norm of gradients: 0.06645882317250312
l2 norm of weights: 2.2272678049675267

Iteration Number: 7938
Loss: 11.023046971856653
l2 norm of gradients: 0.06644807431548377
l2 norm of weights: 2.227295195868933

Iteration Number: 7939
Loss: 11.022756209797112
l2 norm of gradients: 0.0664373294599869
l2 norm of weights: 2.2273225989835574

Iteration Number: 7940
Loss: 11.022465604843598
l2 norm of gradients: 0.06642658860394214
l2 norm of weights: 2.2273500143007867

Iteration Number: 7941
Loss: 11.022175156894356
l2 norm of gradients: 0.06641585174528053
l2 norm of weights: 2.227377441810017

Iteration Number: 7942
Loss: 11.021884865847694
l2 norm of gradients: 0.06640511888193451
l2 norm of weights: 2.2274048815006506

Iteration Number: 7943
Loss: 11.021594731601997
l2 norm of gradients: 0.06639439001183793
l2 norm of weights: 2.2274323333620987

Iteration Number: 7944
Loss: 11.02130475405572
l2 norm of gradients: 0.0663836651329259
l2 norm of weights: 2.2274597973837804

Iteration Number: 7945
Loss: 11.021014933107388
l2 norm of gradients: 0.06637294424313506
l2 norm of weights: 2.227487273555122

Iteration Number: 7946
Loss: 11.020725268655587
l2 norm of gradients: 0.0663622273404034

```
l2 norm of weights: 2.227514761865559
-----
Iteration Number: 7947
Loss: 11.02043576059899
l2 norm of gradients: 0.06635151442267022
l2 norm of weights: 2.227542262304534
-----
Iteration Number: 7948
Loss: 11.020146408836332
l2 norm of gradients: 0.06634080548787623
l2 norm of weights: 2.227569774861497
-----
Iteration Number: 7949
Loss: 11.019857213266409
l2 norm of gradients: 0.06633010053396354
l2 norm of weights: 2.227597299525908
-----
Iteration Number: 7950
Loss: 11.019568173788109
l2 norm of gradients: 0.06631939955887561
l2 norm of weights: 2.227624836287232
-----
Iteration Number: 7951
Loss: 11.019279290300364
l2 norm of gradients: 0.06630870256055728
l2 norm of weights: 2.2276523851349443
-----
Iteration Number: 7952
Loss: 11.018990562702196
l2 norm of gradients: 0.06629800953695483
l2 norm of weights: 2.2276799460585273
-----
Iteration Number: 7953
Loss: 11.018701990892692
l2 norm of gradients: 0.06628732048601575
l2 norm of weights: 2.2277075190474704
-----
Iteration Number: 7954
Loss: 11.018413574771003
l2 norm of gradients: 0.06627663540568904
l2 norm of weights: 2.2277351040912725
-----
Iteration Number: 7955
Loss: 11.018125314236354
l2 norm of gradients: 0.06626595429392505
l2 norm of weights: 2.227762701179439
-----
Iteration Number: 7956
```

Loss: 11.017837209188038
l2 norm of gradients: 0.06625527714867545
l2 norm of weights: 2.2277903103014842

Iteration Number: 7957
Loss: 11.017549259525424
l2 norm of gradients: 0.06624460396789326
l2 norm of weights: 2.2278179314469297

Iteration Number: 7958
Loss: 11.017261465147943
l2 norm of gradients: 0.06623393474953297
l2 norm of weights: 2.2278455646053055

Iteration Number: 7959
Loss: 11.016973825955095
l2 norm of gradients: 0.06622326949155032
l2 norm of weights: 2.227873209766148

Iteration Number: 7960
Loss: 11.016686341846459
l2 norm of gradients: 0.06621260819190244
l2 norm of weights: 2.2279008669190037

Iteration Number: 7961
Loss: 11.016399012721674
l2 norm of gradients: 0.06620195084854788
l2 norm of weights: 2.227928536053425

Iteration Number: 7962
Loss: 11.016111838480452
l2 norm of gradients: 0.06619129745944652
l2 norm of weights: 2.2279562171589733

Iteration Number: 7963
Loss: 11.015824819022576
l2 norm of gradients: 0.06618064802255952
l2 norm of weights: 2.2279839102252175

Iteration Number: 7964
Loss: 11.015537954247899
l2 norm of gradients: 0.0661700025358495
l2 norm of weights: 2.2280116152417344

Iteration Number: 7965
Loss: 11.015251244056337
l2 norm of gradients: 0.06615936099728038
l2 norm of weights: 2.2280393321981076

```

-----
Iteration Number: 7966
Loss: 11.014964688347883
l2 norm of gradients: 0.06614872340481746
l2 norm of weights: 2.2280670610839306
-----
Iteration Number: 7967
Loss: 11.014678287022594
l2 norm of gradients: 0.06613808975642738
l2 norm of weights: 2.228094801888803
-----
Iteration Number: 7968
Loss: 11.014392039980596
l2 norm of gradients: 0.06612746005007808
l2 norm of weights: 2.2281225546023333
-----
Iteration Number: 7969
Loss: 11.01410594712209
l2 norm of gradients: 0.06611683428373898
l2 norm of weights: 2.2281503192141368
-----
Iteration Number: 7970
Loss: 11.013820008347341
l2 norm of gradients: 0.06610621245538074
l2 norm of weights: 2.228178095713837
-----
Iteration Number: 7971
Loss: 11.013534223556682
l2 norm of gradients: 0.06609559456297537
l2 norm of weights: 2.2282058840910657
-----
Iteration Number: 7972
Loss: 11.01324859265052
l2 norm of gradients: 0.06608498060449627
l2 norm of weights: 2.228233684335462
-----
Iteration Number: 7973
Loss: 11.012963115529326
l2 norm of gradients: 0.06607437057791817
l2 norm of weights: 2.2282614964366725
-----
Iteration Number: 7974
Loss: 11.012677792093642
l2 norm of gradients: 0.06606376448121712
l2 norm of weights: 2.2282893203843526
-----
Iteration Number: 7975
Loss: 11.012392622244079

```

```

12 norm of gradients: 0.06605316231237056
12 norm of weights: 2.228317156168164
-----
Iteration Number: 7976
Loss: 11.012107605881317
12 norm of gradients: 0.06604256406935721
12 norm of weights: 2.2283450037777777
-----
Iteration Number: 7977
Loss: 11.011822742906105
12 norm of gradients: 0.06603196975015722
12 norm of weights: 2.228372863202872
-----
Iteration Number: 7978
Loss: 11.01153803321926
12 norm of gradients: 0.06602137935275196
12 norm of weights: 2.228400734433132
-----
Iteration Number: 7979
Loss: 11.011253476721663
12 norm of gradients: 0.0660107928751242
12 norm of weights: 2.2284286174582513
-----
Iteration Number: 7980
Loss: 11.010969073314275
12 norm of gradients: 0.06600021031525803
12 norm of weights: 2.2284565122679316
-----
Iteration Number: 7981
Loss: 11.010684822898114
12 norm of gradients: 0.06598963167113896
12 norm of weights: 2.2284844188518815
-----
Iteration Number: 7982
Loss: 11.010400725374272
12 norm of gradients: 0.06597905694075364
12 norm of weights: 2.228512337199819
-----
Iteration Number: 7983
Loss: 11.010116780643909
12 norm of gradients: 0.06596848612209023
12 norm of weights: 2.228540267301467
-----
Iteration Number: 7984
Loss: 11.009832988608252
12 norm of gradients: 0.06595791921313818
12 norm of weights: 2.228568209146559
-----

```

Iteration Number: 7985
Loss: 11.009549349168598
l2 norm of gradients: 0.06594735621188817
l2 norm of weights: 2.2285961627248345

Iteration Number: 7986
Loss: 11.009265862226313
l2 norm of gradients: 0.06593679711633238
l2 norm of weights: 2.228624128026041

Iteration Number: 7987
Loss: 11.008982527682825
l2 norm of gradients: 0.06592624192446415
l2 norm of weights: 2.2286521050399344

Iteration Number: 7988
Loss: 11.00869934543964
l2 norm of gradients: 0.06591569063427823
l2 norm of weights: 2.228680093756277

Iteration Number: 7989
Loss: 11.008416315398325
l2 norm of gradients: 0.0659051432437707
l2 norm of weights: 2.228708094164841

Iteration Number: 7990
Loss: 11.008133437460517
l2 norm of gradients: 0.06589459975093892
l2 norm of weights: 2.2287361062554036

Iteration Number: 7991
Loss: 11.00785071152792
l2 norm of gradients: 0.06588406015378159
l2 norm of weights: 2.228764130017751

Iteration Number: 7992
Loss: 11.007568137502311
l2 norm of gradients: 0.06587352445029868
l2 norm of weights: 2.2287921654416785

Iteration Number: 7993
Loss: 11.007285715285521
l2 norm of gradients: 0.06586299263849162
l2 norm of weights: 2.228820212516986

Iteration Number: 7994
Loss: 11.007003444779471
l2 norm of gradients: 0.06585246471636304


```
12 norm of weights: 2.228848271233484
-----
Iteration Number: 7995
Loss: 11.006721325886133
12 norm of gradients: 0.06584194068191693
12 norm of weights: 2.2288763415809885
-----
Iteration Number: 7996
Loss: 11.006439358507548
12 norm of gradients: 0.06583142053315846
12 norm of weights: 2.2289044235493245
-----
Iteration Number: 7997
Loss: 11.00615754254583
12 norm of gradients: 0.0658209042680944
12 norm of weights: 2.2289325171283236
-----
Iteration Number: 7998
Loss: 11.005875877903163
12 norm of gradients: 0.06581039188473253
12 norm of weights: 2.2289606223078264
-----
Iteration Number: 7999
Loss: 11.005594364481789
12 norm of gradients: 0.0657998833810821
12 norm of weights: 2.2289887390776806
-----
Iteration Number: 8000
Loss: 11.005313002184028
12 norm of gradients: 0.0657893787551537
12 norm of weights: 2.2290168674277404
-----
Iteration Number: 8001
Loss: 11.005031790912255
12 norm of gradients: 0.06577887800495909
12 norm of weights: 2.229045007347869
-----
Iteration Number: 8002
Loss: 11.004750730568928
12 norm of gradients: 0.0657683811285115
12 norm of weights: 2.229073158827937
-----
Iteration Number: 8003
Loss: 11.004469821056562
12 norm of gradients: 0.0657578881238253
12 norm of weights: 2.2291013218578226
-----
Iteration Number: 8004
```

Loss: 11.00418906227774
l2 norm of gradients: 0.0657473989889163
l2 norm of weights: 2.229129496427411

Iteration Number: 8005
Loss: 11.003908454135113
l2 norm of gradients: 0.06573691372180154
l2 norm of weights: 2.2291576825265955

Iteration Number: 8006
Loss: 11.003627996531405
l2 norm of gradients: 0.06572643232049935
l2 norm of weights: 2.229185880145277

Iteration Number: 8007
Loss: 11.003347689369399
l2 norm of gradients: 0.06571595478302947
l2 norm of weights: 2.2292140892733645

Iteration Number: 8008
Loss: 11.003067532551947
l2 norm of gradients: 0.06570548110741281
l2 norm of weights: 2.229242309900773

Iteration Number: 8009
Loss: 11.002787525981978
l2 norm of gradients: 0.0656950112916716
l2 norm of weights: 2.2292705420174275

Iteration Number: 8010
Loss: 11.002507669562473
l2 norm of gradients: 0.06568454533382945
l2 norm of weights: 2.229298785613258

Iteration Number: 8011
Loss: 11.002227963196486
l2 norm of gradients: 0.06567408323191123
l2 norm of weights: 2.2293270406782044

Iteration Number: 8012
Loss: 11.001948406787145
l2 norm of gradients: 0.06566362498394301
l2 norm of weights: 2.229355307202212

Iteration Number: 8013
Loss: 11.001669000237634
l2 norm of gradients: 0.06565317058795228
l2 norm of weights: 2.2293835851752357

```

-----
Iteration Number: 8014
Loss: 11.001389743451211
l2 norm of gradients: 0.06564272004196775
l2 norm of weights: 2.2294118745872367
-----
Iteration Number: 8015
Loss: 11.0011106363312
l2 norm of gradients: 0.06563227334401943
l2 norm of weights: 2.2294401754281843
-----
Iteration Number: 8016
Loss: 11.000831678780989
l2 norm of gradients: 0.06562183049213868
l2 norm of weights: 2.2294684876880546
-----
Iteration Number: 8017
Loss: 11.000552870704032
l2 norm of gradients: 0.06561139148435804
l2 norm of weights: 2.229496811356832
-----
Iteration Number: 8018
Loss: 11.000274212003855
l2 norm of gradients: 0.06560095631871145
l2 norm of weights: 2.2295251464245087
-----
Iteration Number: 8019
Loss: 10.999995702584048
l2 norm of gradients: 0.065590524993234
l2 norm of weights: 2.2295534928810836
-----
Iteration Number: 8020
Loss: 10.999717342348264
l2 norm of gradients: 0.06558009750596222
l2 norm of weights: 2.229581850716564
-----
Iteration Number: 8021
Loss: 10.999439131200228
l2 norm of gradients: 0.06556967385493383
l2 norm of weights: 2.2296102199209633
-----
Iteration Number: 8022
Loss: 10.999161069043726
l2 norm of gradients: 0.06555925403818783
l2 norm of weights: 2.2296386004843036
-----
Iteration Number: 8023
Loss: 10.99888315578262

```

l2 norm of gradients: 0.06554883805376455
l2 norm of weights: 2.2296669923966155

Iteration Number: 8024
Loss: 10.998605391320824
l2 norm of gradients: 0.0655384258997055
l2 norm of weights: 2.2296953956479344

Iteration Number: 8025
Loss: 10.998327775562336
l2 norm of gradients: 0.06552801757405363
l2 norm of weights: 2.2297238102283052

Iteration Number: 8026
Loss: 10.998050308411202
l2 norm of gradients: 0.06551761307485304
l2 norm of weights: 2.22975223612778

Iteration Number: 8027
Loss: 10.997772989771546
l2 norm of gradients: 0.06550721240014912
l2 norm of weights: 2.229780673336418

Iteration Number: 8028
Loss: 10.997495819547556
l2 norm of gradients: 0.06549681554798858
l2 norm of weights: 2.2298091218442866

Iteration Number: 8029
Loss: 10.997218797643484
l2 norm of gradients: 0.06548642251641937
l2 norm of weights: 2.2298375816414593

Iteration Number: 8030
Loss: 10.996941923963652
l2 norm of gradients: 0.06547603330349071
l2 norm of weights: 2.2298660527180183

Iteration Number: 8031
Loss: 10.996665198412444
l2 norm of gradients: 0.06546564790725308
l2 norm of weights: 2.2298945350640533

Iteration Number: 8032
Loss: 10.996388620894308
l2 norm of gradients: 0.0654552663257583
l2 norm of weights: 2.2299230286696607

Iteration Number: 8033
Loss: 10.996112191313767
l2 norm of gradients: 0.06544488855705938
l2 norm of weights: 2.2299515335249454

Iteration Number: 8034
Loss: 10.9958359095754
l2 norm of gradients: 0.06543451459921064
l2 norm of weights: 2.2299800496200177

Iteration Number: 8035
Loss: 10.99555977558386
l2 norm of gradients: 0.06542414445026759
l2 norm of weights: 2.230008576944998

Iteration Number: 8036
Loss: 10.995283789243858
l2 norm of gradients: 0.06541377810828715
l2 norm of weights: 2.230037115490013

Iteration Number: 8037
Loss: 10.99500795046018
l2 norm of gradients: 0.06540341557132737
l2 norm of weights: 2.2300656652451956

Iteration Number: 8038
Loss: 10.994732259137665
l2 norm of gradients: 0.06539305683744767
l2 norm of weights: 2.230094226200688

Iteration Number: 8039
Loss: 10.994456715181233
l2 norm of gradients: 0.06538270190470856
l2 norm of weights: 2.23012279834664

Iteration Number: 8040
Loss: 10.994181318495858
l2 norm of gradients: 0.06537235077117205
l2 norm of weights: 2.2301513816732066

Iteration Number: 8041
Loss: 10.993906068986586
l2 norm of gradients: 0.06536200343490121
l2 norm of weights: 2.2301799761705525

Iteration Number: 8042
Loss: 10.99363096655852
l2 norm of gradients: 0.06535165989396041

```
l2 norm of weights: 2.230208581828848
-----
Iteration Number: 8043
Loss: 10.993356011116841
l2 norm of gradients: 0.0653413201464154
l2 norm of weights: 2.2302371986382727
-----
Iteration Number: 8044
Loss: 10.993081202566787
l2 norm of gradients: 0.065330984190333
l2 norm of weights: 2.230265826589012
-----
Iteration Number: 8045
Loss: 10.992806540813664
l2 norm of gradients: 0.06532065202378141
l2 norm of weights: 2.2302944656712596
-----
Iteration Number: 8046
Loss: 10.992532025762836
l2 norm of gradients: 0.06531032364483003
l2 norm of weights: 2.230323115875216
-----
Iteration Number: 8047
Loss: 10.99225765731975
l2 norm of gradients: 0.06529999905154955
l2 norm of weights: 2.2303517771910895
-----
Iteration Number: 8048
Loss: 10.991983435389898
l2 norm of gradients: 0.06528967824201189
l2 norm of weights: 2.230380449609096
-----
Iteration Number: 8049
Loss: 10.991709359878852
l2 norm of gradients: 0.06527936121429016
l2 norm of weights: 2.2304091331194584
-----
Iteration Number: 8050
Loss: 10.99143543069224
l2 norm of gradients: 0.06526904796645888
l2 norm of weights: 2.2304378277124064
-----
Iteration Number: 8051
Loss: 10.991161647735762
l2 norm of gradients: 0.06525873849659358
l2 norm of weights: 2.2304665333781784
-----
Iteration Number: 8052
```

```

Loss: 10.990888010915175
l2 norm of gradients: 0.06524843280277129
l2 norm of weights: 2.2304952501070194
-----
Iteration Number: 8053
Loss: 10.99061452013631
l2 norm of gradients: 0.06523813088307007
l2 norm of weights: 2.230523977889182
-----
Iteration Number: 8054
Loss: 10.990341175305057
l2 norm of gradients: 0.06522783273556935
l2 norm of weights: 2.230552716714925
-----
Iteration Number: 8055
Loss: 10.990067976327373
l2 norm of gradients: 0.06521753835834977
l2 norm of weights: 2.2305814665745167
-----
Iteration Number: 8056
Loss: 10.989794923109278
l2 norm of gradients: 0.06520724774949319
l2 norm of weights: 2.2306102274582313
-----
Iteration Number: 8057
Loss: 10.989522015556862
l2 norm of gradients: 0.06519696090708273
l2 norm of weights: 2.2306389993563505
-----
Iteration Number: 8058
Loss: 10.989249253576272
l2 norm of gradients: 0.06518667782920273
l2 norm of weights: 2.2306677822591627
-----
Iteration Number: 8059
Loss: 10.988976637073725
l2 norm of gradients: 0.06517639851393885
l2 norm of weights: 2.230696576156966
-----
Iteration Number: 8060
Loss: 10.9887041659555
l2 norm of gradients: 0.06516612295937779
l2 norm of weights: 2.230725381040063
-----
Iteration Number: 8061
Loss: 10.98843184012795
l2 norm of gradients: 0.06515585116360771
l2 norm of weights: 2.230754196898765

```

```

-----
Iteration Number: 8062
Loss: 10.988159659497475
l2 norm of gradients: 0.06514558312471794
l2 norm of weights: 2.2307830237233905
-----
Iteration Number: 8063
Loss: 10.987887623970552
l2 norm of gradients: 0.06513531884079887
l2 norm of weights: 2.230811861504266
-----
Iteration Number: 8064
Loss: 10.987615733453724
l2 norm of gradients: 0.06512505830994236
l2 norm of weights: 2.2308407102317234
-----
Iteration Number: 8065
Loss: 10.987343987853588
l2 norm of gradients: 0.06511480153024136
l2 norm of weights: 2.2308695698961034
-----
Iteration Number: 8066
Loss: 10.987072387076816
l2 norm of gradients: 0.06510454849979012
l2 norm of weights: 2.2308984404877537
-----
Iteration Number: 8067
Loss: 10.98680093103014
l2 norm of gradients: 0.06509429921668405
l2 norm of weights: 2.23092732199703
-----
Iteration Number: 8068
Loss: 10.986529619620352
l2 norm of gradients: 0.06508405367901987
l2 norm of weights: 2.230956214414293
-----
Iteration Number: 8069
Loss: 10.986258452754317
l2 norm of gradients: 0.06507381188489539
l2 norm of weights: 2.230985117729913
-----
Iteration Number: 8070
Loss: 10.985987430338959
l2 norm of gradients: 0.06506357383240984
l2 norm of weights: 2.231014031934267
-----
Iteration Number: 8071
Loss: 10.985716552281264

```



```

12 norm of gradients: 0.06505333951966344
12 norm of weights: 2.2310429570177384
-----
Iteration Number: 8072
Loss: 10.985445818488287
12 norm of gradients: 0.06504310894475786
12 norm of weights: 2.231071892970719
-----
Iteration Number: 8073
Loss: 10.985175228867147
12 norm of gradients: 0.06503288210579584
12 norm of weights: 2.2311008397836076
-----
Iteration Number: 8074
Loss: 10.984904783325018
12 norm of gradients: 0.0650226590008814
12 norm of weights: 2.2311297974468096
-----
Iteration Number: 8075
Loss: 10.984634481769158
12 norm of gradients: 0.06501243962811977
12 norm of weights: 2.2311587659507373
-----
Iteration Number: 8076
Loss: 10.984364324106863
12 norm of gradients: 0.0650022239856174
12 norm of weights: 2.231187745285812
-----
Iteration Number: 8077
Loss: 10.984094310245514
12 norm of gradients: 0.06499201207148195
12 norm of weights: 2.2312167354424606
-----
Iteration Number: 8078
Loss: 10.983824440092542
12 norm of gradients: 0.06498180388382221
12 norm of weights: 2.2312457364111182
-----
Iteration Number: 8079
Loss: 10.983554713555453
12 norm of gradients: 0.06497159942074839
12 norm of weights: 2.2312747481822264
-----
Iteration Number: 8080
Loss: 10.98328513054181
12 norm of gradients: 0.06496139868037171
12 norm of weights: 2.231303770746235
-----

```

Iteration Number: 8081
Loss: 10.983015690959238
l2 norm of gradients: 0.06495120166080469
l2 norm of weights: 2.2313328040936

Iteration Number: 8082
Loss: 10.98274639471543
l2 norm of gradients: 0.06494100836016108
l2 norm of weights: 2.231361848214785

Iteration Number: 8083
Loss: 10.982477241718144
l2 norm of gradients: 0.06493081877655578
l2 norm of weights: 2.2313909031002606

Iteration Number: 8084
Loss: 10.982208231875198
l2 norm of gradients: 0.064920632908105
l2 norm of weights: 2.231419968740505

Iteration Number: 8085
Loss: 10.981939365094473
l2 norm of gradients: 0.06491045075292598
l2 norm of weights: 2.231449045126004

Iteration Number: 8086
Loss: 10.981670641283912
l2 norm of gradients: 0.06490027230913735
l2 norm of weights: 2.231478132247249

Iteration Number: 8087
Loss: 10.981402060351531
l2 norm of gradients: 0.06489009757485882
l2 norm of weights: 2.23150723009474

Iteration Number: 8088
Loss: 10.981133622205398
l2 norm of gradients: 0.06487992654821136
l2 norm of weights: 2.2315363386589837

Iteration Number: 8089
Loss: 10.98086532675365
l2 norm of gradients: 0.06486975922731716
l2 norm of weights: 2.231565457930494

Iteration Number: 8090
Loss: 10.980597173904487
l2 norm of gradients: 0.06485959561029957

```
l2 norm of weights: 2.2315945878997923
-----
Iteration Number: 8091
Loss: 10.980329163566172
l2 norm of gradients: 0.06484943569528315
l2 norm of weights: 2.2316237285574068
-----
Iteration Number: 8092
Loss: 10.980061295647028
l2 norm of gradients: 0.06483927948039363
l2 norm of weights: 2.2316528798938724
-----
Iteration Number: 8093
Loss: 10.979793570055447
l2 norm of gradients: 0.06482912696375799
l2 norm of weights: 2.2316820418997323
-----
Iteration Number: 8094
Loss: 10.97952598669988
l2 norm of gradients: 0.06481897814350442
l2 norm of weights: 2.231711214565536
-----
Iteration Number: 8095
Loss: 10.979258545488843
l2 norm of gradients: 0.06480883301776227
l2 norm of weights: 2.23174039788184
-----
Iteration Number: 8096
Loss: 10.97899124633091
l2 norm of gradients: 0.06479869158466206
l2 norm of weights: 2.231769591839209
-----
Iteration Number: 8097
Loss: 10.978724089134726
l2 norm of gradients: 0.06478855384233553
l2 norm of weights: 2.2317987964282135
-----
Iteration Number: 8098
Loss: 10.978457073808993
l2 norm of gradients: 0.06477841978891559
l2 norm of weights: 2.2318280116394322
-----
Iteration Number: 8099
Loss: 10.978190200262484
l2 norm of gradients: 0.0647682894225364
l2 norm of weights: 2.2318572374634504
-----
Iteration Number: 8100
```

Loss: 10.97792346840402
l2 norm of gradients: 0.0647581627413333
l2 norm of weights: 2.23188647389086

Iteration Number: 8101
Loss: 10.977656878142499
l2 norm of gradients: 0.06474803974344273
l2 norm of weights: 2.2319157209122618

Iteration Number: 8102
Loss: 10.97739042938687
l2 norm of gradients: 0.06473792042700244
l2 norm of weights: 2.2319449785182615

Iteration Number: 8103
Loss: 10.977124122046162
l2 norm of gradients: 0.06472780479015122
l2 norm of weights: 2.2319742466994734

Iteration Number: 8104
Loss: 10.976857956029443
l2 norm of gradients: 0.06471769283102924
l2 norm of weights: 2.2320035254465185

Iteration Number: 8105
Loss: 10.976591931245867
l2 norm of gradients: 0.06470758454777766
l2 norm of weights: 2.232032814750024

Iteration Number: 8106
Loss: 10.976326047604632
l2 norm of gradients: 0.06469747993853896
l2 norm of weights: 2.2320621146006268

Iteration Number: 8107
Loss: 10.97606030501501
l2 norm of gradients: 0.06468737900145671
l2 norm of weights: 2.2320914249889667

Iteration Number: 8108
Loss: 10.975794703386331
l2 norm of gradients: 0.06467728173467577
l2 norm of weights: 2.232120745905695

Iteration Number: 8109
Loss: 10.975529242627992
l2 norm of gradients: 0.06466718813634205
l2 norm of weights: 2.2321500773414673

```

-----
Iteration Number: 8110
Loss: 10.97526392264944
l2 norm of gradients: 0.06465709820460273
l2 norm of weights: 2.2321794192869464
-----
Iteration Number: 8111
Loss: 10.974998743360198
l2 norm of gradients: 0.06464701193760612
l2 norm of weights: 2.2322087717328034
-----
Iteration Number: 8112
Loss: 10.974733704669848
l2 norm of gradients: 0.06463692933350176
l2 norm of weights: 2.232238134669716
-----
Iteration Number: 8113
Loss: 10.97446880648803
l2 norm of gradients: 0.06462685039044033
l2 norm of weights: 2.232267508088369
-----
Iteration Number: 8114
Loss: 10.974204048724447
l2 norm of gradients: 0.06461677510657364
l2 norm of weights: 2.232296891979453
-----
Iteration Number: 8115
Loss: 10.973939431288871
l2 norm of gradients: 0.06460670348005478
l2 norm of weights: 2.232326286333667
-----
Iteration Number: 8116
Loss: 10.973674954091125
l2 norm of gradients: 0.0645966355090379
l2 norm of weights: 2.2323556911417173
-----
Iteration Number: 8117
Loss: 10.97341061704111
l2 norm of gradients: 0.06458657119167838
l2 norm of weights: 2.232385106394316
-----
Iteration Number: 8118
Loss: 10.973146420048767
l2 norm of gradients: 0.06457651052613281
l2 norm of weights: 2.2324145320821835
-----
Iteration Number: 8119
Loss: 10.972882363024112

```

```

12 norm of gradients: 0.06456645351055883
12 norm of weights: 2.2324439681960455
-----
Iteration Number: 8120
Loss: 10.972618445877233
12 norm of gradients: 0.06455640014311538
12 norm of weights: 2.232473414726637
-----
Iteration Number: 8121
Loss: 10.972354668518259
12 norm of gradients: 0.0645463504219625
12 norm of weights: 2.2325028716646984
-----
Iteration Number: 8122
Loss: 10.972091030857394
12 norm of gradients: 0.06453630434526139
12 norm of weights: 2.232532339000978
-----
Iteration Number: 8123
Loss: 10.971827532804895
12 norm of gradients: 0.06452626191117437
12 norm of weights: 2.2325618167262293
-----
Iteration Number: 8124
Loss: 10.971564174271096
12 norm of gradients: 0.0645162231178651
12 norm of weights: 2.232591304831215
-----
Iteration Number: 8125
Loss: 10.97130095516638
12 norm of gradients: 0.06450618796349819
12 norm of weights: 2.2326208033067045
-----
Iteration Number: 8126
Loss: 10.971037875401187
12 norm of gradients: 0.06449615644623954
12 norm of weights: 2.2326503121434724
-----
Iteration Number: 8127
Loss: 10.970774934886036
12 norm of gradients: 0.06448612856425619
12 norm of weights: 2.232679831332302
-----
Iteration Number: 8128
Loss: 10.970512133531493
12 norm of gradients: 0.06447610431571629
12 norm of weights: 2.2327093608639834
-----

```

Iteration Number: 8129
Loss: 10.970249471248191
l2 norm of gradients: 0.06446608369878921
l2 norm of weights: 2.232738900729313

Iteration Number: 8130
Loss: 10.969986947946822
l2 norm of gradients: 0.0644560667116454
l2 norm of weights: 2.232768450919094

Iteration Number: 8131
Loss: 10.969724563538145
l2 norm of gradients: 0.06444605335245657
l2 norm of weights: 2.2327980114241384

Iteration Number: 8132
Loss: 10.969462317932974
l2 norm of gradients: 0.06443604361939549
l2 norm of weights: 2.232827582235263

Iteration Number: 8133
Loss: 10.96920021104219
l2 norm of gradients: 0.06442603751063614
l2 norm of weights: 2.2328571633432923

Iteration Number: 8134
Loss: 10.96893824277673
l2 norm of gradients: 0.06441603502435364
l2 norm of weights: 2.2328867547390576

Iteration Number: 8135
Loss: 10.968676413047593
l2 norm of gradients: 0.06440603615872427
l2 norm of weights: 2.2329163564133983

Iteration Number: 8136
Loss: 10.968414721765843
l2 norm of gradients: 0.06439604091192543
l2 norm of weights: 2.2329459683571584

Iteration Number: 8137
Loss: 10.968153168842605
l2 norm of gradients: 0.0643860492821357
l2 norm of weights: 2.2329755905611917

Iteration Number: 8138
Loss: 10.96789175418906
l2 norm of gradients: 0.06437606126753481

```

12 norm of weights: 2.233005223016357
-----
Iteration Number: 8139
Loss: 10.967630477716453
12 norm of gradients: 0.06436607686630356
12 norm of weights: 2.23303486571352
-----
Iteration Number: 8140
Loss: 10.967369339336095
12 norm of gradients: 0.064356096076624
12 norm of weights: 2.233064518643554
-----
Iteration Number: 8141
Loss: 10.967108338959346
12 norm of gradients: 0.06434611889667934
12 norm of weights: 2.233094181797339
-----
Iteration Number: 8142
Loss: 10.96684747649764
12 norm of gradients: 0.06433614532465377
12 norm of weights: 2.2331238551657626
-----
Iteration Number: 8143
Loss: 10.966586751862465
12 norm of gradients: 0.0643261753587328
12 norm of weights: 2.233153538739718
-----
Iteration Number: 8144
Loss: 10.96632616496537
12 norm of gradients: 0.06431620899710304
12 norm of weights: 2.2331832325101058
-----
Iteration Number: 8145
Loss: 10.966065715717967
12 norm of gradients: 0.06430624623795214
12 norm of weights: 2.233212936467834
-----
Iteration Number: 8146
Loss: 10.965805404031926
12 norm of gradients: 0.06429628707946904
12 norm of weights: 2.2332426506038168
-----
Iteration Number: 8147
Loss: 10.96554522981898
12 norm of gradients: 0.0642863315198437
12 norm of weights: 2.233272374908976
-----
Iteration Number: 8148

```


Loss: 10.965285192990924
l2 norm of gradients: 0.06427637955726725
l2 norm of weights: 2.2333021093742396

Iteration Number: 8149
Loss: 10.96502529345961
l2 norm of gradients: 0.06426643118993203
l2 norm of weights: 2.233331853990543

Iteration Number: 8150
Loss: 10.964765531136951
l2 norm of gradients: 0.06425648641603139
l2 norm of weights: 2.2333616087488277

Iteration Number: 8151
Loss: 10.964505905934926
l2 norm of gradients: 0.06424654523375994
l2 norm of weights: 2.233391373640043

Iteration Number: 8152
Loss: 10.964246417765569
l2 norm of gradients: 0.06423660764131332
l2 norm of weights: 2.2334211486551445

Iteration Number: 8153
Loss: 10.963987066540973
l2 norm of gradients: 0.06422667363688836
l2 norm of weights: 2.233450933785095

Iteration Number: 8154
Loss: 10.963727852173298
l2 norm of gradients: 0.06421674321868297
l2 norm of weights: 2.233480729020864

Iteration Number: 8155
Loss: 10.96346877457476
l2 norm of gradients: 0.06420681638489632
l2 norm of weights: 2.233510534353427

Iteration Number: 8156
Loss: 10.963209833657634
l2 norm of gradients: 0.06419689313372852
l2 norm of weights: 2.2335403497737683

Iteration Number: 8157
Loss: 10.962951029334258
l2 norm of gradients: 0.06418697346338097
l2 norm of weights: 2.233570175272877

```
-----  
Iteration Number: 8158  
Loss: 10.962692361517032  
l2 norm of gradients: 0.06417705737205609  
l2 norm of weights: 2.2336000108417498  
-----  
Iteration Number: 8159  
Loss: 10.962433830118414  
l2 norm of gradients: 0.06416714485795749  
l2 norm of weights: 2.2336298564713912  
-----  
Iteration Number: 8160  
Loss: 10.96217543505092  
l2 norm of gradients: 0.06415723591928991  
l2 norm of weights: 2.233659712152811  
-----  
Iteration Number: 8161  
Loss: 10.961917176227127  
l2 norm of gradients: 0.06414733055425917  
l2 norm of weights: 2.2336895778770263  
-----  
Iteration Number: 8162  
Loss: 10.961659053559675  
l2 norm of gradients: 0.0641374287610722  
l2 norm of weights: 2.2337194536350617  
-----  
Iteration Number: 8163  
Loss: 10.961401066961264  
l2 norm of gradients: 0.06412753053793714  
l2 norm of weights: 2.2337493394179475  
-----  
Iteration Number: 8164  
Loss: 10.96114321634465  
l2 norm of gradients: 0.06411763588306314  
l2 norm of weights: 2.233779235216722  
-----  
Iteration Number: 8165  
Loss: 10.960885501622656  
l2 norm of gradients: 0.06410774479466055  
l2 norm of weights: 2.2338091410224292  
-----  
Iteration Number: 8166  
Loss: 10.960627922708154  
l2 norm of gradients: 0.06409785727094083  
l2 norm of weights: 2.2338390568261204  
-----  
Iteration Number: 8167  
Loss: 10.960370479514086
```

```

12 norm of gradients: 0.06408797331011656
12 norm of weights: 2.2338689826188536
-----
Iteration Number: 8168
Loss: 10.960113171953449
12 norm of gradients: 0.06407809291040138
12 norm of weights: 2.233898918391694
-----
Iteration Number: 8169
Loss: 10.959855999939297
12 norm of gradients: 0.06406821607001009
12 norm of weights: 2.233928864135713
-----
Iteration Number: 8170
Loss: 10.959598963384755
12 norm of gradients: 0.06405834278715863
12 norm of weights: 2.233958819841989
-----
Iteration Number: 8171
Loss: 10.959342062202994
12 norm of gradients: 0.064048473060064
12 norm of weights: 2.233988785501606
-----
Iteration Number: 8172
Loss: 10.959085296307256
12 norm of gradients: 0.06403860688694435
12 norm of weights: 2.234018761105658
-----
Iteration Number: 8173
Loss: 10.958828665610831
12 norm of gradients: 0.06402874426601894
12 norm of weights: 2.234048746645242
-----
Iteration Number: 8174
Loss: 10.95857217002708
12 norm of gradients: 0.06401888519550808
12 norm of weights: 2.2340787421114645
-----
Iteration Number: 8175
Loss: 10.958315809469418
12 norm of gradients: 0.06400902967363331
12 norm of weights: 2.234108747495437
-----
Iteration Number: 8176
Loss: 10.958059583851318
12 norm of gradients: 0.06399917769861717
12 norm of weights: 2.234138762788278
-----

```

Iteration Number: 8177
Loss: 10.95780349308632
l2 norm of gradients: 0.06398932926868338
l2 norm of weights: 2.234168787981114

Iteration Number: 8178
Loss: 10.957547537088013
l2 norm of gradients: 0.0639794843820567
l2 norm of weights: 2.234198823065077

Iteration Number: 8179
Loss: 10.957291715770047
l2 norm of gradients: 0.06396964303696302
l2 norm of weights: 2.2342288680313067

Iteration Number: 8180
Loss: 10.95703602904614
l2 norm of gradients: 0.06395980523162938
l2 norm of weights: 2.2342589228709477

Iteration Number: 8181
Loss: 10.956780476830065
l2 norm of gradients: 0.0639499709642839
l2 norm of weights: 2.2342889875751535

Iteration Number: 8182
Loss: 10.95652505903565
l2 norm of gradients: 0.06394014023315575
l2 norm of weights: 2.2343190621350826

Iteration Number: 8183
Loss: 10.95626977557679
l2 norm of gradients: 0.0639303130364753
l2 norm of weights: 2.2343491465419016

Iteration Number: 8184
Loss: 10.956014626367423
l2 norm of gradients: 0.0639204893724739
l2 norm of weights: 2.234379240786783

Iteration Number: 8185
Loss: 10.955759611321573
l2 norm of gradients: 0.06391066923938411
l2 norm of weights: 2.234409344860906

Iteration Number: 8186
Loss: 10.955504730353296
l2 norm of gradients: 0.0639008526354395

```

12 norm of weights: 2.2344394587554564
-----
Iteration Number: 8187
Loss: 10.955249983376723
12 norm of gradients: 0.06389103955887485
12 norm of weights: 2.2344695824616276
-----
Iteration Number: 8188
Loss: 10.954995370306039
12 norm of gradients: 0.06388123000792591
12 norm of weights: 2.234499715970619
-----
Iteration Number: 8189
Loss: 10.954740891055492
12 norm of gradients: 0.06387142398082964
12 norm of weights: 2.2345298592736356
-----
Iteration Number: 8190
Loss: 10.95448654553938
12 norm of gradients: 0.06386162147582394
12 norm of weights: 2.2345600123618916
-----
Iteration Number: 8191
Loss: 10.954232333672067
12 norm of gradients: 0.063851822491148
12 norm of weights: 2.2345901752266055
-----
Iteration Number: 8192
Loss: 10.953978255367975
12 norm of gradients: 0.06384202702504198
12 norm of weights: 2.2346203478590043
-----
Iteration Number: 8193
Loss: 10.953724310541585
12 norm of gradients: 0.06383223507574715
12 norm of weights: 2.2346505302503195
-----
Iteration Number: 8194
Loss: 10.953470499107432
12 norm of gradients: 0.06382244664150587
12 norm of weights: 2.2346807223917917
-----
Iteration Number: 8195
Loss: 10.953216820980119
12 norm of gradients: 0.0638126617205616
12 norm of weights: 2.2347109242746663
-----
Iteration Number: 8196

```

Loss: 10.952963276074298
l2 norm of gradients: 0.06380288031115888
l2 norm of weights: 2.2347411358901965

Iteration Number: 8197
Loss: 10.952709864304685
l2 norm of gradients: 0.06379310241154339
l2 norm of weights: 2.2347713572296413

Iteration Number: 8198
Loss: 10.952456585586052
l2 norm of gradients: 0.06378332801996178
l2 norm of weights: 2.2348015882842667

Iteration Number: 8199
Loss: 10.952203439833232
l2 norm of gradients: 0.06377355713466194
l2 norm of weights: 2.2348318290453455

Iteration Number: 8200
Loss: 10.951950426961115
l2 norm of gradients: 0.06376378975389267
l2 norm of weights: 2.2348620795041567

Iteration Number: 8201
Loss: 10.951697546884647
l2 norm of gradients: 0.06375402587590405
l2 norm of weights: 2.2348923396519873

Iteration Number: 8202
Loss: 10.951444799518839
l2 norm of gradients: 0.06374426549894706
l2 norm of weights: 2.234922609480128

Iteration Number: 8203
Loss: 10.951192184778751
l2 norm of gradients: 0.06373450862127386
l2 norm of weights: 2.2349528889798798

Iteration Number: 8204
Loss: 10.950939702579511
l2 norm of gradients: 0.06372475524113766
l2 norm of weights: 2.2349831781425467

Iteration Number: 8205
Loss: 10.950687352836301
l2 norm of gradients: 0.06371500535679281
l2 norm of weights: 2.235013476959442

```

-----
Iteration Number: 8206
Loss: 10.95043513546436
l2 norm of gradients: 0.06370525896649466
l2 norm of weights: 2.235043785421885
-----
Iteration Number: 8207
Loss: 10.950183050378987
l2 norm of gradients: 0.06369551606849964
l2 norm of weights: 2.2350741035212
-----
Iteration Number: 8208
Loss: 10.949931097495536
l2 norm of gradients: 0.06368577666106535
l2 norm of weights: 2.23510443124872
-----
Iteration Number: 8209
Loss: 10.949679276729425
l2 norm of gradients: 0.06367604074245038
l2 norm of weights: 2.2351347685957834
-----
Iteration Number: 8210
Loss: 10.949427587996123
l2 norm of gradients: 0.06366630831091437
l2 norm of weights: 2.2351651155537358
-----
Iteration Number: 8211
Loss: 10.949176031211165
l2 norm of gradients: 0.06365657936471815
l2 norm of weights: 2.235195472113929
-----
Iteration Number: 8212
Loss: 10.948924606290133
l2 norm of gradients: 0.06364685390212352
l2 norm of weights: 2.235225838267721
-----
Iteration Number: 8213
Loss: 10.948673313148682
l2 norm of gradients: 0.06363713192139339
l2 norm of weights: 2.2352562140064767
-----
Iteration Number: 8214
Loss: 10.948422151702513
l2 norm of gradients: 0.06362741342079174
l2 norm of weights: 2.2352865993215683
-----
Iteration Number: 8215
Loss: 10.948171121867384

```

```

12 norm of gradients: 0.0636176983985836
12 norm of weights: 2.2353169942043736
-----
Iteration Number: 8216
Loss: 10.947920223559118
12 norm of gradients: 0.06360798685303515
12 norm of weights: 2.2353473986462773
-----
Iteration Number: 8217
Loss: 10.947669456693593
12 norm of gradients: 0.06359827878241349
12 norm of weights: 2.23537781263867
-----
Iteration Number: 8218
Loss: 10.94741882118675
12 norm of gradients: 0.06358857418498692
12 norm of weights: 2.2354082361729506
-----
Iteration Number: 8219
Loss: 10.947168316954572
12 norm of gradients: 0.06357887305902478
12 norm of weights: 2.235438669240523
-----
Iteration Number: 8220
Loss: 10.946917943913117
12 norm of gradients: 0.06356917540279738
12 norm of weights: 2.235469111832797
-----
Iteration Number: 8221
Loss: 10.946667701978493
12 norm of gradients: 0.06355948121457629
12 norm of weights: 2.235499563941191
-----
Iteration Number: 8222
Loss: 10.946417591066862
12 norm of gradients: 0.06354979049263393
12 norm of weights: 2.2355300255571287
-----
Iteration Number: 8223
Loss: 10.946167611094452
12 norm of gradients: 0.0635401032352439
12 norm of weights: 2.2355604966720404
-----
Iteration Number: 8224
Loss: 10.945917761977544
12 norm of gradients: 0.06353041944068086
12 norm of weights: 2.2355909772773632
-----

```


Iteration Number: 8225
Loss: 10.945668043632473
l2 norm of gradients: 0.06352073910722045
l2 norm of weights: 2.23562146736454

Iteration Number: 8226
Loss: 10.94541845597564
l2 norm of gradients: 0.06351106223313951
l2 norm of weights: 2.235651966925021

Iteration Number: 8227
Loss: 10.945168998923492
l2 norm of gradients: 0.06350138881671581
l2 norm of weights: 2.235682475950263

Iteration Number: 8228
Loss: 10.944919672392546
l2 norm of gradients: 0.06349171885622822
l2 norm of weights: 2.2357129944317284

Iteration Number: 8229
Loss: 10.944670476299361
l2 norm of gradients: 0.06348205234995669
l2 norm of weights: 2.2357435223608864

Iteration Number: 8230
Loss: 10.944421410560574
l2 norm of gradients: 0.06347238929618217
l2 norm of weights: 2.2357740597292133

Iteration Number: 8231
Loss: 10.944172475092861
l2 norm of gradients: 0.06346272969318677
l2 norm of weights: 2.2358046065281916

Iteration Number: 8232
Loss: 10.943923669812962
l2 norm of gradients: 0.06345307353925353
l2 norm of weights: 2.23583516274931

Iteration Number: 8233
Loss: 10.943674994637668
l2 norm of gradients: 0.06344342083266663
l2 norm of weights: 2.2358657283840637

Iteration Number: 8234
Loss: 10.943426449483843
l2 norm of gradients: 0.06343377157171128

```

12 norm of weights: 2.2358963034239543
-----
Iteration Number: 8235
Loss: 10.943178034268394
12 norm of gradients: 0.06342412575467365
12 norm of weights: 2.2359268878604905
-----
Iteration Number: 8236
Loss: 10.942929748908282
12 norm of gradients: 0.06341448337984117
12 norm of weights: 2.2359574816851864
-----
Iteration Number: 8237
Loss: 10.942681593320543
12 norm of gradients: 0.06340484444550212
12 norm of weights: 2.235988084889564
-----
Iteration Number: 8238
Loss: 10.94243356742225
12 norm of gradients: 0.06339520894994591
12 norm of weights: 2.2360186974651506
-----
Iteration Number: 8239
Loss: 10.942185671130547
12 norm of gradients: 0.06338557689146297
12 norm of weights: 2.2360493194034796
-----
Iteration Number: 8240
Loss: 10.941937904362621
12 norm of gradients: 0.06337594826834485
12 norm of weights: 2.2360799506960922
-----
Iteration Number: 8241
Loss: 10.941690267035732
12 norm of gradients: 0.06336632307888404
12 norm of weights: 2.236110591334535
-----
Iteration Number: 8242
Loss: 10.941442759067186
12 norm of gradients: 0.06335670132137418
12 norm of weights: 2.2361412413103614
-----
Iteration Number: 8243
Loss: 10.941195380374353
12 norm of gradients: 0.0633470829941098
12 norm of weights: 2.2361719006151315
-----
Iteration Number: 8244

```

Loss: 10.940948130874645
l2 norm of gradients: 0.06333746809538664
l2 norm of weights: 2.2362025692404113

Iteration Number: 8245
Loss: 10.940701010485549
l2 norm of gradients: 0.06332785662350146
l2 norm of weights: 2.236233247177773

Iteration Number: 8246
Loss: 10.940454019124598
l2 norm of gradients: 0.0633182485767519
l2 norm of weights: 2.236263934418796

Iteration Number: 8247
Loss: 10.940207156709386
l2 norm of gradients: 0.06330864395343687
l2 norm of weights: 2.2362946309550655

Iteration Number: 8248
Loss: 10.939960423157558
l2 norm of gradients: 0.06329904275185613
l2 norm of weights: 2.236325336778174

Iteration Number: 8249
Loss: 10.939713818386817
l2 norm of gradients: 0.06328944497031055
l2 norm of weights: 2.236356051879719

Iteration Number: 8250
Loss: 10.939467342314932
l2 norm of gradients: 0.0632798506071021
l2 norm of weights: 2.236386776251306

Iteration Number: 8251
Loss: 10.939220994859717
l2 norm of gradients: 0.06327025966053368
l2 norm of weights: 2.2364175098845447

Iteration Number: 8252
Loss: 10.938974775939041
l2 norm of gradients: 0.06326067212890929
l2 norm of weights: 2.236448252771053

Iteration Number: 8253
Loss: 10.938728685470846
l2 norm of gradients: 0.06325108801053392
l2 norm of weights: 2.236479004902456

```

-----
Iteration Number: 8254
Loss: 10.938482723373104
l2 norm of gradients: 0.06324150730371364
l2 norm of weights: 2.2365097662703817
-----
Iteration Number: 8255
Loss: 10.93823688956387
l2 norm of gradients: 0.0632319300067555
l2 norm of weights: 2.236540536866468
-----
Iteration Number: 8256
Loss: 10.937991183961238
l2 norm of gradients: 0.0632223561179677
l2 norm of weights: 2.2365713166823578
-----
Iteration Number: 8257
Loss: 10.937745606483363
l2 norm of gradients: 0.06321278563565927
l2 norm of weights: 2.2366021057096996
-----
Iteration Number: 8258
Loss: 10.937500157048454
l2 norm of gradients: 0.06320321855814044
l2 norm of weights: 2.2366329039401496
-----
Iteration Number: 8259
Loss: 10.937254835574787
l2 norm of gradients: 0.06319365488372243
l2 norm of weights: 2.2366637113653693
-----
Iteration Number: 8260
Loss: 10.937009641980676
l2 norm of gradients: 0.06318409461071746
l2 norm of weights: 2.2366945279770274
-----
Iteration Number: 8261
Loss: 10.936764576184505
l2 norm of gradients: 0.06317453773743875
l2 norm of weights: 2.2367253537667984
-----
Iteration Number: 8262
Loss: 10.936519638104706
l2 norm of gradients: 0.06316498426220063
l2 norm of weights: 2.2367561887263636
-----
Iteration Number: 8263
Loss: 10.936274827659775

```

```

12 norm of gradients: 0.06315543418331838
12 norm of weights: 2.2367870328474098
-----
Iteration Number: 8264
Loss: 10.936030144768257
12 norm of gradients: 0.06314588749910836
12 norm of weights: 2.236817886121631
-----
Iteration Number: 8265
Loss: 10.935785589348754
12 norm of gradients: 0.06313634420788791
12 norm of weights: 2.236848748540727
-----
Iteration Number: 8266
Loss: 10.935541161319923
12 norm of gradients: 0.06312680430797542
12 norm of weights: 2.236879620096404
-----
Iteration Number: 8267
Loss: 10.935296860600486
12 norm of gradients: 0.06311726779769028
12 norm of weights: 2.236910500780375
-----
Iteration Number: 8268
Loss: 10.935052687109204
12 norm of gradients: 0.06310773467535294
12 norm of weights: 2.2369413905843585
-----
Iteration Number: 8269
Loss: 10.934808640764908
12 norm of gradients: 0.06309820493928482
12 norm of weights: 2.2369722895000796
-----
Iteration Number: 8270
Loss: 10.934564721486476
12 norm of gradients: 0.06308867858780834
12 norm of weights: 2.2370031975192703
-----
Iteration Number: 8271
Loss: 10.93432092919285
12 norm of gradients: 0.06307915561924707
12 norm of weights: 2.2370341146336683
-----
Iteration Number: 8272
Loss: 10.934077263803015
12 norm of gradients: 0.06306963603192545
12 norm of weights: 2.2370650408350183
-----

```

Iteration Number: 8273
Loss: 10.933833725236028
l2 norm of gradients: 0.063060119824169
l2 norm of weights: 2.2370959761150693

Iteration Number: 8274
Loss: 10.933590313410988
l2 norm of gradients: 0.06305060699430425
l2 norm of weights: 2.237126920465579

Iteration Number: 8275
Loss: 10.933347028247054
l2 norm of gradients: 0.06304109754065874
l2 norm of weights: 2.2371578738783096

Iteration Number: 8276
Loss: 10.933103869663443
l2 norm of gradients: 0.06303159146156108
l2 norm of weights: 2.237188836345031

Iteration Number: 8277
Loss: 10.932860837579419
l2 norm of gradients: 0.06302208875534075
l2 norm of weights: 2.237219807857519

Iteration Number: 8278
Loss: 10.932617931914313
l2 norm of gradients: 0.06301258942032836
l2 norm of weights: 2.2372507884075548

Iteration Number: 8279
Loss: 10.932375152587504
l2 norm of gradients: 0.06300309345485555
l2 norm of weights: 2.2372817779869263

Iteration Number: 8280
Loss: 10.932132499518428
l2 norm of gradients: 0.0629936008572549
l2 norm of weights: 2.2373127765874283

Iteration Number: 8281
Loss: 10.93188997262657
l2 norm of gradients: 0.06298411162586
l2 norm of weights: 2.2373437842008603

Iteration Number: 8282
Loss: 10.93164757183148
l2 norm of gradients: 0.06297462575900549

```

12 norm of weights: 2.237374800819031
-----
Iteration Number: 8283
Loss: 10.93140529705276
12 norm of gradients: 0.06296514325502695
12 norm of weights: 2.2374058264337515
-----
Iteration Number: 8284
Loss: 10.931163148210068
12 norm of gradients: 0.0629556641122611
12 norm of weights: 2.2374368610368425
-----
Iteration Number: 8285
Loss: 10.930921125223112
12 norm of gradients: 0.06294618832904553
12 norm of weights: 2.2374679046201282
-----
Iteration Number: 8286
Loss: 10.930679228011657
12 norm of gradients: 0.06293671590371885
12 norm of weights: 2.237498957175441
-----
Iteration Number: 8287
Loss: 10.930437456495527
12 norm of gradients: 0.06292724683462075
12 norm of weights: 2.2375300186946196
-----
Iteration Number: 8288
Loss: 10.930195810594595
12 norm of gradients: 0.06291778112009189
12 norm of weights: 2.2375610891695072
-----
Iteration Number: 8289
Loss: 10.929954290228796
12 norm of gradients: 0.0629083187584739
12 norm of weights: 2.237592168591954
-----
Iteration Number: 8290
Loss: 10.929712895318113
12 norm of gradients: 0.06289885974810946
12 norm of weights: 2.237623256953817
-----
Iteration Number: 8291
Loss: 10.929471625782588
12 norm of gradients: 0.06288940408734219
12 norm of weights: 2.2376543542469594
-----
Iteration Number: 8292

```

Loss: 10.929230481542312
l2 norm of gradients: 0.06287995177451677
l2 norm of weights: 2.23768546046325

Iteration Number: 8293
Loss: 10.928989462517444
l2 norm of gradients: 0.06287050280797886
l2 norm of weights: 2.2377165755945634

Iteration Number: 8294
Loss: 10.92874856862818
l2 norm of gradients: 0.06286105718607506
l2 norm of weights: 2.2377476996327816

Iteration Number: 8295
Loss: 10.928507799794781
l2 norm of gradients: 0.06285161490715309
l2 norm of weights: 2.237778832569792

Iteration Number: 8296
Loss: 10.928267155937569
l2 norm of gradients: 0.06284217596956154
l2 norm of weights: 2.2378099743974884

Iteration Number: 8297
Loss: 10.9280266369769
l2 norm of gradients: 0.06283274037165008
l2 norm of weights: 2.237841125107771

Iteration Number: 8298
Loss: 10.927786242833204
l2 norm of gradients: 0.06282330811176932
l2 norm of weights: 2.2378722846925454

Iteration Number: 8299
Loss: 10.927545973426959
l2 norm of gradients: 0.06281387918827089
l2 norm of weights: 2.2379034531437245

Iteration Number: 8300
Loss: 10.927305828678696
l2 norm of gradients: 0.0628044535995074
l2 norm of weights: 2.237934630453226

Iteration Number: 8301
Loss: 10.927065808509003
l2 norm of gradients: 0.0627950313438325
l2 norm of weights: 2.2379658166129754


```

-----
Iteration Number: 8302
Loss: 10.926825912838519
l2 norm of gradients: 0.06278561241960078
l2 norm of weights: 2.2379970116149033
-----
Iteration Number: 8303
Loss: 10.926586141587938
l2 norm of gradients: 0.06277619682516777
l2 norm of weights: 2.238028215450946
-----
Iteration Number: 8304
Loss: 10.926346494678013
l2 norm of gradients: 0.06276678455889012
l2 norm of weights: 2.238059428113047
-----
Iteration Number: 8305
Loss: 10.92610697202954
l2 norm of gradients: 0.06275737561912541
l2 norm of weights: 2.238090649593156
-----
Iteration Number: 8306
Loss: 10.925867573563389
l2 norm of gradients: 0.0627479700042321
l2 norm of weights: 2.238121879883227
-----
Iteration Number: 8307
Loss: 10.925628299200465
l2 norm of gradients: 0.06273856771256983
l2 norm of weights: 2.238153118975223
-----
Iteration Number: 8308
Loss: 10.925389148861736
l2 norm of gradients: 0.06272916874249908
l2 norm of weights: 2.238184366861111
-----
Iteration Number: 8309
Loss: 10.92515012246822
l2 norm of gradients: 0.06271977309238139
l2 norm of weights: 2.2382156235328647
-----
Iteration Number: 8310
Loss: 10.924911219940991
l2 norm of gradients: 0.06271038076057922
l2 norm of weights: 2.2382468889824643
-----
Iteration Number: 8311
Loss: 10.924672441201185

```

```

12 norm of gradients: 0.06270099174545608
12 norm of weights: 2.2382781632018958
-----
Iteration Number: 8312
Loss: 10.924433786169976
12 norm of gradients: 0.06269160604537644
12 norm of weights: 2.2383094461831505
-----
Iteration Number: 8313
Loss: 10.924195254768604
12 norm of gradients: 0.06268222365870571
12 norm of weights: 2.2383407379182274
-----
Iteration Number: 8314
Loss: 10.92395684691836
12 norm of gradients: 0.06267284458381035
12 norm of weights: 2.2383720383991306
-----
Iteration Number: 8315
Loss: 10.92371856254059
12 norm of gradients: 0.06266346881905774
12 norm of weights: 2.2384033476178704
-----
Iteration Number: 8316
Loss: 10.92348040155668
12 norm of gradients: 0.06265409636281627
12 norm of weights: 2.238434665566464
-----
Iteration Number: 8317
Loss: 10.923242363888098
12 norm of gradients: 0.06264472721345529
12 norm of weights: 2.238465992236933
-----
Iteration Number: 8318
Loss: 10.923004449456347
12 norm of gradients: 0.06263536136934514
12 norm of weights: 2.238497327621307
-----
Iteration Number: 8319
Loss: 10.922766658182978
12 norm of gradients: 0.06262599882885712
12 norm of weights: 2.23852867171162
-----
Iteration Number: 8320
Loss: 10.922528989989608
12 norm of gradients: 0.06261663959036359
12 norm of weights: 2.2385600244999133
-----

```

Iteration Number: 8321
Loss: 10.92229144479791
l2 norm of gradients: 0.06260728365223771
l2 norm of weights: 2.2385913859782343

Iteration Number: 8322
Loss: 10.922054022529592
l2 norm of gradients: 0.06259793101285381
l2 norm of weights: 2.2386227561386347

Iteration Number: 8323
Loss: 10.921816723106437
l2 norm of gradients: 0.06258858167058705
l2 norm of weights: 2.238654134973175

Iteration Number: 8324
Loss: 10.921579546450268
l2 norm of gradients: 0.06257923562381361
l2 norm of weights: 2.238685522473919

Iteration Number: 8325
Loss: 10.921342492482973
l2 norm of gradients: 0.06256989287091064
l2 norm of weights: 2.238716918632939

Iteration Number: 8326
Loss: 10.92110556112648
l2 norm of gradients: 0.06256055341025633
l2 norm of weights: 2.2387483234423113

Iteration Number: 8327
Loss: 10.920868752302777
l2 norm of gradients: 0.0625512172402297
l2 norm of weights: 2.23877973689412

Iteration Number: 8328
Loss: 10.920632065933908
l2 norm of gradients: 0.0625418843592108
l2 norm of weights: 2.238811158980454

Iteration Number: 8329
Loss: 10.920395501941966
l2 norm of gradients: 0.06253255476558076
l2 norm of weights: 2.2388425896934083

Iteration Number: 8330
Loss: 10.9201590602491
l2 norm of gradients: 0.06252322845772144

```

12 norm of weights: 2.238874029025085
-----
Iteration Number: 8331
Loss: 10.919922740777508
12 norm of gradients: 0.06251390543401592
12 norm of weights: 2.238905476967591
-----
Iteration Number: 8332
Loss: 10.919686543449451
12 norm of gradients: 0.06250458569284809
12 norm of weights: 2.2389369335130405
-----
Iteration Number: 8333
Loss: 10.91945046818723
12 norm of gradients: 0.06249526923260283
12 norm of weights: 2.238968398653552
-----
Iteration Number: 8334
Loss: 10.919214514913206
12 norm of gradients: 0.06248595605166601
12 norm of weights: 2.238999872381251
-----
Iteration Number: 8335
Loss: 10.918978683549797
12 norm of gradients: 0.06247664614842442
12 norm of weights: 2.2390313546882696
-----
Iteration Number: 8336
Loss: 10.918742974019473
12 norm of gradients: 0.062467339521265895
12 norm of weights: 2.239062845566745
-----
Iteration Number: 8337
Loss: 10.918507386244745
12 norm of gradients: 0.06245803616857913
12 norm of weights: 2.239094345008821
-----
Iteration Number: 8338
Loss: 10.918271920148188
12 norm of gradients: 0.0624487360887539
12 norm of weights: 2.2391258530066476
-----
Iteration Number: 8339
Loss: 10.918036575652433
12 norm of gradients: 0.06243943928018078
12 norm of weights: 2.2391573695523785
-----
Iteration Number: 8340

```

Loss: 10.917801352680156
l2 norm of gradients: 0.06243014574125143
l2 norm of weights: 2.2391888946381764

Iteration Number: 8341
Loss: 10.917566251154089
l2 norm of gradients: 0.062420855470358465
l2 norm of weights: 2.2392204282562087

Iteration Number: 8342
Loss: 10.917331270997016
l2 norm of gradients: 0.062411568465895344
l2 norm of weights: 2.239251970398649

Iteration Number: 8343
Loss: 10.917096412131777
l2 norm of gradients: 0.06240228472625664
l2 norm of weights: 2.239283521057676

Iteration Number: 8344
Loss: 10.91686167448126
l2 norm of gradients: 0.06239300424983777
l2 norm of weights: 2.239315080225476

Iteration Number: 8345
Loss: 10.916627057968407
l2 norm of gradients: 0.062383727035035134
l2 norm of weights: 2.23934664789424

Iteration Number: 8346
Loss: 10.916392562516215
l2 norm of gradients: 0.062374453080246084
l2 norm of weights: 2.2393782240561646

Iteration Number: 8347
Loss: 10.916158188047735
l2 norm of gradients: 0.06236518238386892
l2 norm of weights: 2.2394098087034546

Iteration Number: 8348
Loss: 10.915923934486067
l2 norm of gradients: 0.06235591494430296
l2 norm of weights: 2.239441401828318

Iteration Number: 8349
Loss: 10.915689801754358
l2 norm of gradients: 0.06234665075994835
l2 norm of weights: 2.23947300342297

```
-----  
Iteration Number: 8350  
Loss: 10.915455789775823  
l2 norm of gradients: 0.06233738982920631  
l2 norm of weights: 2.239504613479633  
-----  
Iteration Number: 8351  
Loss: 10.915221898473721  
l2 norm of gradients: 0.06232813215047891  
l2 norm of weights: 2.2395362319905328  
-----  
Iteration Number: 8352  
Loss: 10.914988127771355  
l2 norm of gradients: 0.06231887772216926  
l2 norm of weights: 2.239567858947903  
-----  
Iteration Number: 8353  
Loss: 10.914754477592094  
l2 norm of gradients: 0.06230962654268132  
l2 norm of weights: 2.2395994943439823  
-----  
Iteration Number: 8354  
Loss: 10.914520947859357  
l2 norm of gradients: 0.0623003786104201  
l2 norm of weights: 2.239631138171016  
-----  
Iteration Number: 8355  
Loss: 10.914287538496604  
l2 norm of gradients: 0.062291133923791445  
l2 norm of weights: 2.2396627904212543  
-----  
Iteration Number: 8356  
Loss: 10.914054249427366  
l2 norm of gradients: 0.062281892481202285  
l2 norm of weights: 2.2396944510869545  
-----  
Iteration Number: 8357  
Loss: 10.913821080575207  
l2 norm of gradients: 0.06227265428106035  
l2 norm of weights: 2.2397261201603786  
-----  
Iteration Number: 8358  
Loss: 10.913588031863759  
l2 norm of gradients: 0.062263419321774406  
l2 norm of weights: 2.239757797633796  
-----  
Iteration Number: 8359  
Loss: 10.913355103216698
```

```
l2 norm of gradients: 0.06225418760175418
l2 norm of weights: 2.2397894834994805
-----
Iteration Number: 8360
Loss: 10.913122294557748
l2 norm of gradients: 0.0622449591194102
l2 norm of weights: 2.2398211777497123
-----
Iteration Number: 8361
Loss: 10.912889605810697
l2 norm of gradients: 0.062235733873154106
l2 norm of weights: 2.2398528803767785
-----
Iteration Number: 8362
Loss: 10.912657036899382
l2 norm of gradients: 0.0622265118613984
l2 norm of weights: 2.2398845913729706
-----
Iteration Number: 8363
Loss: 10.912424587747683
l2 norm of gradients: 0.06221729308255654
l2 norm of weights: 2.239916310730587
-----
Iteration Number: 8364
Loss: 10.91219225827954
l2 norm of gradients: 0.06220807753504285
l2 norm of weights: 2.2399480384419315
-----
Iteration Number: 8365
Loss: 10.911960048418944
l2 norm of gradients: 0.06219886521727275
l2 norm of weights: 2.2399797744993135
-----
Iteration Number: 8366
Loss: 10.911727958089937
l2 norm of gradients: 0.06218965612766243
l2 norm of weights: 2.2400115188950496
-----
Iteration Number: 8367
Loss: 10.911495987216613
l2 norm of gradients: 0.06218045026462914
l2 norm of weights: 2.240043271621461
-----
Iteration Number: 8368
Loss: 10.911264135723115
l2 norm of gradients: 0.06217124762659099
l2 norm of weights: 2.2400750326708745
-----
```

Iteration Number: 8369
Loss: 10.911032403533646
l2 norm of gradients: 0.06216204821196706
l2 norm of weights: 2.2401068020356245

Iteration Number: 8370
Loss: 10.910800790572452
l2 norm of gradients: 0.06215285201917739
l2 norm of weights: 2.240138579708049

Iteration Number: 8371
Loss: 10.910569296763835
l2 norm of gradients: 0.062143659046642825
l2 norm of weights: 2.2401703656804943

Iteration Number: 8372
Loss: 10.910337922032154
l2 norm of gradients: 0.06213446929278536
l2 norm of weights: 2.24020215994531

Iteration Number: 8373
Loss: 10.910106666301806
l2 norm of gradients: 0.06212528275602773
l2 norm of weights: 2.240233962494854

Iteration Number: 8374
Loss: 10.909875529497253
l2 norm of gradients: 0.062116099434793695
l2 norm of weights: 2.2402657733214877

Iteration Number: 8375
Loss: 10.909644511542997
l2 norm of gradients: 0.062106919327507955
l2 norm of weights: 2.2402975924175808

Iteration Number: 8376
Loss: 10.909413612363602
l2 norm of gradients: 0.06209774243259603
l2 norm of weights: 2.2403294197755064

Iteration Number: 8377
Loss: 10.909182831883681
l2 norm of gradients: 0.06208856874848451
l2 norm of weights: 2.240361255387645

Iteration Number: 8378
Loss: 10.908952170027899
l2 norm of gradients: 0.062079398273600844


```
l2 norm of weights: 2.2403930992463823
-----
Iteration Number: 8379
Loss: 10.908721626720961
l2 norm of gradients: 0.0620702310063734
l2 norm of weights: 2.2404249513441106
-----
Iteration Number: 8380
Loss: 10.908491201887642
l2 norm of gradients: 0.06206106694523149
l2 norm of weights: 2.240456811673227
-----
Iteration Number: 8381
Loss: 10.908260895452754
l2 norm of gradients: 0.06205190608860534
l2 norm of weights: 2.2404886802261346
-----
Iteration Number: 8382
Loss: 10.908030707341167
l2 norm of gradients: 0.06204274843492616
l2 norm of weights: 2.240520556995243
-----
Iteration Number: 8383
Loss: 10.907800637477804
l2 norm of gradients: 0.06203359398262598
l2 norm of weights: 2.240552441972967
-----
Iteration Number: 8384
Loss: 10.907570685787634
l2 norm of gradients: 0.062024442730137866
l2 norm of weights: 2.240584335151728
-----
Iteration Number: 8385
Loss: 10.90734085219568
l2 norm of gradients: 0.06201529467589568
l2 norm of weights: 2.2406162365239513
-----
Iteration Number: 8386
Loss: 10.907111136627018
l2 norm of gradients: 0.062006149818334313
l2 norm of weights: 2.24064814608207
-----
Iteration Number: 8387
Loss: 10.906881539006768
l2 norm of gradients: 0.061997008155889564
l2 norm of weights: 2.240680063818522
-----
Iteration Number: 8388
```

Loss: 10.90665205926011
l2 norm of gradients: 0.06198786968699808
l2 norm of weights: 2.2407119897257517

Iteration Number: 8389
Loss: 10.906422697312273
l2 norm of gradients: 0.06197873441009753
l2 norm of weights: 2.2407439237962086

Iteration Number: 8390
Loss: 10.906193453088534
l2 norm of gradients: 0.06196960232362641
l2 norm of weights: 2.240775866022348

Iteration Number: 8391
Loss: 10.90596432651422
l2 norm of gradients: 0.06196047342602416
l2 norm of weights: 2.240807816396631

Iteration Number: 8392
Loss: 10.905735317514713
l2 norm of gradients: 0.06195134771573122
l2 norm of weights: 2.2408397749115245

Iteration Number: 8393
Loss: 10.905506426015442
l2 norm of gradients: 0.06194222519118882
l2 norm of weights: 2.2408717415595025

Iteration Number: 8394
Loss: 10.905277651941898
l2 norm of gradients: 0.06193310585083915
l2 norm of weights: 2.240903716333042

Iteration Number: 8395
Loss: 10.90504899521961
l2 norm of gradients: 0.06192398969312537
l2 norm of weights: 2.2409356992246283

Iteration Number: 8396
Loss: 10.904820455774159
l2 norm of gradients: 0.06191487671649152
l2 norm of weights: 2.240967690226751

Iteration Number: 8397
Loss: 10.904592033531184
l2 norm of gradients: 0.06190576691938247
l2 norm of weights: 2.2409996893319057

```

-----
Iteration Number: 8398
Loss: 10.90436372841637
l2 norm of gradients: 0.061896660300244155
l2 norm of weights: 2.2410316965325947
-----
Iteration Number: 8399
Loss: 10.904135540355455
l2 norm of gradients: 0.061887556857523285
l2 norm of weights: 2.241063711821325
-----
Iteration Number: 8400
Loss: 10.903907469274225
l2 norm of gradients: 0.061878456589667605
l2 norm of weights: 2.241095735190609
-----
Iteration Number: 8401
Loss: 10.903679515098517
l2 norm of gradients: 0.06186935949512568
l2 norm of weights: 2.241127766632966
-----
Iteration Number: 8402
Loss: 10.903451677754223
l2 norm of gradients: 0.061860265572347
l2 norm of weights: 2.2411598061409204
-----
Iteration Number: 8403
Loss: 10.90322395716728
l2 norm of gradients: 0.061851174819782
l2 norm of weights: 2.241191853707002
-----
Iteration Number: 8404
Loss: 10.90299635326368
l2 norm of gradients: 0.06184208723588198
l2 norm of weights: 2.241223909323747
-----
Iteration Number: 8405
Loss: 10.902768865969465
l2 norm of gradients: 0.06183300281909919
l2 norm of weights: 2.2412559729836974
-----
Iteration Number: 8406
Loss: 10.90254149521072
l2 norm of gradients: 0.061823921567886714
l2 norm of weights: 2.2412880446794
-----
Iteration Number: 8407
Loss: 10.902314240913592

```

```

12 norm of gradients: 0.06181484348069867
12 norm of weights: 2.241320124403408
-----
Iteration Number: 8408
Loss: 10.902087103004277
12 norm of gradients: 0.06180576855598993
12 norm of weights: 2.2413522121482803
-----
Iteration Number: 8409
Loss: 10.901860081409009
12 norm of gradients: 0.061796696792216374
12 norm of weights: 2.2413843079065807
-----
Iteration Number: 8410
Loss: 10.901633176054084
12 norm of gradients: 0.06178762818783476
12 norm of weights: 2.24141641167088
-----
Iteration Number: 8411
Loss: 10.901406386865844
12 norm of gradients: 0.0617785627413027
12 norm of weights: 2.241448523433754
-----
Iteration Number: 8412
Loss: 10.901179713770691
12 norm of gradients: 0.06176950045107883
12 norm of weights: 2.241480643187783
-----
Iteration Number: 8413
Loss: 10.90095315669506
12 norm of gradients: 0.06176044131562255
12 norm of weights: 2.2415127709255556
-----
Iteration Number: 8414
Loss: 10.900726715565444
12 norm of gradients: 0.061751385333394256
12 norm of weights: 2.2415449066396644
-----
Iteration Number: 8415
Loss: 10.900500390308393
12 norm of gradients: 0.061742332502855185
12 norm of weights: 2.2415770503227073
-----
Iteration Number: 8416
Loss: 10.900274180850502
12 norm of gradients: 0.0617332828224675
12 norm of weights: 2.2416092019672886
-----

```

Iteration Number: 8417
Loss: 10.90004808711841
l2 norm of gradients: 0.06172423629069424
l2 norm of weights: 2.2416413615660185

Iteration Number: 8418
Loss: 10.89982210903882
l2 norm of gradients: 0.061715192905999416
l2 norm of weights: 2.2416735291115124

Iteration Number: 8419
Loss: 10.899596246538469
l2 norm of gradients: 0.06170615266684782
l2 norm of weights: 2.2417057045963915

Iteration Number: 8420
Loss: 10.899370499544156
l2 norm of gradients: 0.061697115571705256
l2 norm of weights: 2.2417378880132826

Iteration Number: 8421
Loss: 10.899144867982722
l2 norm of gradients: 0.061688081619038326
l2 norm of weights: 2.2417700793548176

Iteration Number: 8422
Loss: 10.898919351781068
l2 norm of gradients: 0.06167905080731461
l2 norm of weights: 2.2418022786136356

Iteration Number: 8423
Loss: 10.898693950866134
l2 norm of gradients: 0.0616700231350025
l2 norm of weights: 2.241834485782379

Iteration Number: 8424
Loss: 10.898468665164916
l2 norm of gradients: 0.06166099860057132
l2 norm of weights: 2.2418667008536985

Iteration Number: 8425
Loss: 10.898243494604458
l2 norm of gradients: 0.06165197720249133
l2 norm of weights: 2.2418989238202487

Iteration Number: 8426
Loss: 10.898018439111853
l2 norm of gradients: 0.0616429589392336

```

12 norm of weights: 2.2419311546746896
-----
Iteration Number: 8427
Loss: 10.89779349861425
12 norm of gradients: 0.06163394380927016
12 norm of weights: 2.241963393409688
-----
Iteration Number: 8428
Loss: 10.897568673038839
12 norm of gradients: 0.06162493181107391
12 norm of weights: 2.2419956400179157
-----
Iteration Number: 8429
Loss: 10.897343962312863
12 norm of gradients: 0.06161592294311859
12 norm of weights: 2.24202789449205
-----
Iteration Number: 8430
Loss: 10.897119366363619
12 norm of gradients: 0.0616069172038789
12 norm of weights: 2.242060156824774
-----
Iteration Number: 8431
Loss: 10.896894885118446
12 norm of gradients: 0.06159791459183038
12 norm of weights: 2.2420924270087768
-----
Iteration Number: 8432
Loss: 10.896670518504738
12 norm of gradients: 0.06158891510544954
12 norm of weights: 2.242124705036752
-----
Iteration Number: 8433
Loss: 10.896446266449935
12 norm of gradients: 0.06157991874321363
12 norm of weights: 2.2421569909014
-----
Iteration Number: 8434
Loss: 10.896222128881533
12 norm of gradients: 0.061570925503600896
12 norm of weights: 2.242189284595426
-----
Iteration Number: 8435
Loss: 10.895998105727069
12 norm of gradients: 0.061561935385090465
12 norm of weights: 2.2422215861115418
-----
Iteration Number: 8436

```

Loss: 10.895774196914136
l2 norm of gradients: 0.06155294838616228
l2 norm of weights: 2.2422538954424627

Iteration Number: 8437
Loss: 10.895550402370375
l2 norm of gradients: 0.06154396450529729
l2 norm of weights: 2.242286212580912

Iteration Number: 8438
Loss: 10.895326722023471
l2 norm of gradients: 0.061534983740977156
l2 norm of weights: 2.242318537519617

Iteration Number: 8439
Loss: 10.895103155801168
l2 norm of gradients: 0.06152600609168462
l2 norm of weights: 2.242350870251311

Iteration Number: 8440
Loss: 10.89487970363125
l2 norm of gradients: 0.061517031555903094
l2 norm of weights: 2.2423832107687334

Iteration Number: 8441
Loss: 10.894656365441556
l2 norm of gradients: 0.06150806013211706
l2 norm of weights: 2.2424155590646286

Iteration Number: 8442
Loss: 10.894433141159974
l2 norm of gradients: 0.06149909181881174
l2 norm of weights: 2.2424479151317462

Iteration Number: 8443
Loss: 10.89421003071444
l2 norm of gradients: 0.06149012661447334
l2 norm of weights: 2.2424802789628426

Iteration Number: 8444
Loss: 10.893987034032934
l2 norm of gradients: 0.06148116451758889
l2 norm of weights: 2.2425126505506787

Iteration Number: 8445
Loss: 10.893764151043497
l2 norm of gradients: 0.06147220552664626
l2 norm of weights: 2.2425450298880203

```
-----  
Iteration Number: 8446  
Loss: 10.893541381674208  
l2 norm of gradients: 0.061463249640134304  
l2 norm of weights: 2.242577416967641  
-----  
Iteration Number: 8447  
Loss: 10.8933187258532  
l2 norm of gradients: 0.06145429685654268  
l2 norm of weights: 2.242609811782318  
-----  
Iteration Number: 8448  
Loss: 10.893096183508657  
l2 norm of gradients: 0.06144534717436191  
l2 norm of weights: 2.242642214324835  
-----  
Iteration Number: 8449  
Loss: 10.892873754568807  
l2 norm of gradients: 0.06143640059208341  
l2 norm of weights: 2.24267462458798  
-----  
Iteration Number: 8450  
Loss: 10.89265143896193  
l2 norm of gradients: 0.06142745710819953  
l2 norm of weights: 2.2427070425645486  
-----  
Iteration Number: 8451  
Loss: 10.892429236616353  
l2 norm of gradients: 0.06141851672120336  
l2 norm of weights: 2.2427394682473394  
-----  
Iteration Number: 8452  
Loss: 10.892207147460459  
l2 norm of gradients: 0.06140957942958899  
l2 norm of weights: 2.242771901629159  
-----  
Iteration Number: 8453  
Loss: 10.891985171422672  
l2 norm of gradients: 0.06140064523185133  
l2 norm of weights: 2.242804342702818  
-----  
Iteration Number: 8454  
Loss: 10.891763308431464  
l2 norm of gradients: 0.061391714126486176  
l2 norm of weights: 2.2428367914611327  
-----  
Iteration Number: 8455  
Loss: 10.891541558415359
```



```

12 norm of gradients: 0.06138278611199014
12 norm of weights: 2.242869247896925
-----
Iteration Number: 8456
Loss: 10.89131992130293
12 norm of gradients: 0.0613738611868608
12 norm of weights: 2.2429017120030226
-----
Iteration Number: 8457
Loss: 10.891098397022805
12 norm of gradients: 0.06136493934959654
12 norm of weights: 2.2429341837722587
-----
Iteration Number: 8458
Loss: 10.890876985503645
12 norm of gradients: 0.0613560205986966
12 norm of weights: 2.2429666631974716
-----
Iteration Number: 8459
Loss: 10.890655686674172
12 norm of gradients: 0.06134710493266114
12 norm of weights: 2.2429991502715048
-----
Iteration Number: 8460
Loss: 10.89043450046316
12 norm of gradients: 0.061338192349991154
12 norm of weights: 2.2430316449872083
-----
Iteration Number: 8461
Loss: 10.890213426799413
12 norm of gradients: 0.0613292828491885
12 norm of weights: 2.243064147337437
-----
Iteration Number: 8462
Loss: 10.889992465611805
12 norm of gradients: 0.06132037642875593
12 norm of weights: 2.243096657315051
-----
Iteration Number: 8463
Loss: 10.889771616829247
12 norm of gradients: 0.06131147308719703
12 norm of weights: 2.2431291749129163
-----
Iteration Number: 8464
Loss: 10.8895508803807
12 norm of gradients: 0.06130257282301627
12 norm of weights: 2.2431617001239044
-----

```

```

Iteration Number: 8465
Loss: 10.889330256195173
l2 norm of gradients: 0.061293675634718936
l2 norm of weights: 2.2431942329408923
-----
Iteration Number: 8466
Loss: 10.88910974420173
l2 norm of gradients: 0.061284781520811266
l2 norm of weights: 2.243226773356762
-----
Iteration Number: 8467
Loss: 10.888889344329469
l2 norm of gradients: 0.06127589047980029
l2 norm of weights: 2.2432593213644014
-----
Iteration Number: 8468
Loss: 10.888669056507554
l2 norm of gradients: 0.061267002510193895
l2 norm of weights: 2.2432918769567034
-----
Iteration Number: 8469
Loss: 10.888448880665184
l2 norm of gradients: 0.061258117610500956
l2 norm of weights: 2.2433244401265675
-----
Iteration Number: 8470
Loss: 10.88822881673161
l2 norm of gradients: 0.06124923577923097
l2 norm of weights: 2.2433570108668968
-----
Iteration Number: 8471
Loss: 10.888008864636134
l2 norm of gradients: 0.061240357014894505
l2 norm of weights: 2.2433895891706013
-----
Iteration Number: 8472
Loss: 10.88778902430811
l2 norm of gradients: 0.06123148131600293
l2 norm of weights: 2.2434221750305956
-----
Iteration Number: 8473
Loss: 10.88756929567693
l2 norm of gradients: 0.06122260868106843
l2 norm of weights: 2.243454768439801
-----
Iteration Number: 8474
Loss: 10.887349678672036
l2 norm of gradients: 0.061213739108604025

```

l2 norm of weights: 2.243487369391142

Iteration Number: 8475
Loss: 10.887130173222923
l2 norm of gradients: 0.061204872597123754
l2 norm of weights: 2.243519977877552

Iteration Number: 8476
Loss: 10.886910779259134
l2 norm of gradients: 0.061196009145142304
l2 norm of weights: 2.243552593891965

Iteration Number: 8477
Loss: 10.886691496710261
l2 norm of gradients: 0.061187148751175376
l2 norm of weights: 2.243585217427326

Iteration Number: 8478
Loss: 10.886472325505931
l2 norm of gradients: 0.061178291413739365
l2 norm of weights: 2.2436178484765796

Iteration Number: 8479
Loss: 10.886253265575842
l2 norm of gradients: 0.06116943713135172
l2 norm of weights: 2.2436504870326814

Iteration Number: 8480
Loss: 10.886034316849718
l2 norm of gradients: 0.0611605859025306
l2 norm of weights: 2.2436831330885876

Iteration Number: 8481
Loss: 10.885815479257346
l2 norm of gradients: 0.06115173772579505
l2 norm of weights: 2.2437157866372637

Iteration Number: 8482
Loss: 10.88559675272855
l2 norm of gradients: 0.06114289259966499
l2 norm of weights: 2.2437484476716776

Iteration Number: 8483
Loss: 10.885378137193207
l2 norm of gradients: 0.061134050522661146
l2 norm of weights: 2.2437811161848047

Iteration Number: 8484

Loss: 10.885159632581251
l2 norm of gradients: 0.06112521149330515
l2 norm of weights: 2.2438137921696244

Iteration Number: 8485
Loss: 10.884941238822641
l2 norm of gradients: 0.061116375510119454
l2 norm of weights: 2.2438464756191223

Iteration Number: 8486
Loss: 10.884722955847405
l2 norm of gradients: 0.061107542571627364
l2 norm of weights: 2.2438791665262885

Iteration Number: 8487
Loss: 10.884504783585612
l2 norm of gradients: 0.061098712676353
l2 norm of weights: 2.2439118648841205

Iteration Number: 8488
Loss: 10.884286721967372
l2 norm of gradients: 0.06108988582282141
l2 norm of weights: 2.2439445706856183

Iteration Number: 8489
Loss: 10.884068770922857
l2 norm of gradients: 0.06108106200955842
l2 norm of weights: 2.2439772839237895

Iteration Number: 8490
Loss: 10.88385093038227
l2 norm of gradients: 0.061072241235090724
l2 norm of weights: 2.244010004591646

Iteration Number: 8491
Loss: 10.883633200275876
l2 norm of gradients: 0.06106342349794587
l2 norm of weights: 2.244042732682206

Iteration Number: 8492
Loss: 10.883415580533974
l2 norm of gradients: 0.06105460879665223
l2 norm of weights: 2.2440754681884916

Iteration Number: 8493
Loss: 10.883198071086923
l2 norm of gradients: 0.06104579712973905
l2 norm of weights: 2.2441082111035313

```

-----
Iteration Number: 8494
Loss: 10.882980671865123
l2 norm of gradients: 0.06103698849573641
l2 norm of weights: 2.244140961420359
-----
Iteration Number: 8495
Loss: 10.882763382799022
l2 norm of gradients: 0.06102818289317516
l2 norm of weights: 2.244173719132014
-----
Iteration Number: 8496
Loss: 10.882546203819121
l2 norm of gradients: 0.06101938032058717
l2 norm of weights: 2.2442064842315395
-----
Iteration Number: 8497
Loss: 10.882329134855953
l2 norm of gradients: 0.06101058077650495
l2 norm of weights: 2.244239256711986
-----
Iteration Number: 8498
Loss: 10.882112175840117
l2 norm of gradients: 0.06100178425946198
l2 norm of weights: 2.2442720365664086
-----
Iteration Number: 8499
Loss: 10.881895326702251
l2 norm of gradients: 0.060992990767992555
l2 norm of weights: 2.2443048237878673
-----
Iteration Number: 8500
Loss: 10.881678587373038
l2 norm of gradients: 0.060984200300631786
l2 norm of weights: 2.244337618369428
-----
Iteration Number: 8501
Loss: 10.881461957783213
l2 norm of gradients: 0.06097541285591564
l2 norm of weights: 2.2443704203041612
-----
Iteration Number: 8502
Loss: 10.881245437863551
l2 norm of gradients: 0.06096662843238089
l2 norm of weights: 2.2444032295851435
-----
Iteration Number: 8503
Loss: 10.881029027544889

```

```

12 norm of gradients: 0.0609578470285652
12 norm of weights: 2.244436046205457
-----
Iteration Number: 8504
Loss: 10.880812726758087
12 norm of gradients: 0.06094906864300704
12 norm of weights: 2.244468870158188
-----
Iteration Number: 8505
Loss: 10.88059653543408
12 norm of gradients: 0.060940293274245695
12 norm of weights: 2.2445017014364295
-----
Iteration Number: 8506
Loss: 10.880380453503832
12 norm of gradients: 0.06093152092082135
12 norm of weights: 2.244534540033278
-----
Iteration Number: 8507
Loss: 10.880164480898356
12 norm of gradients: 0.06092275158127494
12 norm of weights: 2.244567385941837
-----
Iteration Number: 8508
Loss: 10.879948617548715
12 norm of gradients: 0.06091398525414835
12 norm of weights: 2.2446002391552144
-----
Iteration Number: 8509
Loss: 10.879732863386023
12 norm of gradients: 0.06090522193798418
12 norm of weights: 2.244633099666524
-----
Iteration Number: 8510
Loss: 10.879517218341432
12 norm of gradients: 0.06089646163132595
12 norm of weights: 2.244665967468884
-----
Iteration Number: 8511
Loss: 10.879301682346151
12 norm of gradients: 0.060887704332717944
12 norm of weights: 2.244698842555419
-----
Iteration Number: 8512
Loss: 10.879086255331426
12 norm of gradients: 0.06087895004070529
12 norm of weights: 2.244731724919258
-----

```

Iteration Number: 8513
Loss: 10.878870937228553
l2 norm of gradients: 0.06087019875383402
l2 norm of weights: 2.2447646145535356

Iteration Number: 8514
Loss: 10.878655727968882
l2 norm of gradients: 0.06086145047065092
l2 norm of weights: 2.2447975114513916

Iteration Number: 8515
Loss: 10.878440627483798
l2 norm of gradients: 0.06085270518970363
l2 norm of weights: 2.2448304156059713

Iteration Number: 8516
Loss: 10.878225635704743
l2 norm of gradients: 0.06084396290954059
l2 norm of weights: 2.2448633270104246

Iteration Number: 8517
Loss: 10.878010752563199
l2 norm of gradients: 0.060835223628711146
l2 norm of weights: 2.2448962456579076

Iteration Number: 8518
Loss: 10.877795977990699
l2 norm of gradients: 0.06082648734576535
l2 norm of weights: 2.2449291715415813

Iteration Number: 8519
Loss: 10.877581311918819
l2 norm of gradients: 0.060817754059254256
l2 norm of weights: 2.2449621046546113

Iteration Number: 8520
Loss: 10.877366754279187
l2 norm of gradients: 0.060809023767729564
l2 norm of weights: 2.2449950449901697

Iteration Number: 8521
Loss: 10.87715230500347
l2 norm of gradients: 0.06080029646974389
l2 norm of weights: 2.2450279925414325

Iteration Number: 8522
Loss: 10.876937964023393
l2 norm of gradients: 0.060791572163850706

```

12 norm of weights: 2.245060947301582
-----
Iteration Number: 8523
Loss: 10.876723731270708
12 norm of gradients: 0.06078285084860419
12 norm of weights: 2.2450939092638054
-----
Iteration Number: 8524
Loss: 10.876509606677233
12 norm of gradients: 0.06077413252255951
12 norm of weights: 2.2451268784212948
-----
Iteration Number: 8525
Loss: 10.87629559017483
12 norm of gradients: 0.0607654171842725
12 norm of weights: 2.2451598547672478
-----
Iteration Number: 8526
Loss: 10.876081681695391
12 norm of gradients: 0.06075670483229988
12 norm of weights: 2.245192838294867
-----
Iteration Number: 8527
Loss: 10.875867881170878
12 norm of gradients: 0.060747995465199246
12 norm of weights: 2.245225828997362
-----
Iteration Number: 8528
Loss: 10.875654188533279
12 norm of gradients: 0.06073928908152891
12 norm of weights: 2.2452588268679436
-----
Iteration Number: 8529
Loss: 10.87544060371464
12 norm of gradients: 0.06073058567984811
12 norm of weights: 2.245291831899832
-----
Iteration Number: 8530
Loss: 10.875227126647054
12 norm of gradients: 0.06072188525871683
12 norm of weights: 2.2453248440862508
-----
Iteration Number: 8531
Loss: 10.875013757262648
12 norm of gradients: 0.06071318781669585
12 norm of weights: 2.245357863420428
-----
Iteration Number: 8532

```



```

Loss: 10.874800495493611
l2 norm of gradients: 0.06070449335234691
l2 norm of weights: 2.245390889895598
-----
Iteration Number: 8533
Loss: 10.87458734127217
l2 norm of gradients: 0.06069580186423238
l2 norm of weights: 2.245423923505001
-----
Iteration Number: 8534
Loss: 10.874374294530593
l2 norm of gradients: 0.0606871133509156
l2 norm of weights: 2.2454569642418805
-----
Iteration Number: 8535
Loss: 10.874161355201208
l2 norm of gradients: 0.06067842781096066
l2 norm of weights: 2.245490012099487
-----
Iteration Number: 8536
Loss: 10.873948523216374
l2 norm of gradients: 0.06066974524293243
l2 norm of weights: 2.2455230670710744
-----
Iteration Number: 8537
Loss: 10.87373579850851
l2 norm of gradients: 0.060661065645396695
l2 norm of weights: 2.2455561291499033
-----
Iteration Number: 8538
Loss: 10.873523181010071
l2 norm of gradients: 0.06065238901691995
l2 norm of weights: 2.245589198329239
-----
Iteration Number: 8539
Loss: 10.873310670653563
l2 norm of gradients: 0.06064371535606961
l2 norm of weights: 2.2456222746023515
-----
Iteration Number: 8540
Loss: 10.873098267371535
l2 norm of gradients: 0.06063504466141379
l2 norm of weights: 2.2456553579625176
-----
Iteration Number: 8541
Loss: 10.872885971096585
l2 norm of gradients: 0.060626376931521515
l2 norm of weights: 2.245688448403017

```

```

-----
Iteration Number: 8542
Loss: 10.872673781761355
l2 norm of gradients: 0.06061771216496258
l2 norm of weights: 2.2457215459171356
-----
Iteration Number: 8543
Loss: 10.872461699298531
l2 norm of gradients: 0.060609050360307566
l2 norm of weights: 2.245754650498165
-----
Iteration Number: 8544
Loss: 10.87224972364085
l2 norm of gradients: 0.06060039151612789
l2 norm of weights: 2.245787762139401
-----
Iteration Number: 8545
Loss: 10.872037854721095
l2 norm of gradients: 0.06059173563099583
l2 norm of weights: 2.245820880834145
-----
Iteration Number: 8546
Loss: 10.871826092472086
l2 norm of gradients: 0.060583082703484385
l2 norm of weights: 2.245854006575704
-----
Iteration Number: 8547
Loss: 10.871614436826693
l2 norm of gradients: 0.06057443273216741
l2 norm of weights: 2.24588713935739
-----
Iteration Number: 8548
Loss: 10.871402887717839
l2 norm of gradients: 0.06056578571561957
l2 norm of weights: 2.245920279172519
-----
Iteration Number: 8549
Loss: 10.871191445078486
l2 norm of gradients: 0.06055714165241631
l2 norm of weights: 2.2459534260144127
-----
Iteration Number: 8550
Loss: 10.870980108841643
l2 norm of gradients: 0.06054850054113393
l2 norm of weights: 2.2459865798763996
-----
Iteration Number: 8551
Loss: 10.870768878940359

```

```

12 norm of gradients: 0.0605398623803495
12 norm of weights: 2.246019740751811
-----
Iteration Number: 8552
Loss: 10.870557755307738
12 norm of gradients: 0.06053122716864087
12 norm of weights: 2.2460529086339847
-----
Iteration Number: 8553
Loss: 10.870346737876929
12 norm of gradients: 0.06052259490458678
12 norm of weights: 2.2460860835162624
-----
Iteration Number: 8554
Loss: 10.870135826581118
12 norm of gradients: 0.06051396558676672
12 norm of weights: 2.2461192653919926
-----
Iteration Number: 8555
Loss: 10.869925021353541
12 norm of gradients: 0.06050533921376094
12 norm of weights: 2.2461524542545277
-----
Iteration Number: 8556
Loss: 10.869714322127482
12 norm of gradients: 0.06049671578415059
12 norm of weights: 2.2461856500972255
-----
Iteration Number: 8557
Loss: 10.869503728836271
12 norm of gradients: 0.060488095296517594
12 norm of weights: 2.2462188529134495
-----
Iteration Number: 8558
Loss: 10.869293241413278
12 norm of gradients: 0.06047947774944462
12 norm of weights: 2.246252062696567
-----
Iteration Number: 8559
Loss: 10.869082859791918
12 norm of gradients: 0.06047086314151514
12 norm of weights: 2.2462852794399515
-----
Iteration Number: 8560
Loss: 10.86887258390566
12 norm of gradients: 0.06046225147131355
12 norm of weights: 2.2463185031369814
-----

```

Iteration Number: 8561
Loss: 10.868662413688012
l2 norm of gradients: 0.060453642737424915
l2 norm of weights: 2.2463517337810397

Iteration Number: 8562
Loss: 10.868452349072527
l2 norm of gradients: 0.06044503693843512
l2 norm of weights: 2.246384971365515

Iteration Number: 8563
Loss: 10.868242389992805
l2 norm of gradients: 0.06043643407293094
l2 norm of weights: 2.246418215883801

Iteration Number: 8564
Loss: 10.868032536382493
l2 norm of gradients: 0.060427834139499856
l2 norm of weights: 2.246451467329296

Iteration Number: 8565
Loss: 10.867822788175273
l2 norm of gradients: 0.06041923713673015
l2 norm of weights: 2.2464847256954044

Iteration Number: 8566
Loss: 10.867613145304889
l2 norm of gradients: 0.060410643063210936
l2 norm of weights: 2.2465179909755344

Iteration Number: 8567
Loss: 10.86740360770512
l2 norm of gradients: 0.060402051917532104
l2 norm of weights: 2.2465512631630995

Iteration Number: 8568
Loss: 10.867194175309786
l2 norm of gradients: 0.06039346369828437
l2 norm of weights: 2.246584542251519

Iteration Number: 8569
Loss: 10.866984848052763
l2 norm of gradients: 0.06038487840405923
l2 norm of weights: 2.2466178282342173

Iteration Number: 8570
Loss: 10.86677562586796
l2 norm of gradients: 0.06037629603344893

```
12 norm of weights: 2.2466511211046227
-----
Iteration Number: 8571
Loss: 10.866566508689347
12 norm of gradients: 0.0603677165850466
12 norm of weights: 2.2466844208561696
-----
Iteration Number: 8572
Loss: 10.866357496450922
12 norm of gradients: 0.0603591400574461
12 norm of weights: 2.2467177274822974
-----
Iteration Number: 8573
Loss: 10.866148589086738
12 norm of gradients: 0.06035056644924206
12 norm of weights: 2.2467510409764495
-----
Iteration Number: 8574
Loss: 10.865939786530886
12 norm of gradients: 0.06034199575902996
12 norm of weights: 2.246784361332076
-----
Iteration Number: 8575
Loss: 10.865731088717515
12 norm of gradients: 0.06033342798540609
12 norm of weights: 2.246817688542631
-----
Iteration Number: 8576
Loss: 10.865522495580802
12 norm of gradients: 0.06032486312696742
12 norm of weights: 2.2468510226015734
-----
Iteration Number: 8577
Loss: 10.86531400705498
12 norm of gradients: 0.060316301182311835
12 norm of weights: 2.246884363502368
-----
Iteration Number: 8578
Loss: 10.86510562307432
12 norm of gradients: 0.06030774215003791
12 norm of weights: 2.2469177112384835
-----
Iteration Number: 8579
Loss: 10.86489734357315
12 norm of gradients: 0.06029918602874513
12 norm of weights: 2.246951065803395
-----
Iteration Number: 8580
```

Loss: 10.864689168485825
l2 norm of gradients: 0.060290632817033625
l2 norm of weights: 2.2469844271905814

Iteration Number: 8581
Loss: 10.864481097746761
l2 norm of gradients: 0.060282082513504406
l2 norm of weights: 2.2470177953935275

Iteration Number: 8582
Loss: 10.864273131290402
l2 norm of gradients: 0.060273535116759244
l2 norm of weights: 2.247051170405723

Iteration Number: 8583
Loss: 10.864065269051256
l2 norm of gradients: 0.060264990625400715
l2 norm of weights: 2.2470845522206617

Iteration Number: 8584
Loss: 10.86385751096386
l2 norm of gradients: 0.06025644903803218
l2 norm of weights: 2.247117940831844

Iteration Number: 8585
Loss: 10.863649856962805
l2 norm of gradients: 0.06024791035325774
l2 norm of weights: 2.247151336232773

Iteration Number: 8586
Loss: 10.863442306982721
l2 norm of gradients: 0.060239374569682314
l2 norm of weights: 2.247184738416959

Iteration Number: 8587
Loss: 10.863234860958285
l2 norm of gradients: 0.06023084168591164
l2 norm of weights: 2.2472181473779167

Iteration Number: 8588
Loss: 10.863027518824216
l2 norm of gradients: 0.06022231170055217
l2 norm of weights: 2.2472515631091654

Iteration Number: 8589
Loss: 10.862820280515272
l2 norm of gradients: 0.060213784612211184
l2 norm of weights: 2.247284985604229

```

-----
Iteration Number: 8590
Loss: 10.86261314596628
l2 norm of gradients: 0.06020526041949672
l2 norm of weights: 2.2473184148566374
-----
Iteration Number: 8591
Loss: 10.86240611511208
l2 norm of gradients: 0.060196739121017646
l2 norm of weights: 2.2473518508599253
-----
Iteration Number: 8592
Loss: 10.862199187887578
l2 norm of gradients: 0.06018822071538357
l2 norm of weights: 2.2473852936076315
-----
Iteration Number: 8593
Loss: 10.861992364227714
l2 norm of gradients: 0.060179705201204886
l2 norm of weights: 2.247418743093301
-----
Iteration Number: 8594
Loss: 10.861785644067469
l2 norm of gradients: 0.060171192577092754
l2 norm of weights: 2.2474521993104823
-----
Iteration Number: 8595
Loss: 10.861579027341884
l2 norm of gradients: 0.06016268284165914
l2 norm of weights: 2.2474856622527306
-----
Iteration Number: 8596
Loss: 10.861372513986028
l2 norm of gradients: 0.060154175993516794
l2 norm of weights: 2.247519131913605
-----
Iteration Number: 8597
Loss: 10.861166103935025
l2 norm of gradients: 0.06014567203127919
l2 norm of weights: 2.247552608286669
-----
Iteration Number: 8598
Loss: 10.860959797124037
l2 norm of gradients: 0.060137170953560685
l2 norm of weights: 2.2475860913654926
-----
Iteration Number: 8599
Loss: 10.860753593488269

```

l2 norm of gradients: 0.060128672758976306
l2 norm of weights: 2.24761958114365

Iteration Number: 8600
Loss: 10.860547492962978
l2 norm of gradients: 0.06012017744614192
l2 norm of weights: 2.247653077614719

Iteration Number: 8601
Loss: 10.860341495483459
l2 norm of gradients: 0.060111685013674136
l2 norm of weights: 2.2476865807722857

Iteration Number: 8602
Loss: 10.860135600985052
l2 norm of gradients: 0.06010319546019034
l2 norm of weights: 2.247720090609937

Iteration Number: 8603
Loss: 10.859929809403136
l2 norm of gradients: 0.06009470878430872
l2 norm of weights: 2.2477536071212683

Iteration Number: 8604
Loss: 10.859724120673144
l2 norm of gradients: 0.06008622498464823
l2 norm of weights: 2.2477871302998778

Iteration Number: 8605
Loss: 10.85951853473055
l2 norm of gradients: 0.06007774405982856
l2 norm of weights: 2.2478206601393693

Iteration Number: 8606
Loss: 10.859313051510867
l2 norm of gradients: 0.060069266008470235
l2 norm of weights: 2.2478541966333516

Iteration Number: 8607
Loss: 10.859107670949655
l2 norm of gradients: 0.060060790829194535
l2 norm of weights: 2.247887739775438

Iteration Number: 8608
Loss: 10.858902392982516
l2 norm of gradients: 0.060052318520623475
l2 norm of weights: 2.247921289559248

Iteration Number: 8609
Loss: 10.858697217545103
l2 norm of gradients: 0.06004384908137986
l2 norm of weights: 2.2479548459784042

Iteration Number: 8610
Loss: 10.858492144573106
l2 norm of gradients: 0.06003538251008727
l2 norm of weights: 2.247988409026535

Iteration Number: 8611
Loss: 10.858287174002253
l2 norm of gradients: 0.060026918805370096
l2 norm of weights: 2.248021978697274

Iteration Number: 8612
Loss: 10.858082305768331
l2 norm of gradients: 0.06001845796585335
l2 norm of weights: 2.2480555549842594

Iteration Number: 8613
Loss: 10.857877539807161
l2 norm of gradients: 0.06000999999016304
l2 norm of weights: 2.248089137881134

Iteration Number: 8614
Loss: 10.857672876054608
l2 norm of gradients: 0.06000154487692579
l2 norm of weights: 2.2481227273815456

Iteration Number: 8615
Loss: 10.85746831444658
l2 norm of gradients: 0.059993092624769
l2 norm of weights: 2.2481563234791477

Iteration Number: 8616
Loss: 10.857263854919035
l2 norm of gradients: 0.05998464323232086
l2 norm of weights: 2.2481899261675977

Iteration Number: 8617
Loss: 10.857059497407967
l2 norm of gradients: 0.05997619669821041
l2 norm of weights: 2.2482235354405584

Iteration Number: 8618
Loss: 10.856855241849418
l2 norm of gradients: 0.059967753021067244

```
l2 norm of weights: 2.248257151291697
-----
Iteration Number: 8619
Loss: 10.856651088179472
l2 norm of gradients: 0.05995931219952194
l2 norm of weights: 2.2482907737146864
-----
Iteration Number: 8620
Loss: 10.85644703633426
l2 norm of gradients: 0.059950874232205756
l2 norm of weights: 2.2483244027032034
-----
Iteration Number: 8621
Loss: 10.856243086249945
l2 norm of gradients: 0.05994243911775068
l2 norm of weights: 2.248358038250931
-----
Iteration Number: 8622
Loss: 10.856039237862749
l2 norm of gradients: 0.0599340068547895
l2 norm of weights: 2.248391680351555
-----
Iteration Number: 8623
Loss: 10.855835491108927
l2 norm of gradients: 0.05992557744195578
l2 norm of weights: 2.2484253289987683
-----
Iteration Number: 8624
Loss: 10.855631845924782
l2 norm of gradients: 0.05991715087788383
l2 norm of weights: 2.248458984186267
-----
Iteration Number: 8625
Loss: 10.855428302246656
l2 norm of gradients: 0.05990872716120871
l2 norm of weights: 2.248492645907753
-----
Iteration Number: 8626
Loss: 10.85522486001094
l2 norm of gradients: 0.05990030629056628
l2 norm of weights: 2.248526314156933
-----
Iteration Number: 8627
Loss: 10.855021519154063
l2 norm of gradients: 0.05989188826459308
l2 norm of weights: 2.248559988927518
-----
Iteration Number: 8628
```

```

Loss: 10.8548182796125
l2 norm of gradients: 0.059883473081926535
l2 norm of weights: 2.248593670213224
-----
Iteration Number: 8629
Loss: 10.854615141322771
l2 norm of gradients: 0.059875060741204696
l2 norm of weights: 2.2486273580077727
-----
Iteration Number: 8630
Loss: 10.854412104221435
l2 norm of gradients: 0.059866651241066486
l2 norm of weights: 2.248661052304889
-----
Iteration Number: 8631
Loss: 10.854209168245097
l2 norm of gradients: 0.05985824458015147
l2 norm of weights: 2.2486947530983046
-----
Iteration Number: 8632
Loss: 10.854006333330403
l2 norm of gradients: 0.05984984075710013
l2 norm of weights: 2.248728460381754
-----
Iteration Number: 8633
Loss: 10.853803599414046
l2 norm of gradients: 0.059841439770553534
l2 norm of weights: 2.2487621741489785
-----
Iteration Number: 8634
Loss: 10.853600966432754
l2 norm of gradients: 0.059833041619153635
l2 norm of weights: 2.2487958943937225
-----
Iteration Number: 8635
Loss: 10.85339843432331
l2 norm of gradients: 0.059824646301543084
l2 norm of weights: 2.248829621109736
-----
Iteration Number: 8636
Loss: 10.853196003022532
l2 norm of gradients: 0.0598162538163653
l2 norm of weights: 2.2488633542907746
-----
Iteration Number: 8637
Loss: 10.85299367246728
l2 norm of gradients: 0.059807864162264454
l2 norm of weights: 2.248897093930597

```

```

-----
Iteration Number: 8638
Loss: 10.85279144259446
l2 norm of gradients: 0.05979947733788545
l2 norm of weights: 2.2489308400229686
-----
Iteration Number: 8639
Loss: 10.85258931334102
l2 norm of gradients: 0.059791093341874
l2 norm of weights: 2.2489645925616575
-----
Iteration Number: 8640
Loss: 10.852387284643957
l2 norm of gradients: 0.05978271217287654
l2 norm of weights: 2.248998351540439
-----
Iteration Number: 8641
Loss: 10.852185356440298
l2 norm of gradients: 0.059774333829540234
l2 norm of weights: 2.2490321169530905
-----
Iteration Number: 8642
Loss: 10.85198352866712
l2 norm of gradients: 0.05976595831051302
l2 norm of weights: 2.2490658887933965
-----
Iteration Number: 8643
Loss: 10.851781801261547
l2 norm of gradients: 0.05975758561444361
l2 norm of weights: 2.2490996670551455
-----
Iteration Number: 8644
Loss: 10.851580174160743
l2 norm of gradients: 0.05974921573998143
l2 norm of weights: 2.2491334517321304
-----
Iteration Number: 8645
Loss: 10.851378647301905
l2 norm of gradients: 0.059740848685776705
l2 norm of weights: 2.24916724281815
-----
Iteration Number: 8646
Loss: 10.85117722062229
l2 norm of gradients: 0.05973248445048038
l2 norm of weights: 2.249201040307006
-----
Iteration Number: 8647
Loss: 10.85097589405918

```

```

12 norm of gradients: 0.05972412303274406
12 norm of weights: 2.2492348441925065
-----
Iteration Number: 8648
Loss: 10.850774667549917
12 norm of gradients: 0.05971576443122031
12 norm of weights: 2.2492686544684637
-----
Iteration Number: 8649
Loss: 10.850573541031872
12 norm of gradients: 0.05970740864456227
12 norm of weights: 2.2493024711286944
-----
Iteration Number: 8650
Loss: 10.850372514442467
12 norm of gradients: 0.059699055671423834
12 norm of weights: 2.2493362941670214
-----
Iteration Number: 8651
Loss: 10.85017158771916
12 norm of gradients: 0.059690705510459734
12 norm of weights: 2.2493701235772705
-----
Iteration Number: 8652
Loss: 10.849970760799453
12 norm of gradients: 0.05968235816032541
12 norm of weights: 2.2494039593532733
-----
Iteration Number: 8653
Loss: 10.849770033620894
12 norm of gradients: 0.05967401361967703
12 norm of weights: 2.2494378014888663
-----
Iteration Number: 8654
Loss: 10.849569406121073
12 norm of gradients: 0.059665671887171555
12 norm of weights: 2.24947164997789
-----
Iteration Number: 8655
Loss: 10.849368878237623
12 norm of gradients: 0.059657332961466567
12 norm of weights: 2.2495055048141896
-----
Iteration Number: 8656
Loss: 10.84916844990821
12 norm of gradients: 0.05964899684122056
12 norm of weights: 2.249539365991617
-----

```

```

Iteration Number: 8657
Loss: 10.848968121070557
l2 norm of gradients: 0.05964066352509265
l2 norm of weights: 2.2495732335040266
-----
Iteration Number: 8658
Loss: 10.84876789166242
l2 norm of gradients: 0.059632333011742776
l2 norm of weights: 2.2496071073452777
-----
Iteration Number: 8659
Loss: 10.848567761621602
l2 norm of gradients: 0.05962400529983152
l2 norm of weights: 2.2496409875092356
-----
Iteration Number: 8660
Loss: 10.84836773088594
l2 norm of gradients: 0.05961568038802038
l2 norm of weights: 2.2496748739897696
-----
Iteration Number: 8661
Loss: 10.848167799393323
l2 norm of gradients: 0.05960735827497136
l2 norm of weights: 2.2497087667807536
-----
Iteration Number: 8662
Loss: 10.847967967081678
l2 norm of gradients: 0.059599038959347386
l2 norm of weights: 2.249742665876066
-----
Iteration Number: 8663
Loss: 10.847768233888974
l2 norm of gradients: 0.0595907224398121
l2 norm of weights: 2.2497765712695914
-----
Iteration Number: 8664
Loss: 10.847568599753226
l2 norm of gradients: 0.059582408715029817
l2 norm of weights: 2.2498104829552177
-----
Iteration Number: 8665
Loss: 10.847369064612481
l2 norm of gradients: 0.05957409778366563
l2 norm of weights: 2.249844400926837
-----
Iteration Number: 8666
Loss: 10.84716962840484
l2 norm of gradients: 0.05956578964438539

```

```

12 norm of weights: 2.249878325178348
-----
Iteration Number: 8667
Loss: 10.84697029106844
12 norm of gradients: 0.05955748429585565
12 norm of weights: 2.249912255703653
-----
Iteration Number: 8668
Loss: 10.846771052541465
12 norm of gradients: 0.05954918173674373
12 norm of weights: 2.2499461924966586
-----
Iteration Number: 8669
Loss: 10.84657191276213
12 norm of gradients: 0.05954088196571764
12 norm of weights: 2.249980135551277
-----
Iteration Number: 8670
Loss: 10.846372871668704
12 norm of gradients: 0.0595325849814462
12 norm of weights: 2.2500140848614243
-----
Iteration Number: 8671
Loss: 10.846173929199493
12 norm of gradients: 0.05952429078259889
12 norm of weights: 2.250048040421022
-----
Iteration Number: 8672
Loss: 10.845975085292842
12 norm of gradients: 0.05951599936784601
12 norm of weights: 2.250082002223996
-----
Iteration Number: 8673
Loss: 10.845776339887147
12 norm of gradients: 0.059507710735858484
12 norm of weights: 2.2501159702642766
-----
Iteration Number: 8674
Loss: 10.845577692920834
12 norm of gradients: 0.0594994248853081
12 norm of weights: 2.2501499445357993
-----
Iteration Number: 8675
Loss: 10.84537914433238
12 norm of gradients: 0.05949114181486728
12 norm of weights: 2.250183925032504
-----
Iteration Number: 8676

```

Loss: 10.845180694060302
l2 norm of gradients: 0.05948286152320928
l2 norm of weights: 2.250217911748335

Iteration Number: 8677
Loss: 10.844982342043151
l2 norm of gradients: 0.05947458400900791
l2 norm of weights: 2.2502519046772425

Iteration Number: 8678
Loss: 10.844784088219532
l2 norm of gradients: 0.059466309270937906
l2 norm of weights: 2.2502859038131793

Iteration Number: 8679
Loss: 10.844585932528085
l2 norm of gradients: 0.059458037307674634
l2 norm of weights: 2.2503199091501043

Iteration Number: 8680
Loss: 10.844387874907493
l2 norm of gradients: 0.05944976811789424
l2 norm of weights: 2.2503539206819814

Iteration Number: 8681
Loss: 10.844189915296479
l2 norm of gradients: 0.05944150170027353
l2 norm of weights: 2.2503879384027776

Iteration Number: 8682
Loss: 10.843992053633812
l2 norm of gradients: 0.059433238053490114
l2 norm of weights: 2.2504219623064667

Iteration Number: 8683
Loss: 10.843794289858293
l2 norm of gradients: 0.0594249771762223
l2 norm of weights: 2.2504559923870247

Iteration Number: 8684
Loss: 10.843596623908779
l2 norm of gradients: 0.059416719067149106
l2 norm of weights: 2.250490028638434

Iteration Number: 8685
Loss: 10.843399055724154
l2 norm of gradients: 0.059408463724950365
l2 norm of weights: 2.2505240710546817


```

-----
Iteration Number: 8686
Loss: 10.843201585243358
l2 norm of gradients: 0.05940021114830649
l2 norm of weights: 2.250558119629758
-----
Iteration Number: 8687
Loss: 10.843004212405358
l2 norm of gradients: 0.059391961335898774
l2 norm of weights: 2.250592174357659
-----
Iteration Number: 8688
Loss: 10.842806937149172
l2 norm of gradients: 0.05938371428640913
l2 norm of weights: 2.2506262352323856
-----
Iteration Number: 8689
Loss: 10.842609759413858
l2 norm of gradients: 0.05937546999852027
l2 norm of weights: 2.250660302247943
-----
Iteration Number: 8690
Loss: 10.842412679138512
l2 norm of gradients: 0.05936722847091559
l2 norm of weights: 2.2506943753983397
-----
Iteration Number: 8691
Loss: 10.84221569626228
l2 norm of gradients: 0.059358989702279194
l2 norm of weights: 2.250728454677591
-----
Iteration Number: 8692
Loss: 10.842018810724332
l2 norm of gradients: 0.059350753691295964
l2 norm of weights: 2.2507625400797164
-----
Iteration Number: 8693
Loss: 10.841822022463898
l2 norm of gradients: 0.05934252043665148
l2 norm of weights: 2.250796631598738
-----
Iteration Number: 8694
Loss: 10.841625331420236
l2 norm of gradients: 0.05933428993703205
l2 norm of weights: 2.250830729228685
-----
Iteration Number: 8695
Loss: 10.84142873753266

```

```

12 norm of gradients: 0.05932606219112469
12 norm of weights: 2.2508648329635905
-----
Iteration Number: 8696
Loss: 10.84123224074051
12 norm of gradients: 0.05931783719761716
12 norm of weights: 2.2508989427974906
-----
Iteration Number: 8697
Loss: 10.841035840983173
12 norm of gradients: 0.05930961495519791
12 norm of weights: 2.2509330587244283
-----
Iteration Number: 8698
Loss: 10.840839538200077
12 norm of gradients: 0.059301395462556174
12 norm of weights: 2.2509671807384497
-----
Iteration Number: 8699
Loss: 10.840643332330696
12 norm of gradients: 0.059293178718381856
12 norm of weights: 2.2510013088336067
-----
Iteration Number: 8700
Loss: 10.840447223314538
12 norm of gradients: 0.05928496472136559
12 norm of weights: 2.2510354430039548
-----
Iteration Number: 8701
Loss: 10.840251211091157
12 norm of gradients: 0.05927675347019873
12 norm of weights: 2.2510695832435537
-----
Iteration Number: 8702
Loss: 10.840055295600143
12 norm of gradients: 0.05926854496357335
12 norm of weights: 2.251103729546469
-----
Iteration Number: 8703
Loss: 10.839859476781138
12 norm of gradients: 0.05926033920018228
12 norm of weights: 2.2511378819067707
-----
Iteration Number: 8704
Loss: 10.839663754573804
12 norm of gradients: 0.05925213617871899
12 norm of weights: 2.251172040318532
-----

```

Iteration Number: 8705
Loss: 10.839468128917865
l2 norm of gradients: 0.05924393589787774
l2 norm of weights: 2.251206204775832

Iteration Number: 8706
Loss: 10.839272599753079
l2 norm of gradients: 0.05923573835635349
l2 norm of weights: 2.2512403752727543

Iteration Number: 8707
Loss: 10.839077167019244
l2 norm of gradients: 0.05922754355284188
l2 norm of weights: 2.251274551803386

Iteration Number: 8708
Loss: 10.838881830656195
l2 norm of gradients: 0.059219351486039344
l2 norm of weights: 2.2513087343618206

Iteration Number: 8709
Loss: 10.838686590603816
l2 norm of gradients: 0.05921116215464291
l2 norm of weights: 2.251342922942154

Iteration Number: 8710
Loss: 10.838491446802024
l2 norm of gradients: 0.05920297555735048
l2 norm of weights: 2.2513771175384885

Iteration Number: 8711
Loss: 10.838296399190785
l2 norm of gradients: 0.059194791692860516
l2 norm of weights: 2.25141131814493

Iteration Number: 8712
Loss: 10.8381014477101
l2 norm of gradients: 0.05918661055987234
l2 norm of weights: 2.2514455247555887

Iteration Number: 8713
Loss: 10.837906592300008
l2 norm of gradients: 0.05917843215708582
l2 norm of weights: 2.2514797373645807

Iteration Number: 8714
Loss: 10.8377118329006
l2 norm of gradients: 0.059170256483201705

```

12 norm of weights: 2.251513955966025
-----
Iteration Number: 8715
Loss: 10.837517169451994
12 norm of gradients: 0.059162083536921366
12 norm of weights: 2.251548180554046
-----
Iteration Number: 8716
Loss: 10.837322601894357
12 norm of gradients: 0.05915391331694692
12 norm of weights: 2.2515824111227727
-----
Iteration Number: 8717
Loss: 10.837128130167898
12 norm of gradients: 0.059145745821981136
12 norm of weights: 2.251616647666338
-----
Iteration Number: 8718
Loss: 10.836933754212858
12 norm of gradients: 0.059137581050727535
12 norm of weights: 2.251650890178881
-----
Iteration Number: 8719
Loss: 10.836739473969532
12 norm of gradients: 0.0591294190018904
12 norm of weights: 2.2516851386545427
-----
Iteration Number: 8720
Loss: 10.836545289378238
12 norm of gradients: 0.05912125967417465
12 norm of weights: 2.2517193930874706
-----
Iteration Number: 8721
Loss: 10.836351200379355
12 norm of gradients: 0.05911310306628596
12 norm of weights: 2.2517536534718166
-----
Iteration Number: 8722
Loss: 10.836157206913283
12 norm of gradients: 0.059104949176930696
12 norm of weights: 2.2517879198017363
-----
Iteration Number: 8723
Loss: 10.835963308920473
12 norm of gradients: 0.05909679800481589
12 norm of weights: 2.2518221920713897
-----
Iteration Number: 8724

```

Loss: 10.83576950634142
l2 norm of gradients: 0.05908864954864934
l2 norm of weights: 2.251856470274943

Iteration Number: 8725
Loss: 10.835575799116645
l2 norm of gradients: 0.059080503807139584
l2 norm of weights: 2.2518907544065647

Iteration Number: 8726
Loss: 10.83538218718673
l2 norm of gradients: 0.05907236077899576
l2 norm of weights: 2.251925044460429

Iteration Number: 8727
Loss: 10.835188670492277
l2 norm of gradients: 0.05906422046292785
l2 norm of weights: 2.251959340430715

Iteration Number: 8728
Loss: 10.83499524897394
l2 norm of gradients: 0.059056082857646346
l2 norm of weights: 2.2519936423116054

Iteration Number: 8729
Loss: 10.834801922572408
l2 norm of gradients: 0.05904794796186267
l2 norm of weights: 2.252027950097287

Iteration Number: 8730
Loss: 10.834608691228418
l2 norm of gradients: 0.05903981577428878
l2 norm of weights: 2.252062263781953

Iteration Number: 8731
Loss: 10.834415554882739
l2 norm of gradients: 0.059031686293637455
l2 norm of weights: 2.2520965833597995

Iteration Number: 8732
Loss: 10.834222513476185
l2 norm of gradients: 0.05902355951862206
l2 norm of weights: 2.252130908825027

Iteration Number: 8733
Loss: 10.834029566949605
l2 norm of gradients: 0.059015435447956836
l2 norm of weights: 2.2521652401718417

```

-----
Iteration Number: 8734
Loss: 10.833836715243892
l2 norm of gradients: 0.05900731408035651
l2 norm of weights: 2.2521995773944523
-----
Iteration Number: 8735
Loss: 10.833643958299987
l2 norm of gradients: 0.05899919541453667
l2 norm of weights: 2.252233920487075
-----
Iteration Number: 8736
Loss: 10.833451296058856
l2 norm of gradients: 0.05899107944921358
l2 norm of weights: 2.2522682694439276
-----
Iteration Number: 8737
Loss: 10.83325872846151
l2 norm of gradients: 0.05898296618310418
l2 norm of weights: 2.2523026242592334
-----
Iteration Number: 8738
Loss: 10.833066255449008
l2 norm of gradients: 0.0589748556149261
l2 norm of weights: 2.252336984927221
-----
Iteration Number: 8739
Loss: 10.832873876962438
l2 norm of gradients: 0.058966747743397656
l2 norm of weights: 2.2523713514421213
-----
Iteration Number: 8740
Loss: 10.832681592942938
l2 norm of gradients: 0.058958642567238
l2 norm of weights: 2.2524057237981725
-----
Iteration Number: 8741
Loss: 10.832489403331678
l2 norm of gradients: 0.058950540085166814
l2 norm of weights: 2.252440101989615
-----
Iteration Number: 8742
Loss: 10.832297308069872
l2 norm of gradients: 0.05894244029590451
l2 norm of weights: 2.252474486010694
-----
Iteration Number: 8743
Loss: 10.832105307098775

```

```

12 norm of gradients: 0.05893434319817225
12 norm of weights: 2.252508875855661
-----
Iteration Number: 8744
Loss: 10.831913400359676
12 norm of gradients: 0.058926248790691936
12 norm of weights: 2.25254327151877
-----
Iteration Number: 8745
Loss: 10.83172158779391
12 norm of gradients: 0.05891815707218606
12 norm of weights: 2.252577672994279
-----
Iteration Number: 8746
Loss: 10.831529869342852
12 norm of gradients: 0.05891006804137792
12 norm of weights: 2.252612080276452
-----
Iteration Number: 8747
Loss: 10.831338244947911
12 norm of gradients: 0.05890198169699137
12 norm of weights: 2.252646493359558
-----
Iteration Number: 8748
Loss: 10.831146714550538
12 norm of gradients: 0.05889389803775112
12 norm of weights: 2.2526809122378677
-----
Iteration Number: 8749
Loss: 10.830955278092233
12 norm of gradients: 0.05888581706238246
12 norm of weights: 2.2527153369056587
-----
Iteration Number: 8750
Loss: 10.830763935514515
12 norm of gradients: 0.05887773876961141
12 norm of weights: 2.2527497673572117
-----
Iteration Number: 8751
Loss: 10.830572686758968
12 norm of gradients: 0.05886966315816471
12 norm of weights: 2.2527842035868124
-----
Iteration Number: 8752
Loss: 10.830381531767197
12 norm of gradients: 0.05886159022676975
12 norm of weights: 2.2528186455887504
-----

```

Iteration Number: 8753
Loss: 10.830190470480854
l2 norm of gradients: 0.05885351997415467
l2 norm of weights: 2.2528530933573214

Iteration Number: 8754
Loss: 10.82999950284163
l2 norm of gradients: 0.05884545239904826
l2 norm of weights: 2.2528875468868224

Iteration Number: 8755
Loss: 10.829808628791255
l2 norm of gradients: 0.05883738750017997
l2 norm of weights: 2.2529220061715582

Iteration Number: 8756
Loss: 10.829617848271496
l2 norm of gradients: 0.05882932527628006
l2 norm of weights: 2.252956471205835

Iteration Number: 8757
Loss: 10.829427161224167
l2 norm of gradients: 0.05882126572607933
l2 norm of weights: 2.252990941983966

Iteration Number: 8758
Loss: 10.82923656759111
l2 norm of gradients: 0.05881320884830943
l2 norm of weights: 2.253025418500267

Iteration Number: 8759
Loss: 10.829046067314222
l2 norm of gradients: 0.05880515464170257
l2 norm of weights: 2.253059900749059

Iteration Number: 8760
Loss: 10.828855660335421
l2 norm of gradients: 0.05879710310499172
l2 norm of weights: 2.253094388724667

Iteration Number: 8761
Loss: 10.828665346596681
l2 norm of gradients: 0.05878905423691049
l2 norm of weights: 2.253128882421421

Iteration Number: 8762
Loss: 10.828475126040006
l2 norm of gradients: 0.05878100803619328


```

12 norm of weights: 2.2531633818336547
-----
Iteration Number: 8763
Loss: 10.82828499860744
12 norm of gradients: 0.058772964501575056
12 norm of weights: 2.2531978869557068
-----
Iteration Number: 8764
Loss: 10.828094964241073
12 norm of gradients: 0.058764923631791526
12 norm of weights: 2.253232397781919
-----
Iteration Number: 8765
Loss: 10.827905022883026
12 norm of gradients: 0.0587568854255791
12 norm of weights: 2.2532669143066397
-----
Iteration Number: 8766
Loss: 10.827715174475463
12 norm of gradients: 0.05874884988167487
12 norm of weights: 2.2533014365242194
-----
Iteration Number: 8767
Loss: 10.82752541896059
12 norm of gradients: 0.0587408169988166
12 norm of weights: 2.253335964429015
-----
Iteration Number: 8768
Loss: 10.827335756280645
12 norm of gradients: 0.058732786775742736
12 norm of weights: 2.2533704980153857
-----
Iteration Number: 8769
Loss: 10.82714618637791
12 norm of gradients: 0.058724759211192426
12 norm of weights: 2.2534050372776964
-----
Iteration Number: 8770
Loss: 10.826956709194707
12 norm of gradients: 0.05871673430390552
12 norm of weights: 2.253439582210316
-----
Iteration Number: 8771
Loss: 10.8267673246734
12 norm of gradients: 0.05870871205262247
12 norm of weights: 2.2534741328076184
-----
Iteration Number: 8772

```

Loss: 10.82657803275638
l2 norm of gradients: 0.05870069245608455
l2 norm of weights: 2.25350868906398

Iteration Number: 8773
Loss: 10.826388833386089
l2 norm of gradients: 0.05869267551303361
l2 norm of weights: 2.2535432509737845

Iteration Number: 8774
Loss: 10.826199726505008
l2 norm of gradients: 0.058684661222212206
l2 norm of weights: 2.2535778185314164

Iteration Number: 8775
Loss: 10.826010712055647
l2 norm of gradients: 0.058676649582363634
l2 norm of weights: 2.253612391731268

Iteration Number: 8776
Loss: 10.825821789980562
l2 norm of gradients: 0.058668640592231765
l2 norm of weights: 2.253646970567733

Iteration Number: 8777
Loss: 10.825632960222352
l2 norm of gradients: 0.058660634250561235
l2 norm of weights: 2.2536815550352114

Iteration Number: 8778
Loss: 10.825444222723645
l2 norm of gradients: 0.05865263055609736
l2 norm of weights: 2.2537161451281067

Iteration Number: 8779
Loss: 10.825255577427118
l2 norm of gradients: 0.058644629507586116
l2 norm of weights: 2.2537507408408275

Iteration Number: 8780
Loss: 10.825067024275478
l2 norm of gradients: 0.05863663110377414
l2 norm of weights: 2.2537853421677854

Iteration Number: 8781
Loss: 10.824878563211476
l2 norm of gradients: 0.05862863534340878
l2 norm of weights: 2.2538199491033977

```

-----
Iteration Number: 8782
Loss: 10.8246901941779
l2 norm of gradients: 0.058620642225238044
l2 norm of weights: 2.2538545616420844
-----
Iteration Number: 8783
Loss: 10.824501917117582
l2 norm of gradients: 0.05861265174801067
l2 norm of weights: 2.253889179778272
-----
Iteration Number: 8784
Loss: 10.824313731973382
l2 norm of gradients: 0.05860466391047599
l2 norm of weights: 2.2539238035063893
-----
Iteration Number: 8785
Loss: 10.82412563868821
l2 norm of gradients: 0.05859667871138407
l2 norm of weights: 2.2539584328208706
-----
Iteration Number: 8786
Loss: 10.82393763720501
l2 norm of gradients: 0.05858869614948565
l2 norm of weights: 2.2539930677161544
-----
Iteration Number: 8787
Loss: 10.823749727466762
l2 norm of gradients: 0.05858071622353216
l2 norm of weights: 2.2540277081866824
-----
Iteration Number: 8788
Loss: 10.823561909416487
l2 norm of gradients: 0.05857273893227565
l2 norm of weights: 2.254062354226902
-----
Iteration Number: 8789
Loss: 10.823374182997249
l2 norm of gradients: 0.0585647642744689
l2 norm of weights: 2.254097005831264
-----
Iteration Number: 8790
Loss: 10.823186548152144
l2 norm of gradients: 0.05855679224886537
l2 norm of weights: 2.2541316629942245
-----
Iteration Number: 8791
Loss: 10.822999004824307

```

```

12 norm of gradients: 0.05854882285421914
12 norm of weights: 2.254166325710243
-----
Iteration Number: 8792
Loss: 10.822811552956923
12 norm of gradients: 0.058540856089285026
12 norm of weights: 2.2542009939737824
-----
Iteration Number: 8793
Loss: 10.822624192493196
12 norm of gradients: 0.058532891952818494
12 norm of weights: 2.2542356677793127
-----
Iteration Number: 8794
Loss: 10.822436923376381
12 norm of gradients: 0.058524930443575644
12 norm of weights: 2.2542703471213055
-----
Iteration Number: 8795
Loss: 10.822249745549778
12 norm of gradients: 0.058516971560313345
12 norm of weights: 2.2543050319942375
-----
Iteration Number: 8796
Loss: 10.822062658956707
12 norm of gradients: 0.05850901530178906
12 norm of weights: 2.2543397223925905
-----
Iteration Number: 8797
Loss: 10.821875663540538
12 norm of gradients: 0.058501061666760966
12 norm of weights: 2.2543744183108494
-----
Iteration Number: 8798
Loss: 10.821688759244687
12 norm of gradients: 0.05849311065398785
12 norm of weights: 2.254409119743504
-----
Iteration Number: 8799
Loss: 10.821501946012589
12 norm of gradients: 0.05848516226222926
12 norm of weights: 2.254443826685048
-----
Iteration Number: 8800
Loss: 10.821315223787733
12 norm of gradients: 0.05847721649024533
12 norm of weights: 2.25447853912998
-----

```

Iteration Number: 8801
Loss: 10.821128592513638
l2 norm of gradients: 0.05846927333679695
l2 norm of weights: 2.2545132570728024

Iteration Number: 8802
Loss: 10.820942052133868
l2 norm of gradients: 0.05846133280064559
l2 norm of weights: 2.2545479805080215

Iteration Number: 8803
Loss: 10.820755602592017
l2 norm of gradients: 0.05845339488055345
l2 norm of weights: 2.2545827094301485

Iteration Number: 8804
Loss: 10.820569243831727
l2 norm of gradients: 0.058445459575283396
l2 norm of weights: 2.2546174438336988

Iteration Number: 8805
Loss: 10.820382975796669
l2 norm of gradients: 0.05843752688359897
l2 norm of weights: 2.254652183713192

Iteration Number: 8806
Loss: 10.82019679843056
l2 norm of gradients: 0.05842959680426436
l2 norm of weights: 2.254686929063151

Iteration Number: 8807
Loss: 10.82001071167715
l2 norm of gradients: 0.058421669336044396
l2 norm of weights: 2.254721679878104

Iteration Number: 8808
Loss: 10.819824715480225
l2 norm of gradients: 0.05841374447770461
l2 norm of weights: 2.2547564361525834

Iteration Number: 8809
Loss: 10.819638809783621
l2 norm of gradients: 0.058405822228011224
l2 norm of weights: 2.2547911978811257

Iteration Number: 8810
Loss: 10.819452994531197
l2 norm of gradients: 0.05839790258573108

```
12 norm of weights: 2.254825965058272
-----
Iteration Number: 8811
Loss: 10.819267269666858
12 norm of gradients: 0.05838998554963173
12 norm of weights: 2.2548607376785657
-----
Iteration Number: 8812
Loss: 10.81908163513455
12 norm of gradients: 0.05838207111848136
12 norm of weights: 2.254895515736557
-----
Iteration Number: 8813
Loss: 10.818896090878253
12 norm of gradients: 0.0583741592910488
12 norm of weights: 2.2549302992267988
-----
Iteration Number: 8814
Loss: 10.818710636841981
12 norm of gradients: 0.05836625006610358
12 norm of weights: 2.254965088143849
-----
Iteration Number: 8815
Loss: 10.818525272969795
12 norm of gradients: 0.05835834344241593
12 norm of weights: 2.254999882482269
-----
Iteration Number: 8816
Loss: 10.818339999205786
12 norm of gradients: 0.05835043941875668
12 norm of weights: 2.2550346822366247
-----
Iteration Number: 8817
Loss: 10.818154815494086
12 norm of gradients: 0.05834253799389734
12 norm of weights: 2.2550694874014865
-----
Iteration Number: 8818
Loss: 10.81796972177887
12 norm of gradients: 0.05833463916661008
12 norm of weights: 2.2551042979714286
-----
Iteration Number: 8819
Loss: 10.817784718004336
12 norm of gradients: 0.05832674293566776
12 norm of weights: 2.2551391139410293
-----
Iteration Number: 8820
```

Loss: 10.817599804114739
l2 norm of gradients: 0.058318849299843885
l2 norm of weights: 2.255173935304872

Iteration Number: 8821
Loss: 10.817414980054359
l2 norm of gradients: 0.058310958257912625
l2 norm of weights: 2.255208762057543

Iteration Number: 8822
Loss: 10.817230245767513
l2 norm of gradients: 0.05830306980864876
l2 norm of weights: 2.255243594193634

Iteration Number: 8823
Loss: 10.817045601198569
l2 norm of gradients: 0.05829518395082784
l2 norm of weights: 2.2552784317077395

Iteration Number: 8824
Loss: 10.81686104629192
l2 norm of gradients: 0.058287300683226
l2 norm of weights: 2.25531327459446

Iteration Number: 8825
Loss: 10.816676580992
l2 norm of gradients: 0.058279420004620046
l2 norm of weights: 2.2553481228483983

Iteration Number: 8826
Loss: 10.816492205243279
l2 norm of gradients: 0.05827154191378741
l2 norm of weights: 2.2553829764641633

Iteration Number: 8827
Loss: 10.81630791899027
l2 norm of gradients: 0.05826366640950629
l2 norm of weights: 2.255417835436366

Iteration Number: 8828
Loss: 10.816123722177522
l2 norm of gradients: 0.058255793490555347
l2 norm of weights: 2.255452699759623

Iteration Number: 8829
Loss: 10.815939614749617
l2 norm of gradients: 0.058247923155714115
l2 norm of weights: 2.2554875694285554

```

-----
Iteration Number: 8830
Loss: 10.815755596651174
l2 norm of gradients: 0.058240055403762696
l2 norm of weights: 2.2555224444377866
-----
Iteration Number: 8831
Loss: 10.815571667826863
l2 norm of gradients: 0.05823219023348183
l2 norm of weights: 2.2555573247819463
-----
Iteration Number: 8832
Loss: 10.815387828221377
l2 norm of gradients: 0.05822432764365295
l2 norm of weights: 2.2555922104556663
-----
Iteration Number: 8833
Loss: 10.815204077779448
l2 norm of gradients: 0.05821646763305807
l2 norm of weights: 2.2556271014535847
-----
Iteration Number: 8834
Loss: 10.815020416445856
l2 norm of gradients: 0.05820861020047998
l2 norm of weights: 2.2556619977703423
-----
Iteration Number: 8835
Loss: 10.814836844165404
l2 norm of gradients: 0.058200755344702024
l2 norm of weights: 2.255696899400584
-----
Iteration Number: 8836
Loss: 10.814653360882948
l2 norm of gradients: 0.05819290306450825
l2 norm of weights: 2.2557318063389604
-----
Iteration Number: 8837
Loss: 10.814469966543365
l2 norm of gradients: 0.05818505335868336
l2 norm of weights: 2.255766718580124
-----
Iteration Number: 8838
Loss: 10.814286661091582
l2 norm of gradients: 0.058177206226012655
l2 norm of weights: 2.255801636118733
-----
Iteration Number: 8839
Loss: 10.814103444472561

```



```

12 norm of gradients: 0.05816936166528218
12 norm of weights: 2.2558365589494493
-----
Iteration Number: 8840
Loss: 10.813920316631295
12 norm of gradients: 0.05816151967527858
12 norm of weights: 2.255871487066939
-----
Iteration Number: 8841
Loss: 10.81373727751282
12 norm of gradients: 0.05815368025478912
12 norm of weights: 2.2559064204658723
-----
Iteration Number: 8842
Loss: 10.813554327062203
12 norm of gradients: 0.05814584340260182
12 norm of weights: 2.2559413591409236
-----
Iteration Number: 8843
Loss: 10.81337146522456
12 norm of gradients: 0.05813800911750522
12 norm of weights: 2.255976303086771
-----
Iteration Number: 8844
Loss: 10.813188691945038
12 norm of gradients: 0.0581301773982886
12 norm of weights: 2.2560112522980975
-----
Iteration Number: 8845
Loss: 10.813006007168816
12 norm of gradients: 0.05812234824374191
12 norm of weights: 2.2560462067695894
-----
Iteration Number: 8846
Loss: 10.812823410841114
12 norm of gradients: 0.05811452165265569
12 norm of weights: 2.256081166495938
-----
Iteration Number: 8847
Loss: 10.812640902907194
12 norm of gradients: 0.05810669762382113
12 norm of weights: 2.2561161314718374
-----
Iteration Number: 8848
Loss: 10.812458483312348
12 norm of gradients: 0.0580988761560301
12 norm of weights: 2.2561511016919873
-----

```

Iteration Number: 8849
Loss: 10.81227615200191
l2 norm of gradients: 0.05809105724807509
l2 norm of weights: 2.2561860771510904

Iteration Number: 8850
Loss: 10.812093908921248
l2 norm of gradients: 0.05808324089874929
l2 norm of weights: 2.256221057843854

Iteration Number: 8851
Loss: 10.81191175401577
l2 norm of gradients: 0.058075427106846475
l2 norm of weights: 2.25625604376499

Iteration Number: 8852
Loss: 10.811729687230914
l2 norm of gradients: 0.05806761587116113
l2 norm of weights: 2.2562910349092133

Iteration Number: 8853
Loss: 10.811547708512165
l2 norm of gradients: 0.058059807190488345
l2 norm of weights: 2.2563260312712434

Iteration Number: 8854
Loss: 10.81136581780504
l2 norm of gradients: 0.058052001063623876
l2 norm of weights: 2.2563610328458044

Iteration Number: 8855
Loss: 10.811184015055092
l2 norm of gradients: 0.058044197489364076
l2 norm of weights: 2.2563960396276235

Iteration Number: 8856
Loss: 10.811002300207909
l2 norm of gradients: 0.05803639646650601
l2 norm of weights: 2.2564310516114325

Iteration Number: 8857
Loss: 10.810820673209125
l2 norm of gradients: 0.05802859799384741
l2 norm of weights: 2.2564660687919678

Iteration Number: 8858
Loss: 10.810639134004395
l2 norm of gradients: 0.058020802070186527

```

12 norm of weights: 2.2565010911639685
-----
Iteration Number: 8859
Loss: 10.810457682539433
12 norm of gradients: 0.05801300869432237
12 norm of weights: 2.2565361187221793
-----
Iteration Number: 8860
Loss: 10.81027631875997
12 norm of gradients: 0.05800521786505459
12 norm of weights: 2.256571151461348
-----
Iteration Number: 8861
Loss: 10.810095042611785
12 norm of gradients: 0.05799742958118344
12 norm of weights: 2.256606189376227
-----
Iteration Number: 8862
Loss: 10.809913854040683
12 norm of gradients: 0.05798964384150974
12 norm of weights: 2.2566412324615723
-----
Iteration Number: 8863
Loss: 10.809732752992522
12 norm of gradients: 0.057981860644835154
12 norm of weights: 2.2566762807121443
-----
Iteration Number: 8864
Loss: 10.809551739413177
12 norm of gradients: 0.05797407998996181
12 norm of weights: 2.2567113341227074
-----
Iteration Number: 8865
Loss: 10.809370813248579
12 norm of gradients: 0.0579663018756926
12 norm of weights: 2.25674639268803
-----
Iteration Number: 8866
Loss: 10.809189974444681
12 norm of gradients: 0.05795852630083095
12 norm of weights: 2.256781456402885
-----
Iteration Number: 8867
Loss: 10.809009222947486
12 norm of gradients: 0.05795075326418097
12 norm of weights: 2.2568165252620482
-----
Iteration Number: 8868

```

```

Loss: 10.808828558703018
12 norm of gradients: 0.05794298276454744
12 norm of weights: 2.2568515992603007
-----
Iteration Number: 8869
Loss: 10.808647981657346
12 norm of gradients: 0.05793521480073579
12 norm of weights: 2.2568866783924264
-----
Iteration Number: 8870
Loss: 10.80846749175658
12 norm of gradients: 0.057927449371551996
12 norm of weights: 2.256921762653215
-----
Iteration Number: 8871
Loss: 10.80828708894686
12 norm of gradients: 0.05791968647580276
12 norm of weights: 2.2569568520374585
-----
Iteration Number: 8872
Loss: 10.808106773174366
12 norm of gradients: 0.05791192611229542
12 norm of weights: 2.256991946539954
-----
Iteration Number: 8873
Loss: 10.807926544385305
12 norm of gradients: 0.05790416827983795
12 norm of weights: 2.257027046155502
-----
Iteration Number: 8874
Loss: 10.807746402525936
12 norm of gradients: 0.05789641297723887
12 norm of weights: 2.2570621508789075
-----
Iteration Number: 8875
Loss: 10.807566347542544
12 norm of gradients: 0.05788866020330744
12 norm of weights: 2.2570972607049793
-----
Iteration Number: 8876
Loss: 10.807386379381452
12 norm of gradients: 0.057880909956853584
12 norm of weights: 2.25713237562853
-----
Iteration Number: 8877
Loss: 10.80720649798902
12 norm of gradients: 0.05787316223668775
12 norm of weights: 2.257167495644377

```

```

-----
Iteration Number: 8878
Loss: 10.80702670331165
l2 norm of gradients: 0.05786541704162107
l2 norm of weights: 2.257202620747341
-----
Iteration Number: 8879
Loss: 10.806846995295766
l2 norm of gradients: 0.05785767437046536
l2 norm of weights: 2.2572377509322465
-----
Iteration Number: 8880
Loss: 10.806667373887848
l2 norm of gradients: 0.05784993422203304
l2 norm of weights: 2.2572728861939226
-----
Iteration Number: 8881
Loss: 10.806487839034391
l2 norm of gradients: 0.057842196595137126
l2 norm of weights: 2.2573080265272027
-----
Iteration Number: 8882
Loss: 10.806308390681947
l2 norm of gradients: 0.05783446148859132
l2 norm of weights: 2.2573431719269235
-----
Iteration Number: 8883
Loss: 10.806129028777084
l2 norm of gradients: 0.05782672890120994
l2 norm of weights: 2.2573783223879254
-----
Iteration Number: 8884
Loss: 10.805949753266425
l2 norm of gradients: 0.057818998831807945
l2 norm of weights: 2.2574134779050543
-----
Iteration Number: 8885
Loss: 10.805770564096614
l2 norm of gradients: 0.057811271279200896
l2 norm of weights: 2.2574486384731585
-----
Iteration Number: 8886
Loss: 10.805591461214343
l2 norm of gradients: 0.057803546242205045
l2 norm of weights: 2.2574838040870913
-----
Iteration Number: 8887
Loss: 10.805412444566333

```

l2 norm of gradients: 0.057795823719637186
l2 norm of weights: 2.257518974741709

Iteration Number: 8888
Loss: 10.805233514099342
l2 norm of gradients: 0.05778810371031488
l2 norm of weights: 2.2575541504318735

Iteration Number: 8889
Loss: 10.805054669760164
l2 norm of gradients: 0.0577803862130562
l2 norm of weights: 2.2575893311524484

Iteration Number: 8890
Loss: 10.804875911495634
l2 norm of gradients: 0.05777267122667989
l2 norm of weights: 2.257624516898304

Iteration Number: 8891
Loss: 10.804697239252615
l2 norm of gradients: 0.05776495875000535
l2 norm of weights: 2.2576597076643123

Iteration Number: 8892
Loss: 10.804518652978011
l2 norm of gradients: 0.05775724878185259
l2 norm of weights: 2.2576949034453504

Iteration Number: 8893
Loss: 10.804340152618764
l2 norm of gradients: 0.05774954132104221
l2 norm of weights: 2.2577301042362987

Iteration Number: 8894
Loss: 10.804161738121843
l2 norm of gradients: 0.0577418363663955
l2 norm of weights: 2.2577653100320427

Iteration Number: 8895
Loss: 10.803983409434267
l2 norm of gradients: 0.05773413391673438
l2 norm of weights: 2.2578005208274705

Iteration Number: 8896
Loss: 10.803805166503079
l2 norm of gradients: 0.05772643397088136
l2 norm of weights: 2.2578357366174755

Iteration Number: 8897
Loss: 10.803627009275358
l2 norm of gradients: 0.057718736527659596
l2 norm of weights: 2.2578709573969538

Iteration Number: 8898
Loss: 10.80344893769823
l2 norm of gradients: 0.05771104158589287
l2 norm of weights: 2.257906183160806

Iteration Number: 8899
Loss: 10.803270951718842
l2 norm of gradients: 0.05770334914440557
l2 norm of weights: 2.2579414139039367

Iteration Number: 8900
Loss: 10.803093051284392
l2 norm of gradients: 0.05769565920202277
l2 norm of weights: 2.2579766496212548

Iteration Number: 8901
Loss: 10.802915236342098
l2 norm of gradients: 0.0576879717575701
l2 norm of weights: 2.2580118903076727

Iteration Number: 8902
Loss: 10.802737506839229
l2 norm of gradients: 0.05768028680987391
l2 norm of weights: 2.2580471359581065

Iteration Number: 8903
Loss: 10.802559862723077
l2 norm of gradients: 0.05767260435776106
l2 norm of weights: 2.2580823865674766

Iteration Number: 8904
Loss: 10.802382303940977
l2 norm of gradients: 0.057664924400059096
l2 norm of weights: 2.258117642130708

Iteration Number: 8905
Loss: 10.8022048304403
l2 norm of gradients: 0.05765724693559621
l2 norm of weights: 2.2581529026427285

Iteration Number: 8906
Loss: 10.802027442168448
l2 norm of gradients: 0.057649571963201175

```

12 norm of weights: 2.2581881680984703
-----
Iteration Number: 8907
Loss: 10.801850139072865
12 norm of gradients: 0.05764189948170341
12 norm of weights: 2.258223438492869
-----
Iteration Number: 8908
Loss: 10.801672921101021
12 norm of gradients: 0.05763422948993301
12 norm of weights: 2.2582587138208656
-----
Iteration Number: 8909
Loss: 10.801495788200434
12 norm of gradients: 0.05762656198672055
12 norm of weights: 2.2582939940774036
-----
Iteration Number: 8910
Loss: 10.801318740318646
12 norm of gradients: 0.057618896970897365
12 norm of weights: 2.2583292792574303
-----
Iteration Number: 8911
Loss: 10.801141777403238
12 norm of gradients: 0.05761123444129535
12 norm of weights: 2.258364569355899
-----
Iteration Number: 8912
Loss: 10.800964899401835
12 norm of gradients: 0.05760357439674707
12 norm of weights: 2.2583998643677643
-----
Iteration Number: 8913
Loss: 10.800788106262086
12 norm of gradients: 0.05759591683608564
12 norm of weights: 2.2584351642879854
-----
Iteration Number: 8914
Loss: 10.800611397931682
12 norm of gradients: 0.05758826175814487
12 norm of weights: 2.2584704691115274
-----
Iteration Number: 8915
Loss: 10.800434774358346
12 norm of gradients: 0.05758060916175914
12 norm of weights: 2.2585057788333565
-----
Iteration Number: 8916

```



```

Loss: 10.800258235489835
l2 norm of gradients: 0.05757295904576344
l2 norm of weights: 2.2585410934484442
-----
Iteration Number: 8917
Loss: 10.80008178127395
l2 norm of gradients: 0.05756531140899343
l2 norm of weights: 2.2585764129517663
-----
Iteration Number: 8918
Loss: 10.799905411658518
l2 norm of gradients: 0.05755766625028538
l2 norm of weights: 2.258611737338302
-----
Iteration Number: 8919
Loss: 10.799729126591405
l2 norm of gradients: 0.05755002356847617
l2 norm of weights: 2.2586470666030336
-----
Iteration Number: 8920
Loss: 10.799552926020514
l2 norm of gradients: 0.05754238336240324
l2 norm of weights: 2.258682400740949
-----
Iteration Number: 8921
Loss: 10.799376809893781
l2 norm of gradients: 0.05753474563090479
l2 norm of weights: 2.258717739747038
-----
Iteration Number: 8922
Loss: 10.799200778159179
l2 norm of gradients: 0.057527110372819476
l2 norm of weights: 2.258753083616296
-----
Iteration Number: 8923
Loss: 10.79902483076471
l2 norm of gradients: 0.057519477586986674
l2 norm of weights: 2.258788432343722
-----
Iteration Number: 8924
Loss: 10.798848967658422
l2 norm of gradients: 0.05751184727224634
l2 norm of weights: 2.2588237859243177
-----
Iteration Number: 8925
Loss: 10.79867318878839
l2 norm of gradients: 0.057504219427439114
l2 norm of weights: 2.2588591443530897

```

```

-----
Iteration Number: 8926
Loss: 10.798497494102728
l2 norm of gradients: 0.05749659405140613
l2 norm of weights: 2.258894507625049
-----
Iteration Number: 8927
Loss: 10.798321883549582
l2 norm of gradients: 0.05748897114298921
l2 norm of weights: 2.258929875735208
-----
Iteration Number: 8928
Loss: 10.798146357077131
l2 norm of gradients: 0.057481350701030824
l2 norm of weights: 2.258965248678586
-----
Iteration Number: 8929
Loss: 10.797970914633606
l2 norm of gradients: 0.05747373272437399
l2 norm of weights: 2.2590006264502054
-----
Iteration Number: 8930
Loss: 10.797795556167245
l2 norm of gradients: 0.057466117211862386
l2 norm of weights: 2.25903600904509
-----
Iteration Number: 8931
Loss: 10.797620281626344
l2 norm of gradients: 0.05745850416234026
l2 norm of weights: 2.2590713964582716
-----
Iteration Number: 8932
Loss: 10.797445090959227
l2 norm of gradients: 0.05745089357465254
l2 norm of weights: 2.2591067886847815
-----
Iteration Number: 8933
Loss: 10.797269984114248
l2 norm of gradients: 0.057443285447644724
l2 norm of weights: 2.2591421857196585
-----
Iteration Number: 8934
Loss: 10.797094961039804
l2 norm of gradients: 0.057435679780162924
l2 norm of weights: 2.2591775875579434
-----
Iteration Number: 8935
Loss: 10.79692002168432

```

l2 norm of gradients: 0.057428076571053845
l2 norm of weights: 2.259212994194681

Iteration Number: 8936
Loss: 10.796745165996263
l2 norm of gradients: 0.05742047581916487
l2 norm of weights: 2.2592484056249202

Iteration Number: 8937
Loss: 10.796570393924126
l2 norm of gradients: 0.05741287752334395
l2 norm of weights: 2.2592838218437135

Iteration Number: 8938
Loss: 10.796395705416446
l2 norm of gradients: 0.057405281682439614
l2 norm of weights: 2.259319242846118

Iteration Number: 8939
Loss: 10.796221100421786
l2 norm of gradients: 0.05739768829530108
l2 norm of weights: 2.2593546686271933

Iteration Number: 8940
Loss: 10.79604657888875
l2 norm of gradients: 0.05739009736077812
l2 norm of weights: 2.259390099182004

Iteration Number: 8941
Loss: 10.79587214076598
l2 norm of gradients: 0.057382508877721124
l2 norm of weights: 2.2594255345056182

Iteration Number: 8942
Loss: 10.795697786002144
l2 norm of gradients: 0.05737492284498112
l2 norm of weights: 2.2594609745931074

Iteration Number: 8943
Loss: 10.795523514545948
l2 norm of gradients: 0.057367339261409724
l2 norm of weights: 2.259496419439548

Iteration Number: 8944
Loss: 10.795349326346136
l2 norm of gradients: 0.057359758125859174
l2 norm of weights: 2.2595318690400186

Iteration Number: 8945
Loss: 10.795175221351483
l2 norm of gradients: 0.05735217943718231
l2 norm of weights: 2.2595673233896028

Iteration Number: 8946
Loss: 10.795001199510796
l2 norm of gradients: 0.05734460319423255
l2 norm of weights: 2.259602782483388

Iteration Number: 8947
Loss: 10.794827260772928
l2 norm of gradients: 0.05733702939586391
l2 norm of weights: 2.2596382463164653

Iteration Number: 8948
Loss: 10.794653405086756
l2 norm of gradients: 0.05732945804093118
l2 norm of weights: 2.2596737148839288

Iteration Number: 8949
Loss: 10.794479632401199
l2 norm of gradients: 0.05732188912828955
l2 norm of weights: 2.259709188180878

Iteration Number: 8950
Loss: 10.794305942665195
l2 norm of gradients: 0.05731432265679486
l2 norm of weights: 2.259744666202414

Iteration Number: 8951
Loss: 10.794132335827738
l2 norm of gradients: 0.05730675862530363
l2 norm of weights: 2.259780148943644

Iteration Number: 8952
Loss: 10.793958811837845
l2 norm of gradients: 0.057299197032673
l2 norm of weights: 2.259815636399678

Iteration Number: 8953
Loss: 10.793785370644565
l2 norm of gradients: 0.05729163787776058
l2 norm of weights: 2.259851128565629

Iteration Number: 8954
Loss: 10.79361201219699
l2 norm of gradients: 0.0572840811594247

```

12 norm of weights: 2.2598866254366152
-----
Iteration Number: 8955
Loss: 10.79343873644424
12 norm of gradients: 0.05727652687652428
12 norm of weights: 2.2599221270077576
-----
Iteration Number: 8956
Loss: 10.793265543335473
12 norm of gradients: 0.05726897502791881
12 norm of weights: 2.259957633274182
-----
Iteration Number: 8957
Loss: 10.79309243281988
12 norm of gradients: 0.057261425612468435
12 norm of weights: 2.259993144231016
-----
Iteration Number: 8958
Loss: 10.792919404846682
12 norm of gradients: 0.057253878629033805
12 norm of weights: 2.2600286598733943
-----
Iteration Number: 8959
Loss: 10.792746459365143
12 norm of gradients: 0.05724633407647635
12 norm of weights: 2.2600641801964514
-----
Iteration Number: 8960
Loss: 10.792573596324555
12 norm of gradients: 0.05723879195365787
12 norm of weights: 2.2600997051953287
-----
Iteration Number: 8961
Loss: 10.792400815674247
12 norm of gradients: 0.05723125225944099
12 norm of weights: 2.26013523486517
-----
Iteration Number: 8962
Loss: 10.792228117363582
12 norm of gradients: 0.05722371499268879
12 norm of weights: 2.2601707692011233
-----
Iteration Number: 8963
Loss: 10.792055501341958
12 norm of gradients: 0.05721618015226502
12 norm of weights: 2.2602063081983395
-----
Iteration Number: 8964

```

Loss: 10.791882967558806
l2 norm of gradients: 0.05720864773703395
l2 norm of weights: 2.2602418518519745

Iteration Number: 8965
Loss: 10.791710515963588
l2 norm of gradients: 0.057201117745860626
l2 norm of weights: 2.2602774001571873

Iteration Number: 8966
Loss: 10.791538146505808
l2 norm of gradients: 0.057193590177610484
l2 norm of weights: 2.260312953109141

Iteration Number: 8967
Loss: 10.791365859134995
l2 norm of gradients: 0.057186065031149695
l2 norm of weights: 2.2603485107030017

Iteration Number: 8968
Loss: 10.791193653800724
l2 norm of gradients: 0.057178542305345005
l2 norm of weights: 2.2603840729339404

Iteration Number: 8969
Loss: 10.791021530452591
l2 norm of gradients: 0.057171021999063695
l2 norm of weights: 2.260419639797131

Iteration Number: 8970
Loss: 10.790849489040236
l2 norm of gradients: 0.057163504111173775
l2 norm of weights: 2.2604552112877507

Iteration Number: 8971
Loss: 10.790677529513324
l2 norm of gradients: 0.05715598864054373
l2 norm of weights: 2.260490787400982

Iteration Number: 8972
Loss: 10.790505651821563
l2 norm of gradients: 0.05714847558604268
l2 norm of weights: 2.2605263681320102

Iteration Number: 8973
Loss: 10.790333855914696
l2 norm of gradients: 0.05714096494654038
l2 norm of weights: 2.260561953476024

```

-----
Iteration Number: 8974
Loss: 10.790162141742488
l2 norm of gradients: 0.05713345672090713
l2 norm of weights: 2.2605975434282164
-----
Iteration Number: 8975
Loss: 10.789990509254748
l2 norm of gradients: 0.05712595090801387
l2 norm of weights: 2.260633137983784
-----
Iteration Number: 8976
Loss: 10.78981895840132
l2 norm of gradients: 0.05711844750673208
l2 norm of weights: 2.2606687371379266
-----
Iteration Number: 8977
Loss: 10.78964748913207
l2 norm of gradients: 0.05711094651593394
l2 norm of weights: 2.2607043408858494
-----
Iteration Number: 8978
Loss: 10.789476101396916
l2 norm of gradients: 0.057103447934492096
l2 norm of weights: 2.2607399492227587
-----
Iteration Number: 8979
Loss: 10.789304795145794
l2 norm of gradients: 0.057095951761279906
l2 norm of weights: 2.2607755621438677
-----
Iteration Number: 8980
Loss: 10.78913357032868
l2 norm of gradients: 0.05708845799517124
l2 norm of weights: 2.26081117964439
-----
Iteration Number: 8981
Loss: 10.788962426895589
l2 norm of gradients: 0.05708096663504063
l2 norm of weights: 2.2608468017195453
-----
Iteration Number: 8982
Loss: 10.788791364796559
l2 norm of gradients: 0.05707347767976312
l2 norm of weights: 2.2608824283645563
-----
Iteration Number: 8983
Loss: 10.788620383981671

```

```

12 norm of gradients: 0.05706599112821443
12 norm of weights: 2.260918059574649
-----
Iteration Number: 8984
Loss: 10.788449484401035
12 norm of gradients: 0.057058506979270805
12 norm of weights: 2.2609536953450533
-----
Iteration Number: 8985
Loss: 10.788278666004798
12 norm of gradients: 0.05705102523180918
12 norm of weights: 2.2609893356710034
-----
Iteration Number: 8986
Loss: 10.788107928743138
12 norm of gradients: 0.05704354588470696
12 norm of weights: 2.2610249805477367
-----
Iteration Number: 8987
Loss: 10.787937272566268
12 norm of gradients: 0.05703606893684222
12 norm of weights: 2.261060629970494
-----
Iteration Number: 8988
Loss: 10.787766697424429
12 norm of gradients: 0.057028594387093616
12 norm of weights: 2.2610962839345206
-----
Iteration Number: 8989
Loss: 10.787596203267906
12 norm of gradients: 0.05702112223434038
12 norm of weights: 2.261131942435065
-----
Iteration Number: 8990
Loss: 10.787425790047015
12 norm of gradients: 0.05701365247746235
12 norm of weights: 2.2611676054673793
-----
Iteration Number: 8991
Loss: 10.7872554577121
12 norm of gradients: 0.057006185115339975
12 norm of weights: 2.261203273026719
-----
Iteration Number: 8992
Loss: 10.78708520621354
12 norm of gradients: 0.056998720146854226
12 norm of weights: 2.2612389451083446
-----

```


Iteration Number: 8993
Loss: 10.78691503550175
l2 norm of gradients: 0.05699125757088674
l2 norm of weights: 2.2612746217075186

Iteration Number: 8994
Loss: 10.786744945527182
l2 norm of gradients: 0.056983797386319694
l2 norm of weights: 2.261310302819508

Iteration Number: 8995
Loss: 10.786574936240314
l2 norm of gradients: 0.05697633959203589
l2 norm of weights: 2.2613459884395843

Iteration Number: 8996
Loss: 10.78640500759166
l2 norm of gradients: 0.05696888418691871
l2 norm of weights: 2.261381678563021

Iteration Number: 8997
Loss: 10.786235159531774
l2 norm of gradients: 0.056961431169852084
l2 norm of weights: 2.2614173731850964

Iteration Number: 8998
Loss: 10.78606539201123
l2 norm of gradients: 0.05695398053972057
l2 norm of weights: 2.261453072301092

Iteration Number: 8999
Loss: 10.78589570498065
l2 norm of gradients: 0.05694653229540934
l2 norm of weights: 2.261488775906294

Iteration Number: 9000
Loss: 10.78572609839068
l2 norm of gradients: 0.05693908643580411
l2 norm of weights: 2.26152448399599

Iteration Number: 9001
Loss: 10.785556572192
l2 norm of gradients: 0.056931642959791154
l2 norm of weights: 2.2615601965654735

Iteration Number: 9002
Loss: 10.785387126335332
l2 norm of gradients: 0.05692420186625741

```

12 norm of weights: 2.2615959136100408
-----
Iteration Number: 9003
Loss: 10.785217760771419
12 norm of gradients: 0.056916763154090375
12 norm of weights: 2.261631635124991
-----
Iteration Number: 9004
Loss: 10.785048475451044
12 norm of gradients: 0.05690932682217813
12 norm of weights: 2.2616673611056295
-----
Iteration Number: 9005
Loss: 10.784879270325021
12 norm of gradients: 0.05690189286940925
12 norm of weights: 2.261703091547262
-----
Iteration Number: 9006
Loss: 10.784710145344206
12 norm of gradients: 0.05689446129467309
12 norm of weights: 2.2617388264452
-----
Iteration Number: 9007
Loss: 10.784541100459478
12 norm of gradients: 0.056887032096859406
12 norm of weights: 2.2617745657947577
-----
Iteration Number: 9008
Loss: 10.784372135621743
12 norm of gradients: 0.05687960527485864
12 norm of weights: 2.261810309591254
-----
Iteration Number: 9009
Loss: 10.784203250781962
12 norm of gradients: 0.056872180827561854
12 norm of weights: 2.2618460578300104
-----
Iteration Number: 9010
Loss: 10.784034445891109
12 norm of gradients: 0.05686475875386053
12 norm of weights: 2.261881810506352
-----
Iteration Number: 9011
Loss: 10.783865720900204
12 norm of gradients: 0.05685733905264688
12 norm of weights: 2.2619175676156087
-----
Iteration Number: 9012

```

```

Loss: 10.783697075760292
l2 norm of gradients: 0.056849921722813654
l2 norm of weights: 2.261953329153113
-----
Iteration Number: 9013
Loss: 10.783528510422451
l2 norm of gradients: 0.056842506763254155
l2 norm of weights: 2.2619890951142003
-----
Iteration Number: 9014
Loss: 10.783360024837805
l2 norm of gradients: 0.05683509417286234
l2 norm of weights: 2.2620248654942117
-----
Iteration Number: 9015
Loss: 10.783191618957488
l2 norm of gradients: 0.056827683950532705
l2 norm of weights: 2.2620606402884906
-----
Iteration Number: 9016
Loss: 10.78302329273269
l2 norm of gradients: 0.056820276095160285
l2 norm of weights: 2.262096419492384
-----
Iteration Number: 9017
Loss: 10.78285504611462
l2 norm of gradients: 0.05681287060564079
l2 norm of weights: 2.262132203101243
-----
Iteration Number: 9018
Loss: 10.782686879054523
l2 norm of gradients: 0.05680546748087045
l2 norm of weights: 2.262167991110422
-----
Iteration Number: 9019
Loss: 10.782518791503689
l2 norm of gradients: 0.05679806671974611
l2 norm of weights: 2.262203783515279
-----
Iteration Number: 9020
Loss: 10.782350783413413
l2 norm of gradients: 0.05679066832116509
l2 norm of weights: 2.2622395803111757
-----
Iteration Number: 9021
Loss: 10.782182854735055
l2 norm of gradients: 0.05678327228402547
l2 norm of weights: 2.262275381493477

```

```

-----
Iteration Number: 9022
Loss: 10.782015005419984
l2 norm of gradients: 0.056775878607225744
l2 norm of weights: 2.262311187057553
-----
Iteration Number: 9023
Loss: 10.781847235419617
l2 norm of gradients: 0.05676848728966508
l2 norm of weights: 2.2623469969987746
-----
Iteration Number: 9024
Loss: 10.781679544685392
l2 norm of gradients: 0.05676109833024318
l2 norm of weights: 2.2623828113125186
-----
Iteration Number: 9025
Loss: 10.781511933168789
l2 norm of gradients: 0.0567537117278604
l2 norm of weights: 2.262418629994165
-----
Iteration Number: 9026
Loss: 10.781344400821316
l2 norm of gradients: 0.056746327481417586
l2 norm of weights: 2.2624544530390973
-----
Iteration Number: 9027
Loss: 10.781176947594517
l2 norm of gradients: 0.05673894558981613
l2 norm of weights: 2.2624902804427016
-----
Iteration Number: 9028
Loss: 10.781009573439967
l2 norm of gradients: 0.05673156605195815
l2 norm of weights: 2.262526112200369
-----
Iteration Number: 9029
Loss: 10.78084227830927
l2 norm of gradients: 0.05672418886674622
l2 norm of weights: 2.2625619483074924
-----
Iteration Number: 9030
Loss: 10.780675062154074
l2 norm of gradients: 0.05671681403308355
l2 norm of weights: 2.2625977887594706
-----
Iteration Number: 9031
Loss: 10.780507924926047

```

```

12 norm of gradients: 0.05670944154987386
12 norm of weights: 2.2626336335517045
-----
Iteration Number: 9032
Loss: 10.780340866576893
12 norm of gradients: 0.05670207141602152
12 norm of weights: 2.2626694826795983
-----
Iteration Number: 9033
Loss: 10.780173887058352
12 norm of gradients: 0.05669470363043143
12 norm of weights: 2.2627053361385614
-----
Iteration Number: 9034
Loss: 10.780006986322197
12 norm of gradients: 0.056687338192009065
12 norm of weights: 2.262741193924005
-----
Iteration Number: 9035
Loss: 10.779840164320234
12 norm of gradients: 0.05667997509966051
12 norm of weights: 2.2627770560313443
-----
Iteration Number: 9036
Loss: 10.779673421004293
12 norm of gradients: 0.056672614352292386
12 norm of weights: 2.262812922455999
-----
Iteration Number: 9037
Loss: 10.77950675632625
12 norm of gradients: 0.056665255948811936
12 norm of weights: 2.2628487931933914
-----
Iteration Number: 9038
Loss: 10.779340170237997
12 norm of gradients: 0.05665789988812692
12 norm of weights: 2.262884668238948
-----
Iteration Number: 9039
Loss: 10.779173662691477
12 norm of gradients: 0.056650546169145694
12 norm of weights: 2.262920547588098
-----
Iteration Number: 9040
Loss: 10.779007233638657
12 norm of gradients: 0.0566431947907772
12 norm of weights: 2.2629564312362747
-----

```

Iteration Number: 9041
Loss: 10.778840883031524
l2 norm of gradients: 0.05663584575193094
l2 norm of weights: 2.2629923191789154

Iteration Number: 9042
Loss: 10.778674610822124
l2 norm of gradients: 0.05662849905151699
l2 norm of weights: 2.2630282114114597

Iteration Number: 9043
Loss: 10.778508416962511
l2 norm of gradients: 0.056621154688446015
l2 norm of weights: 2.263064107929353

Iteration Number: 9044
Loss: 10.77834230140479
l2 norm of gradients: 0.056613812661629195
l2 norm of weights: 2.263100008728041

Iteration Number: 9045
Loss: 10.77817626410108
l2 norm of gradients: 0.05660647296997837
l2 norm of weights: 2.2631359138029756

Iteration Number: 9046
Loss: 10.77801030500355
l2 norm of gradients: 0.05659913561240587
l2 norm of weights: 2.2631718231496114

Iteration Number: 9047
Loss: 10.777844424064385
l2 norm of gradients: 0.05659180058782463
l2 norm of weights: 2.2632077367634063

Iteration Number: 9048
Loss: 10.777678621235822
l2 norm of gradients: 0.056584467895148156
l2 norm of weights: 2.263243654639822

Iteration Number: 9049
Loss: 10.777512896470109
l2 norm of gradients: 0.05657713753329053
l2 norm of weights: 2.2632795767743232

Iteration Number: 9050
Loss: 10.77734724971954
l2 norm of gradients: 0.056569809501166406

```

12 norm of weights: 2.2633155031623793
-----
Iteration Number: 9051
Loss: 10.777181680936437
12 norm of gradients: 0.05656248379769099
12 norm of weights: 2.2633514337994622
-----
Iteration Number: 9052
Loss: 10.777016190073159
12 norm of gradients: 0.056555160421780035
12 norm of weights: 2.263387368681047
-----
Iteration Number: 9053
Loss: 10.776850777082087
12 norm of gradients: 0.05654783937234988
12 norm of weights: 2.2634233078026136
-----
Iteration Number: 9054
Loss: 10.77668544191564
12 norm of gradients: 0.05654052064831751
12 norm of weights: 2.2634592511596443
-----
Iteration Number: 9055
Loss: 10.776520184526277
12 norm of gradients: 0.05653320424860037
12 norm of weights: 2.263495198747626
-----
Iteration Number: 9056
Loss: 10.776355004866476
12 norm of gradients: 0.05652589017211652
12 norm of weights: 2.263531150562048
-----
Iteration Number: 9057
Loss: 10.77618990288875
12 norm of gradients: 0.05651857841778452
12 norm of weights: 2.2635671065984035
-----
Iteration Number: 9058
Loss: 10.776024878545652
12 norm of gradients: 0.05651126898452363
12 norm of weights: 2.26360306685219
-----
Iteration Number: 9059
Loss: 10.775859931789757
12 norm of gradients: 0.0565039618712536
12 norm of weights: 2.2636390313189065
-----
Iteration Number: 9060

```

Loss: 10.77569506257368
l2 norm of gradients: 0.056496657076894705
l2 norm of weights: 2.2636749999940577

Iteration Number: 9061
Loss: 10.775530270850066
l2 norm of gradients: 0.05648935460036785
l2 norm of weights: 2.2637109728731506

Iteration Number: 9062
Loss: 10.775365556571584
l2 norm of gradients: 0.05648205444059448
l2 norm of weights: 2.2637469499516962

Iteration Number: 9063
Loss: 10.775200919690953
l2 norm of gradients: 0.05647475659649659
l2 norm of weights: 2.263782931225209

Iteration Number: 9064
Loss: 10.775036360160906
l2 norm of gradients: 0.056467461066996794
l2 norm of weights: 2.2638189166892055

Iteration Number: 9065
Loss: 10.774871877934213
l2 norm of gradients: 0.0564601678510182
l2 norm of weights: 2.2638549063392084

Iteration Number: 9066
Loss: 10.77470747296368
l2 norm of gradients: 0.05645287694748455
l2 norm of weights: 2.2638909001707415

Iteration Number: 9067
Loss: 10.774543145202145
l2 norm of gradients: 0.0564455883553201
l2 norm of weights: 2.2639268981793332

Iteration Number: 9068
Loss: 10.774378894602473
l2 norm of gradients: 0.05643830207344965
l2 norm of weights: 2.2639629003605157

Iteration Number: 9069
Loss: 10.774214721117563
l2 norm of gradients: 0.05643101810079864
l2 norm of weights: 2.2639989067098236


```

-----
Iteration Number: 9070
Loss: 10.774050624700347
l2 norm of gradients: 0.05642373643629299
l2 norm of weights: 2.264034917222796
-----
Iteration Number: 9071
Loss: 10.773886605303788
l2 norm of gradients: 0.05641645707885923
l2 norm of weights: 2.2640709318949743
-----
Iteration Number: 9072
Loss: 10.773722662880882
l2 norm of gradients: 0.056409180027424455
l2 norm of weights: 2.2641069507219047
-----
Iteration Number: 9073
Loss: 10.773558797384652
l2 norm of gradients: 0.0564019052809163
l2 norm of weights: 2.264142973699136
-----
Iteration Number: 9074
Loss: 10.773395008768155
l2 norm of gradients: 0.05639463283826296
l2 norm of weights: 2.2641790008222205
-----
Iteration Number: 9075
Loss: 10.773231296984488
l2 norm of gradients: 0.056387362698393215
l2 norm of weights: 2.2642150320867147
-----
Iteration Number: 9076
Loss: 10.77306766198677
l2 norm of gradients: 0.056380094860236374
l2 norm of weights: 2.264251067488177
-----
Iteration Number: 9077
Loss: 10.772904103728155
l2 norm of gradients: 0.05637282932272235
l2 norm of weights: 2.264287107022172
-----
Iteration Number: 9078
Loss: 10.772740622161821
l2 norm of gradients: 0.05636556608478157
l2 norm of weights: 2.264323150684264
-----
Iteration Number: 9079
Loss: 10.772577217240993

```

```

12 norm of gradients: 0.05635830514534499
12 norm of weights: 2.264359198470024
-----
Iteration Number: 9080
Loss: 10.772413888918917
12 norm of gradients: 0.056351046503344264
12 norm of weights: 2.2643952503750255
-----
Iteration Number: 9081
Loss: 10.772250637148872
12 norm of gradients: 0.05634379015771144
12 norm of weights: 2.2644313063948442
-----
Iteration Number: 9082
Loss: 10.77208746188417
12 norm of gradients: 0.05633653610737923
12 norm of weights: 2.2644673665250608
-----
Iteration Number: 9083
Loss: 10.771924363078154
12 norm of gradients: 0.056329284351280866
12 norm of weights: 2.2645034307612586
-----
Iteration Number: 9084
Loss: 10.771761340684197
12 norm of gradients: 0.056322034888350136
12 norm of weights: 2.2645394990990244
-----
Iteration Number: 9085
Loss: 10.771598394655708
12 norm of gradients: 0.0563147877175214
12 norm of weights: 2.2645755715339493
-----
Iteration Number: 9086
Loss: 10.771435524946122
12 norm of gradients: 0.05630754283772955
12 norm of weights: 2.264611648061627
-----
Iteration Number: 9087
Loss: 10.771272731508907
12 norm of gradients: 0.056300300247910065
12 norm of weights: 2.264647728677654
-----
Iteration Number: 9088
Loss: 10.771110014297566
12 norm of gradients: 0.05629305994699899
12 norm of weights: 2.2646838133776312
-----

```

Iteration Number: 9089
Loss: 10.770947373265631
l2 norm of gradients: 0.056285821933932854
l2 norm of weights: 2.2647199021571636

Iteration Number: 9090
Loss: 10.770784808366665
l2 norm of gradients: 0.05627858620764881
l2 norm of weights: 2.264755995011858

Iteration Number: 9091
Loss: 10.770622319554262
l2 norm of gradients: 0.05627135276708457
l2 norm of weights: 2.2647920919373257

Iteration Number: 9092
Loss: 10.770459906782044
l2 norm of gradients: 0.05626412161117836
l2 norm of weights: 2.2648281929291807

Iteration Number: 9093
Loss: 10.770297570003674
l2 norm of gradients: 0.05625689273886894
l2 norm of weights: 2.264864297983041

Iteration Number: 9094
Loss: 10.77013530917284
l2 norm of gradients: 0.05624966614909573
l2 norm of weights: 2.264900407094528

Iteration Number: 9095
Loss: 10.76997312424326
l2 norm of gradients: 0.056242441840798574
l2 norm of weights: 2.264936520259266

Iteration Number: 9096
Loss: 10.769811015168687
l2 norm of gradients: 0.05623521981291797
l2 norm of weights: 2.2649726374728827

Iteration Number: 9097
Loss: 10.769648981902899
l2 norm of gradients: 0.0562280000643949
l2 norm of weights: 2.2650087587310104

Iteration Number: 9098
Loss: 10.769487024399716
l2 norm of gradients: 0.056220782594170994

```

12 norm of weights: 2.265044884029283
-----
Iteration Number: 9099
Loss: 10.769325142612976
12 norm of gradients: 0.056213567401188276
12 norm of weights: 2.2650810133633397
-----
Iteration Number: 9100
Loss: 10.76916333649656
12 norm of gradients: 0.05620635448438949
12 norm of weights: 2.2651171467288216
-----
Iteration Number: 9101
Loss: 10.769001606004373
12 norm of gradients: 0.05619914384271781
12 norm of weights: 2.2651532841213733
-----
Iteration Number: 9102
Loss: 10.768839951090357
12 norm of gradients: 0.05619193547511705
12 norm of weights: 2.2651894255366436
-----
Iteration Number: 9103
Loss: 10.768678371708475
12 norm of gradients: 0.056184729380531456
12 norm of weights: 2.2652255709702844
-----
Iteration Number: 9104
Loss: 10.768516867812732
12 norm of gradients: 0.05617752555790599
12 norm of weights: 2.2652617204179504
-----
Iteration Number: 9105
Loss: 10.76835543935716
12 norm of gradients: 0.056170324006186066
12 norm of weights: 2.265297873875301
-----
Iteration Number: 9106
Loss: 10.768194086295818
12 norm of gradients: 0.056163124724317615
12 norm of weights: 2.265334031337997
-----
Iteration Number: 9107
Loss: 10.768032808582804
12 norm of gradients: 0.056155927711247194
12 norm of weights: 2.2653701928017043
-----
Iteration Number: 9108

```

Loss: 10.767871606172235
l2 norm of gradients: 0.056148732965921856
l2 norm of weights: 2.2654063582620916

Iteration Number: 9109
Loss: 10.767710479018275
l2 norm of gradients: 0.05614154048728926
l2 norm of weights: 2.2654425277148307

Iteration Number: 9110
Loss: 10.76754942707511
l2 norm of gradients: 0.05613435027429754
l2 norm of weights: 2.265478701155598

Iteration Number: 9111
Loss: 10.767388450296954
l2 norm of gradients: 0.0561271623258954
l2 norm of weights: 2.2655148785800705

Iteration Number: 9112
Loss: 10.767227548638052
l2 norm of gradients: 0.05611997664103219
l2 norm of weights: 2.2655510599839315

Iteration Number: 9113
Loss: 10.767066722052691
l2 norm of gradients: 0.05611279321865759
l2 norm of weights: 2.2655872453628665

Iteration Number: 9114
Loss: 10.76690597049518
l2 norm of gradients: 0.05610561205772213
l2 norm of weights: 2.2656234347125643

Iteration Number: 9115
Loss: 10.766745293919858
l2 norm of gradients: 0.05609843315717662
l2 norm of weights: 2.2656596280287165

Iteration Number: 9116
Loss: 10.766584692281095
l2 norm of gradients: 0.05609125651597252
l2 norm of weights: 2.265695825307019

Iteration Number: 9117
Loss: 10.766424165533298
l2 norm of gradients: 0.05608408213306187
l2 norm of weights: 2.2657320265431715

Iteration Number: 9118
Loss: 10.766263713630895
l2 norm of gradients: 0.056076910007397174
l2 norm of weights: 2.265768231732875

Iteration Number: 9119
Loss: 10.766103336528358
l2 norm of gradients: 0.05606974013793156
l2 norm of weights: 2.2658044408718365

Iteration Number: 9120
Loss: 10.765943034180177
l2 norm of gradients: 0.056062572523618674
l2 norm of weights: 2.2658406539557636

Iteration Number: 9121
Loss: 10.76578280654088
l2 norm of gradients: 0.05605540716341269
l2 norm of weights: 2.2658768709803696

Iteration Number: 9122
Loss: 10.765622653565018
l2 norm of gradients: 0.05604824405626832
l2 norm of weights: 2.2659130919413695

Iteration Number: 9123
Loss: 10.765462575207186
l2 norm of gradients: 0.05604108320114087
l2 norm of weights: 2.265949316834482

Iteration Number: 9124
Loss: 10.765302571421998
l2 norm of gradients: 0.056033924596986195
l2 norm of weights: 2.2659855456554308

Iteration Number: 9125
Loss: 10.7651426421641
l2 norm of gradients: 0.05602676824276056
l2 norm of weights: 2.2660217783999403

Iteration Number: 9126
Loss: 10.764982787388174
l2 norm of gradients: 0.05601961413742094
l2 norm of weights: 2.2660580150637393

Iteration Number: 9127
Loss: 10.76482300704893

```

12 norm of gradients: 0.05601246227992476
12 norm of weights: 2.2660942556425616
-----
Iteration Number: 9128
Loss: 10.76466330110111
12 norm of gradients: 0.05600531266923005
12 norm of weights: 2.266130500132141
-----
Iteration Number: 9129
Loss: 10.76450366949948
12 norm of gradients: 0.055998165304295286
12 norm of weights: 2.2661667485282173
-----
Iteration Number: 9130
Loss: 10.764344112198849
12 norm of gradients: 0.05599102018407961
12 norm of weights: 2.266203000826533
-----
Iteration Number: 9131
Loss: 10.76418462915404
12 norm of gradients: 0.05598387730754258
12 norm of weights: 2.2662392570228334
-----
Iteration Number: 9132
Loss: 10.76402522031992
12 norm of gradients: 0.0559767366736444
12 norm of weights: 2.266275517112867
-----
Iteration Number: 9133
Loss: 10.763865885651382
12 norm of gradients: 0.05596959828134576
12 norm of weights: 2.2663117810923863
-----
Iteration Number: 9134
Loss: 10.76370662510335
12 norm of gradients: 0.055962462129607884
12 norm of weights: 2.266348048957147
-----
Iteration Number: 9135
Loss: 10.763547438630775
12 norm of gradients: 0.05595532821739259
12 norm of weights: 2.2663843207029073
-----
Iteration Number: 9136
Loss: 10.763388326188643
12 norm of gradients: 0.055948196543662135
12 norm of weights: 2.26642059632543
-----

```

Iteration Number: 9137
Loss: 10.763229287731969
l2 norm of gradients: 0.0559410671073795
l2 norm of weights: 2.2664568758204804

Iteration Number: 9138
Loss: 10.763070323215796
l2 norm of gradients: 0.05593393990750796
l2 norm of weights: 2.2664931591838267

Iteration Number: 9139
Loss: 10.762911432595201
l2 norm of gradients: 0.055926814943011514
l2 norm of weights: 2.266529446411241

Iteration Number: 9140
Loss: 10.762752615825292
l2 norm of gradients: 0.05591969221285466
l2 norm of weights: 2.2665657374984995

Iteration Number: 9141
Loss: 10.7625938728612
l2 norm of gradients: 0.05591257171600242
l2 norm of weights: 2.2666020324413796

Iteration Number: 9142
Loss: 10.762435203658091
l2 norm of gradients: 0.05590545345142027
l2 norm of weights: 2.2666383312356637

Iteration Number: 9143
Loss: 10.762276608171168
l2 norm of gradients: 0.05589833741807441
l2 norm of weights: 2.2666746338771366

Iteration Number: 9144
Loss: 10.762118086355652
l2 norm of gradients: 0.055891223614931415
l2 norm of weights: 2.266710940361587

Iteration Number: 9145
Loss: 10.761959638166804
l2 norm of gradients: 0.055884112040958486
l2 norm of weights: 2.2667472506848068

Iteration Number: 9146
Loss: 10.761801263559907
l2 norm of gradients: 0.055877002695123294


```

12 norm of weights: 2.266783564842591
-----
Iteration Number: 9147
Loss: 10.761642962490278
12 norm of gradients: 0.05586989557639407
12 norm of weights: 2.2668198828307373
-----
Iteration Number: 9148
Loss: 10.76148473491327
12 norm of gradients: 0.05586279068373968
12 norm of weights: 2.266856204645048
-----
Iteration Number: 9149
Loss: 10.761326580784257
12 norm of gradients: 0.05585568801612936
12 norm of weights: 2.266892530281327
-----
Iteration Number: 9150
Loss: 10.761168500058645
12 norm of gradients: 0.05584858757253298
12 norm of weights: 2.2669288597353834
-----
Iteration Number: 9151
Loss: 10.761010492691877
12 norm of gradients: 0.05584148935192095
12 norm of weights: 2.2669651930030277
-----
Iteration Number: 9152
Loss: 10.760852558639414
12 norm of gradients: 0.05583439335326413
12 norm of weights: 2.2670015300800745
-----
Iteration Number: 9153
Loss: 10.76069469785676
12 norm of gradients: 0.05582729957553409
12 norm of weights: 2.2670378709623424
-----
Iteration Number: 9154
Loss: 10.760536910299443
12 norm of gradients: 0.05582020801770273
12 norm of weights: 2.267074215645652
-----
Iteration Number: 9155
Loss: 10.760379195923015
12 norm of gradients: 0.0558131186787426
12 norm of weights: 2.267110564125827
-----
Iteration Number: 9156

```

```

Loss: 10.760221554683072
l2 norm of gradients: 0.05580603155762675
l2 norm of weights: 2.267146916398697
-----
Iteration Number: 9157
Loss: 10.760063986535227
l2 norm of gradients: 0.055798946653328775
l2 norm of weights: 2.2671832724600907
-----
Iteration Number: 9158
Loss: 10.759906491435128
l2 norm of gradients: 0.05579186396482275
l2 norm of weights: 2.267219632305843
-----
Iteration Number: 9159
Loss: 10.759749069338453
l2 norm of gradients: 0.05578478349108342
l2 norm of weights: 2.267255995931792
-----
Iteration Number: 9160
Loss: 10.759591720200913
l2 norm of gradients: 0.055777705231085946
l2 norm of weights: 2.2672923633337785
-----
Iteration Number: 9161
Loss: 10.759434443978243
l2 norm of gradients: 0.055770629183806056
l2 norm of weights: 2.267328734507644
-----
Iteration Number: 9162
Loss: 10.759277240626211
l2 norm of gradients: 0.05576355534821998
l2 norm of weights: 2.267365109449238
-----
Iteration Number: 9163
Loss: 10.759120110100612
l2 norm of gradients: 0.055756483723304516
l2 norm of weights: 2.26740148815441
-----
Iteration Number: 9164
Loss: 10.758963052357279
l2 norm of gradients: 0.055749414308036975
l2 norm of weights: 2.2674378706190135
-----
Iteration Number: 9165
Loss: 10.758806067352065
l2 norm of gradients: 0.055742347101395216
l2 norm of weights: 2.267474256838905

```

```

-----
Iteration Number: 9166
Loss: 10.758649155040859
l2 norm of gradients: 0.0557352821023576
l2 norm of weights: 2.267510646809945
-----
Iteration Number: 9167
Loss: 10.758492315379575
l2 norm of gradients: 0.05572821930990308
l2 norm of weights: 2.2675470405279965
-----
Iteration Number: 9168
Loss: 10.75833554832416
l2 norm of gradients: 0.05572115872301103
l2 norm of weights: 2.267583437988926
-----
Iteration Number: 9169
Loss: 10.758178853830591
l2 norm of gradients: 0.05571410034066149
l2 norm of weights: 2.2676198391886033
-----
Iteration Number: 9170
Loss: 10.758022231854873
l2 norm of gradients: 0.05570704416183485
l2 norm of weights: 2.267656244122901
-----
Iteration Number: 9171
Loss: 10.757865682353044
l2 norm of gradients: 0.05569999018551222
l2 norm of weights: 2.2676926527876953
-----
Iteration Number: 9172
Loss: 10.757709205281168
l2 norm of gradients: 0.05569293841067515
l2 norm of weights: 2.267729065178866
-----
Iteration Number: 9173
Loss: 10.757552800595336
l2 norm of gradients: 0.05568588883630571
l2 norm of weights: 2.267765481292295
-----
Iteration Number: 9174
Loss: 10.757396468251676
l2 norm of gradients: 0.055678841461386484
l2 norm of weights: 2.2678019011238684
-----
Iteration Number: 9175
Loss: 10.75724020820634

```

l2 norm of gradients: 0.055671796284900675
l2 norm of weights: 2.2678383246694747

Iteration Number: 9176
Loss: 10.757084020415517
l2 norm of gradients: 0.05566475330583189
l2 norm of weights: 2.267874751925006

Iteration Number: 9177
Loss: 10.756927904835413
l2 norm of gradients: 0.05565771252316435
l2 norm of weights: 2.267911182886359

Iteration Number: 9178
Loss: 10.756771861422274
l2 norm of gradients: 0.05565067393588276
l2 norm of weights: 2.2679476175494306

Iteration Number: 9179
Loss: 10.756615890132375
l2 norm of gradients: 0.05564363754297237
l2 norm of weights: 2.267984055910124

Iteration Number: 9180
Loss: 10.756459990922012
l2 norm of gradients: 0.05563660334341894
l2 norm of weights: 2.268020497964342

Iteration Number: 9181
Loss: 10.756304163747519
l2 norm of gradients: 0.05562957133620882
l2 norm of weights: 2.2680569437079954

Iteration Number: 9182
Loss: 10.756148408565256
l2 norm of gradients: 0.055622541520328736
l2 norm of weights: 2.2680933931369935

Iteration Number: 9183
Loss: 10.755992725331616
l2 norm of gradients: 0.055615513894766155
l2 norm of weights: 2.268129846247252

Iteration Number: 9184
Loss: 10.755837114003016
l2 norm of gradients: 0.05560848845850887
l2 norm of weights: 2.2681663030346875

Iteration Number: 9185
Loss: 10.755681574535904
l2 norm of gradients: 0.05560146521054529
l2 norm of weights: 2.2682027634952213

Iteration Number: 9186
Loss: 10.75552610688676
l2 norm of gradients: 0.055594444149864385
l2 norm of weights: 2.2682392276247785

Iteration Number: 9187
Loss: 10.755370711012093
l2 norm of gradients: 0.05558742527545554
l2 norm of weights: 2.2682756954192844

Iteration Number: 9188
Loss: 10.755215386868432
l2 norm of gradients: 0.05558040858630874
l2 norm of weights: 2.268312166874671

Iteration Number: 9189
Loss: 10.755060134412355
l2 norm of gradients: 0.055573394081414505
l2 norm of weights: 2.268348641986871

Iteration Number: 9190
Loss: 10.75490495360045
l2 norm of gradients: 0.05556638175976384
l2 norm of weights: 2.268385120751821

Iteration Number: 9191
Loss: 10.754749844389345
l2 norm of gradients: 0.05555937162034828
l2 norm of weights: 2.2684216031654616

Iteration Number: 9192
Loss: 10.754594806735692
l2 norm of gradients: 0.055552363662159876
l2 norm of weights: 2.2684580892237354

Iteration Number: 9193
Loss: 10.754439840596174
l2 norm of gradients: 0.0555453578841912
l2 norm of weights: 2.268494578922589

Iteration Number: 9194
Loss: 10.754284945927509
l2 norm of gradients: 0.05553835428543541

```

12 norm of weights: 2.268531072257971
-----
Iteration Number: 9195
Loss: 10.754130122686432
12 norm of gradients: 0.0555313528648861
12 norm of weights: 2.268567569225835
-----
Iteration Number: 9196
Loss: 10.753975370829714
12 norm of gradients: 0.05552435362153741
12 norm of weights: 2.2686040698221364
-----
Iteration Number: 9197
Loss: 10.753820690314162
12 norm of gradients: 0.055517356554384015
12 norm of weights: 2.2686405740428333
-----
Iteration Number: 9198
Loss: 10.7536660810966
12 norm of gradients: 0.055510361662421105
12 norm of weights: 2.268677081883888
-----
Iteration Number: 9199
Loss: 10.75351154313389
12 norm of gradients: 0.055503368944644395
12 norm of weights: 2.2687135933412663
-----
Iteration Number: 9200
Loss: 10.753357076382912
12 norm of gradients: 0.055496378400050124
12 norm of weights: 2.2687501084109356
-----
Iteration Number: 9201
Loss: 10.75320268080059
12 norm of gradients: 0.055489390027635005
12 norm of weights: 2.268786627088868
-----
Iteration Number: 9202
Loss: 10.753048356343868
12 norm of gradients: 0.05548240382639634
12 norm of weights: 2.268823149371038
-----
Iteration Number: 9203
Loss: 10.75289410296972
12 norm of gradients: 0.0554754197953319
12 norm of weights: 2.268859675253423
-----
Iteration Number: 9204

```

```

Loss: 10.752739920635147
l2 norm of gradients: 0.05546843793344
l2 norm of weights: 2.268896204732004
-----
Iteration Number: 9205
Loss: 10.752585809297182
l2 norm of gradients: 0.05546145823971947
l2 norm of weights: 2.2689327378027646
-----
Iteration Number: 9206
Loss: 10.752431768912892
l2 norm of gradients: 0.055454480713169635
l2 norm of weights: 2.2689692744616923
-----
Iteration Number: 9207
Loss: 10.75227779943936
l2 norm of gradients: 0.05544750535279035
l2 norm of weights: 2.269005814704778
-----
Iteration Number: 9208
Loss: 10.752123900833716
l2 norm of gradients: 0.05544053215758202
l2 norm of weights: 2.269042358528013
-----
Iteration Number: 9209
Loss: 10.751970073053098
l2 norm of gradients: 0.05543356112654551
l2 norm of weights: 2.2690789059273957
-----
Iteration Number: 9210
Loss: 10.75181631605469
l2 norm of gradients: 0.055426592258682246
l2 norm of weights: 2.269115456898925
-----
Iteration Number: 9211
Loss: 10.75166262979569
l2 norm of gradients: 0.05541962555299418
l2 norm of weights: 2.2691520114386043
-----
Iteration Number: 9212
Loss: 10.751509014233344
l2 norm of gradients: 0.055412661008483695
l2 norm of weights: 2.2691885695424383
-----
Iteration Number: 9213
Loss: 10.751355469324904
l2 norm of gradients: 0.055405698624153836
l2 norm of weights: 2.269225131206436

```

```

-----
Iteration Number: 9214
Loss: 10.751201995027674
l2 norm of gradients: 0.055398738399008024
l2 norm of weights: 2.269261696426611
-----
Iteration Number: 9215
Loss: 10.751048591298972
l2 norm of gradients: 0.05539178033205026
l2 norm of weights: 2.2692982651989766
-----
Iteration Number: 9216
Loss: 10.750895258096142
l2 norm of gradients: 0.05538482442228507
l2 norm of weights: 2.269334837519552
-----
Iteration Number: 9217
Loss: 10.750741995376572
l2 norm of gradients: 0.0553778706687174
l2 norm of weights: 2.2693714133843583
-----
Iteration Number: 9218
Loss: 10.750588803097662
l2 norm of gradients: 0.055370919070352895
l2 norm of weights: 2.26940799278942
-----
Iteration Number: 9219
Loss: 10.750435681216851
l2 norm of gradients: 0.055363969626197496
l2 norm of weights: 2.2694445757307653
-----
Iteration Number: 9220
Loss: 10.75028262969161
l2 norm of gradients: 0.05535702233525787
l2 norm of weights: 2.269481162204424
-----
Iteration Number: 9221
Loss: 10.750129648479424
l2 norm of gradients: 0.05535007719654104
l2 norm of weights: 2.26951775220643
-----
Iteration Number: 9222
Loss: 10.749976737537821
l2 norm of gradients: 0.055343134209054595
l2 norm of weights: 2.269554345732821
-----
Iteration Number: 9223
Loss: 10.74982389682435

```


l2 norm of gradients: 0.055336193371806654
l2 norm of weights: 2.2695909427796357

Iteration Number: 9224
Loss: 10.74967112629659
l2 norm of gradients: 0.05532925468380584
l2 norm of weights: 2.269627543342918

Iteration Number: 9225
Loss: 10.749518425912152
l2 norm of gradients: 0.055322318144061244
l2 norm of weights: 2.269664147418714

Iteration Number: 9226
Loss: 10.749365795628671
l2 norm of gradients: 0.05531538375158254
l2 norm of weights: 2.269700755003072

Iteration Number: 9227
Loss: 10.749213235403811
l2 norm of gradients: 0.05530845150537988
l2 norm of weights: 2.2697373660920452

Iteration Number: 9228
Loss: 10.74906074519527
l2 norm of gradients: 0.05530152140446391
l2 norm of weights: 2.2697739806816886

Iteration Number: 9229
Loss: 10.748908324960768
l2 norm of gradients: 0.05529459344784583
l2 norm of weights: 2.2698105987680606

Iteration Number: 9230
Loss: 10.748755974658057
l2 norm of gradients: 0.055287667634537276
l2 norm of weights: 2.269847220347223

Iteration Number: 9231
Loss: 10.748603694244917
l2 norm of gradients: 0.05528074396355052
l2 norm of weights: 2.2698838454152406

Iteration Number: 9232
Loss: 10.748451483679151
l2 norm of gradients: 0.05527382243389818
l2 norm of weights: 2.2699204739681798

Iteration Number: 9233
Loss: 10.748299342918603
l2 norm of gradients: 0.05526690304459352
l2 norm of weights: 2.2699571060021126

Iteration Number: 9234
Loss: 10.74814727192113
l2 norm of gradients: 0.0552599857946503
l2 norm of weights: 2.269993741513112

Iteration Number: 9235
Loss: 10.747995270644633
l2 norm of gradients: 0.05525307068308272
l2 norm of weights: 2.270030380497256

Iteration Number: 9236
Loss: 10.747843339047028
l2 norm of gradients: 0.055246157708905515
l2 norm of weights: 2.270067022950623

Iteration Number: 9237
Loss: 10.747691477086267
l2 norm of gradients: 0.05523924687113395
l2 norm of weights: 2.2701036688692966

Iteration Number: 9238
Loss: 10.747539684720328
l2 norm of gradients: 0.05523233816878378
l2 norm of weights: 2.270140318249363

Iteration Number: 9239
Loss: 10.747387961907217
l2 norm of gradients: 0.055225431600871275
l2 norm of weights: 2.2701769710869106

Iteration Number: 9240
Loss: 10.747236308604968
l2 norm of gradients: 0.055218527166413234
l2 norm of weights: 2.2702136273780322

Iteration Number: 9241
Loss: 10.74708472477165
l2 norm of gradients: 0.05521162486442693
l2 norm of weights: 2.2702502871188224

Iteration Number: 9242
Loss: 10.746933210365345
l2 norm of gradients: 0.05520472469393017

```

12 norm of weights: 2.27028695030538
-----
Iteration Number: 9243
Loss: 10.746781765344178
12 norm of gradients: 0.0551978266539412
12 norm of weights: 2.270323616933806
-----
Iteration Number: 9244
Loss: 10.746630389666299
12 norm of gradients: 0.05519093074347885
12 norm of weights: 2.2703602870002046
-----
Iteration Number: 9245
Loss: 10.74647908328988
12 norm of gradients: 0.055184036961562474
12 norm of weights: 2.2703969605006833
-----
Iteration Number: 9246
Loss: 10.746327846173129
12 norm of gradients: 0.05517714530721187
12 norm of weights: 2.270433637431352
-----
Iteration Number: 9247
Loss: 10.746176678274272
12 norm of gradients: 0.05517025577944735
12 norm of weights: 2.270470317788324
-----
Iteration Number: 9248
Loss: 10.74602557955158
12 norm of gradients: 0.05516336837728975
12 norm of weights: 2.2705070015677165
-----
Iteration Number: 9249
Loss: 10.745874549963332
12 norm of gradients: 0.0551564830997604
12 norm of weights: 2.2705436887656485
-----
Iteration Number: 9250
Loss: 10.745723589467852
12 norm of gradients: 0.05514959994588114
12 norm of weights: 2.2705803793782424
-----
Iteration Number: 9251
Loss: 10.745572698023478
12 norm of gradients: 0.055142718914674316
12 norm of weights: 2.2706170734016236
-----
Iteration Number: 9252

```

Loss: 10.745421875588587
l2 norm of gradients: 0.055135840005162806
l2 norm of weights: 2.2706537708319208

Iteration Number: 9253
Loss: 10.74527112212158
l2 norm of gradients: 0.05512896321636993
l2 norm of weights: 2.2706904716652656

Iteration Number: 9254
Loss: 10.745120437580887
l2 norm of gradients: 0.055122088547319564
l2 norm of weights: 2.2707271758977927

Iteration Number: 9255
Loss: 10.74496982192496
l2 norm of gradients: 0.05511521599703603
l2 norm of weights: 2.270763883525639

Iteration Number: 9256
Loss: 10.744819275112288
l2 norm of gradients: 0.05510834556454425
l2 norm of weights: 2.2708005945449456

Iteration Number: 9257
Loss: 10.744668797101387
l2 norm of gradients: 0.05510147724886952
l2 norm of weights: 2.2708373089518563

Iteration Number: 9258
Loss: 10.744518387850789
l2 norm of gradients: 0.05509461104903779
l2 norm of weights: 2.270874026742517

Iteration Number: 9259
Loss: 10.744368047319073
l2 norm of gradients: 0.055087746964075375
l2 norm of weights: 2.2709107479130783

Iteration Number: 9260
Loss: 10.74421777546483
l2 norm of gradients: 0.05508088499300919
l2 norm of weights: 2.2709474724596914

Iteration Number: 9261
Loss: 10.744067572246687
l2 norm of gradients: 0.05507402513486659
l2 norm of weights: 2.2709842003785132

```

-----
Iteration Number: 9262
Loss: 10.743917437623296
l2 norm of gradients: 0.05506716738867541
l2 norm of weights: 2.2710209316657015
-----
Iteration Number: 9263
Loss: 10.743767371553334
l2 norm of gradients: 0.0550603117534641
l2 norm of weights: 2.2710576663174185
-----
Iteration Number: 9264
Loss: 10.743617373995512
l2 norm of gradients: 0.055053458228261465
l2 norm of weights: 2.271094404329828
-----
Iteration Number: 9265
Loss: 10.743467444908568
l2 norm of gradients: 0.05504660681209697
l2 norm of weights: 2.2711311456990986
-----
Iteration Number: 9266
Loss: 10.743317584251265
l2 norm of gradients: 0.05503975750400042
l2 norm of weights: 2.2711678904214
-----
Iteration Number: 9267
Loss: 10.743167791982394
l2 norm of gradients: 0.05503291030300226
l2 norm of weights: 2.271204638492906
-----
Iteration Number: 9268
Loss: 10.74301806806077
l2 norm of gradients: 0.05502606520813331
l2 norm of weights: 2.2712413899097936
-----
Iteration Number: 9269
Loss: 10.742868412445251
l2 norm of gradients: 0.05501922221842502
l2 norm of weights: 2.2712781446682415
-----
Iteration Number: 9270
Loss: 10.742718825094698
l2 norm of gradients: 0.05501238133290916
l2 norm of weights: 2.271314902764433
-----
Iteration Number: 9271
Loss: 10.742569305968024

```

```

12 norm of gradients: 0.05500554255061818
12 norm of weights: 2.2713516641945524
-----
Iteration Number: 9272
Loss: 10.742419855024158
12 norm of gradients: 0.054998705870584946
12 norm of weights: 2.27138842895479
-----
Iteration Number: 9273
Loss: 10.742270472222051
12 norm of gradients: 0.05499187129184282
12 norm of weights: 2.2714251970413355
-----
Iteration Number: 9274
Loss: 10.7421211575207
12 norm of gradients: 0.05498503881342574
12 norm of weights: 2.2714619684503843
-----
Iteration Number: 9275
Loss: 10.741971910879105
12 norm of gradients: 0.05497820843436795
12 norm of weights: 2.2714987431781335
-----
Iteration Number: 9276
Loss: 10.741822732256319
12 norm of gradients: 0.05497138015370442
12 norm of weights: 2.271535521220783
-----
Iteration Number: 9277
Loss: 10.741673621611401
12 norm of gradients: 0.05496455397047048
12 norm of weights: 2.2715723025745373
-----
Iteration Number: 9278
Loss: 10.741524578903451
12 norm of gradients: 0.05495772988370199
12 norm of weights: 2.271609087235601
-----
Iteration Number: 9279
Loss: 10.741375604091592
12 norm of gradients: 0.05495090789243532
12 norm of weights: 2.271645875200185
-----
Iteration Number: 9280
Loss: 10.741226697134977
12 norm of gradients: 0.05494408799570727
12 norm of weights: 2.2716826664645002
-----

```

Iteration Number: 9281
Loss: 10.741077857992781
l2 norm of gradients: 0.05493727019255526
l2 norm of weights: 2.271719461024762

Iteration Number: 9282
Loss: 10.740929086624213
l2 norm of gradients: 0.05493045448201707
l2 norm of weights: 2.271756258877189

Iteration Number: 9283
Loss: 10.7407803829885
l2 norm of gradients: 0.05492364086313112
l2 norm of weights: 2.271793060018002

Iteration Number: 9284
Loss: 10.740631747044915
l2 norm of gradients: 0.05491682933493617
l2 norm of weights: 2.2718298644434247

Iteration Number: 9285
Loss: 10.740483178752735
l2 norm of gradients: 0.05491001989647158
l2 norm of weights: 2.271866672149684

Iteration Number: 9286
Loss: 10.74033467807128
l2 norm of gradients: 0.05490321254677718
l2 norm of weights: 2.27190348313301

Iteration Number: 9287
Loss: 10.740186244959895
l2 norm of gradients: 0.05489640728489329
l2 norm of weights: 2.271940297389636

Iteration Number: 9288
Loss: 10.740037879377946
l2 norm of gradients: 0.05488960410986075
l2 norm of weights: 2.2719771149157966

Iteration Number: 9289
Loss: 10.739889581284837
l2 norm of gradients: 0.05488280302072079
l2 norm of weights: 2.2720139357077316

Iteration Number: 9290
Loss: 10.739741350639989
l2 norm of gradients: 0.05487600401651529

```

12 norm of weights: 2.2720507597616817
-----
Iteration Number: 9291
Loss: 10.739593187402852
12 norm of gradients: 0.05486920709628649
12 norm of weights: 2.272087587073892
-----
Iteration Number: 9292
Loss: 10.739445091532911
12 norm of gradients: 0.05486241225907719
12 norm of weights: 2.27212441764061
-----
Iteration Number: 9293
Loss: 10.739297062989674
12 norm of gradients: 0.05485561950393069
12 norm of weights: 2.272161251458086
-----
Iteration Number: 9294
Loss: 10.739149101732671
12 norm of gradients: 0.05484882882989079
12 norm of weights: 2.272198088522573
-----
Iteration Number: 9295
Loss: 10.739001207721465
12 norm of gradients: 0.05484204023600164
12 norm of weights: 2.272234928830328
-----
Iteration Number: 9296
Loss: 10.738853380915645
12 norm of gradients: 0.05483525372130812
12 norm of weights: 2.27227177237761
-----
Iteration Number: 9297
Loss: 10.738705621274827
12 norm of gradients: 0.054828469284855394
12 norm of weights: 2.2723086191606803
-----
Iteration Number: 9298
Loss: 10.738557928758654
12 norm of gradients: 0.05482168692568924
12 norm of weights: 2.2723454691758045
-----
Iteration Number: 9299
Loss: 10.738410303326798
12 norm of gradients: 0.05481490664285586
12 norm of weights: 2.2723823224192503
-----
Iteration Number: 9300

```


Loss: 10.738262744938957
l2 norm of gradients: 0.054808128435402025
l2 norm of weights: 2.2724191788872896

Iteration Number: 9301
Loss: 10.738115253554849
l2 norm of gradients: 0.05480135230237487
l2 norm of weights: 2.2724560385761943

Iteration Number: 9302
Loss: 10.737967829134236
l2 norm of gradients: 0.05479457824282214
l2 norm of weights: 2.272492901482243

Iteration Number: 9303
Loss: 10.73782047163689
l2 norm of gradients: 0.05478780625579204
l2 norm of weights: 2.2725297676017133

Iteration Number: 9304
Loss: 10.73767318102262
l2 norm of gradients: 0.054781036340333225
l2 norm of weights: 2.2725666369308892

Iteration Number: 9305
Loss: 10.737525957251254
l2 norm of gradients: 0.05477426849549485
l2 norm of weights: 2.2726035094660553

Iteration Number: 9306
Loss: 10.73737880028266
l2 norm of gradients: 0.05476750272032658
l2 norm of weights: 2.2726403852035

Iteration Number: 9307
Loss: 10.73723171007672
l2 norm of gradients: 0.05476073901387859
l2 norm of weights: 2.2726772641395154

Iteration Number: 9308
Loss: 10.737084686593347
l2 norm of gradients: 0.05475397737520147
l2 norm of weights: 2.272714146270394

Iteration Number: 9309
Loss: 10.736937729792484
l2 norm of gradients: 0.054747217803346376
l2 norm of weights: 2.272751031592434

```

-----
Iteration Number: 9310
Loss: 10.7367908396341
l2 norm of gradients: 0.05474046029736489
l2 norm of weights: 2.2727879201019343
-----
Iteration Number: 9311
Loss: 10.736644016078191
l2 norm of gradients: 0.054733704856309164
l2 norm of weights: 2.272824811795198
-----
Iteration Number: 9312
Loss: 10.736497259084775
l2 norm of gradients: 0.05472695147923169
l2 norm of weights: 2.2728617066685306
-----
Iteration Number: 9313
Loss: 10.736350568613902
l2 norm of gradients: 0.05472020016518562
l2 norm of weights: 2.272898604718241
-----
Iteration Number: 9314
Loss: 10.73620394462565
l2 norm of gradients: 0.054713450913224486
l2 norm of weights: 2.27293550594064
-----
Iteration Number: 9315
Loss: 10.73605738708012
l2 norm of gradients: 0.054706703722402314
l2 norm of weights: 2.2729724103320423
-----
Iteration Number: 9316
Loss: 10.73591089593744
l2 norm of gradients: 0.05469995859177368
l2 norm of weights: 2.273009317888765
-----
Iteration Number: 9317
Loss: 10.73576447115777
l2 norm of gradients: 0.05469321552039358
l2 norm of weights: 2.273046228607128
-----
Iteration Number: 9318
Loss: 10.735618112701289
l2 norm of gradients: 0.0546864745073175
l2 norm of weights: 2.2730831424834537
-----
Iteration Number: 9319
Loss: 10.735471820528206

```

```

12 norm of gradients: 0.05467973555160142
12 norm of weights: 2.2731200595140684
-----
Iteration Number: 9320
Loss: 10.735325594598764
12 norm of gradients: 0.054672998652301866
12 norm of weights: 2.2731569796953006
-----
Iteration Number: 9321
Loss: 10.735179434873219
12 norm of gradients: 0.054666263808475744
12 norm of weights: 2.2731939030234822
-----
Iteration Number: 9322
Loss: 10.735033341311867
12 norm of gradients: 0.054659531019180564
12 norm of weights: 2.273230829494947
-----
Iteration Number: 9323
Loss: 10.73488731387502
12 norm of gradients: 0.054652800283474134
12 norm of weights: 2.2732677591060315
-----
Iteration Number: 9324
Loss: 10.734741352523027
12 norm of gradients: 0.05464607160041496
12 norm of weights: 2.2733046918530775
-----
Iteration Number: 9325
Loss: 10.734595457216255
12 norm of gradients: 0.05463934496906192
12 norm of weights: 2.2733416277324263
-----
Iteration Number: 9326
Loss: 10.7344496279151
12 norm of gradients: 0.054632620388474364
12 norm of weights: 2.273378566740425
-----
Iteration Number: 9327
Loss: 10.734303864579983
12 norm of gradients: 0.05462589785771219
12 norm of weights: 2.2734155088734207
-----
Iteration Number: 9328
Loss: 10.734158167171364
12 norm of gradients: 0.0546191773758357
12 norm of weights: 2.2734524541277663
-----

```

Iteration Number: 9329
Loss: 10.734012535649708
l2 norm of gradients: 0.054612458941905764
l2 norm of weights: 2.273489402499815

Iteration Number: 9330
Loss: 10.733866969975525
l2 norm of gradients: 0.05460574255498366
l2 norm of weights: 2.273526353985925

Iteration Number: 9331
Loss: 10.733721470109343
l2 norm of gradients: 0.05459902821413121
l2 norm of weights: 2.273563308582456

Iteration Number: 9332
Loss: 10.73357603601172
l2 norm of gradients: 0.054592315918410614
l2 norm of weights: 2.27360026628577

Iteration Number: 9333
Loss: 10.73343066764324
l2 norm of gradients: 0.05458560566688471
l2 norm of weights: 2.2736372270922334

Iteration Number: 9334
Loss: 10.733285364964507
l2 norm of gradients: 0.05457889745861666
l2 norm of weights: 2.2736741909982148

Iteration Number: 9335
Loss: 10.733140127936164
l2 norm of gradients: 0.054572191292670276
l2 norm of weights: 2.2737111580000855

Iteration Number: 9336
Loss: 10.732994956518867
l2 norm of gradients: 0.054565487168109654
l2 norm of weights: 2.273748128094219

Iteration Number: 9337
Loss: 10.73284985067331
l2 norm of gradients: 0.05455878508399952
l2 norm of weights: 2.2737851012769936

Iteration Number: 9338
Loss: 10.732704810360207
l2 norm of gradients: 0.054552085039405056

```

12 norm of weights: 2.2738220775447884
-----
Iteration Number: 9339
Loss: 10.732559835540298
12 norm of gradients: 0.054545387033391846
12 norm of weights: 2.2738590568939854
-----
Iteration Number: 9340
Loss: 10.732414926174355
12 norm of gradients: 0.05453869106502605
12 norm of weights: 2.2738960393209715
-----
Iteration Number: 9341
Loss: 10.732270082223168
12 norm of gradients: 0.05453199713337422
12 norm of weights: 2.2739330248221337
-----
Iteration Number: 9342
Loss: 10.73212530364756
12 norm of gradients: 0.054525305237503464
12 norm of weights: 2.273970013393864
-----
Iteration Number: 9343
Loss: 10.73198059040838
12 norm of gradients: 0.05451861537648133
12 norm of weights: 2.2740070050325567
-----
Iteration Number: 9344
Loss: 10.731835942466502
12 norm of gradients: 0.05451192754937582
12 norm of weights: 2.2740439997346074
-----
Iteration Number: 9345
Loss: 10.731691359782822
12 norm of gradients: 0.0545052417552555
12 norm of weights: 2.274080997496416
-----
Iteration Number: 9346
Loss: 10.731546842318268
12 norm of gradients: 0.054498557993189325
12 norm of weights: 2.2741179983143853
-----
Iteration Number: 9347
Loss: 10.731402390033793
12 norm of gradients: 0.05449187626224675
12 norm of weights: 2.2741550021849206
-----
Iteration Number: 9348

```

Loss: 10.731258002890378
l2 norm of gradients: 0.054485196561497765
l2 norm of weights: 2.274192009104429

Iteration Number: 9349
Loss: 10.731113680849026
l2 norm of gradients: 0.054478518890012734
l2 norm of weights: 2.2742290190693226

Iteration Number: 9350
Loss: 10.730969423870768
l2 norm of gradients: 0.054471843246862606
l2 norm of weights: 2.2742660320760137

Iteration Number: 9351
Loss: 10.73082523191666
l2 norm of gradients: 0.05446516963111872
l2 norm of weights: 2.27430304812092

Iteration Number: 9352
Loss: 10.730681104947791
l2 norm of gradients: 0.054458498041852946
l2 norm of weights: 2.27434006720046

Iteration Number: 9353
Loss: 10.730537042925265
l2 norm of gradients: 0.054451828478137614
l2 norm of weights: 2.2743770893110553

Iteration Number: 9354
Loss: 10.73039304581022
l2 norm of gradients: 0.05444516093904548
l2 norm of weights: 2.274414114449131

Iteration Number: 9355
Loss: 10.730249113563824
l2 norm of gradients: 0.05443849542364991
l2 norm of weights: 2.2744511426111154

Iteration Number: 9356
Loss: 10.730105246147252
l2 norm of gradients: 0.05443183193102456
l2 norm of weights: 2.2744881737934377

Iteration Number: 9357
Loss: 10.729961443521733
l2 norm of gradients: 0.05442517046024372
l2 norm of weights: 2.274525207992532

```

-----
Iteration Number: 9358
Loss: 10.729817705648498
l2 norm of gradients: 0.05441851101038207
l2 norm of weights: 2.274562245204834
-----
Iteration Number: 9359
Loss: 10.729674032488818
l2 norm of gradients: 0.054411853580514834
l2 norm of weights: 2.2745992854267825
-----
Iteration Number: 9360
Loss: 10.729530424003983
l2 norm of gradients: 0.054405198169717575
l2 norm of weights: 2.2746363286548186
-----
Iteration Number: 9361
Loss: 10.72938688015531
l2 norm of gradients: 0.0543985447770665
l2 norm of weights: 2.2746733748853867
-----
Iteration Number: 9362
Loss: 10.72924340090415
l2 norm of gradients: 0.054391893401638185
l2 norm of weights: 2.274710424114934
-----
Iteration Number: 9363
Loss: 10.729099986211871
l2 norm of gradients: 0.054385244042509676
l2 norm of weights: 2.2747474763399107
-----
Iteration Number: 9364
Loss: 10.728956636039866
l2 norm of gradients: 0.05437859669875854
l2 norm of weights: 2.2747845315567683
-----
Iteration Number: 9365
Loss: 10.728813350349558
l2 norm of gradients: 0.05437195136946283
l2 norm of weights: 2.2748215897619635
-----
Iteration Number: 9366
Loss: 10.7286701291024
l2 norm of gradients: 0.054365308053700975
l2 norm of weights: 2.2748586509519533
-----
Iteration Number: 9367
Loss: 10.72852697225986

```

```

12 norm of gradients: 0.05435866675055202
12 norm of weights: 2.2748957151231997
-----
Iteration Number: 9368
Loss: 10.728383879783449
12 norm of gradients: 0.05435202745909532
12 norm of weights: 2.2749327822721654
-----
Iteration Number: 9369
Loss: 10.72824085163468
12 norm of gradients: 0.05434539017841079
12 norm of weights: 2.274969852395318
-----
Iteration Number: 9370
Loss: 10.728097887775116
12 norm of gradients: 0.05433875490757888
12 norm of weights: 2.275006925489125
-----
Iteration Number: 9371
Loss: 10.727954988166324
12 norm of gradients: 0.05433212164568036
12 norm of weights: 2.27504400155006
-----
Iteration Number: 9372
Loss: 10.727812152769916
12 norm of gradients: 0.054325490391796606
12 norm of weights: 2.275081080574597
-----
Iteration Number: 9373
Loss: 10.72766938154752
12 norm of gradients: 0.05431886114500938
12 norm of weights: 2.2751181625592127
-----
Iteration Number: 9374
Loss: 10.727526674460792
12 norm of gradients: 0.054312233904400986
12 norm of weights: 2.275155247500389
-----
Iteration Number: 9375
Loss: 10.727384031471408
12 norm of gradients: 0.05430560866905412
12 norm of weights: 2.2751923353946073
-----
Iteration Number: 9376
Loss: 10.72724145254108
12 norm of gradients: 0.054298985438052026
12 norm of weights: 2.2752294262383543
-----

```


Iteration Number: 9377
Loss: 10.72709893763154
l2 norm of gradients: 0.05429236421047832
l2 norm of weights: 2.2752665200281177

Iteration Number: 9378
Loss: 10.726956486704545
l2 norm of gradients: 0.05428574498541716
l2 norm of weights: 2.2753036167603895

Iteration Number: 9379
Loss: 10.726814099721878
l2 norm of gradients: 0.05427912776195321
l2 norm of weights: 2.275340716431663

Iteration Number: 9380
Loss: 10.72667177664535
l2 norm of gradients: 0.05427251253917147
l2 norm of weights: 2.2753778190384346

Iteration Number: 9381
Loss: 10.726529517436798
l2 norm of gradients: 0.054265899316157565
l2 norm of weights: 2.2754149245772046

Iteration Number: 9382
Loss: 10.726387322058077
l2 norm of gradients: 0.05425928809199746
l2 norm of weights: 2.2754520330444743

Iteration Number: 9383
Loss: 10.726245190471083
l2 norm of gradients: 0.05425267886577769
l2 norm of weights: 2.2754891444367495

Iteration Number: 9384
Loss: 10.72610312263772
l2 norm of gradients: 0.05424607163658519
l2 norm of weights: 2.2755262587505367

Iteration Number: 9385
Loss: 10.725961118519933
l2 norm of gradients: 0.054239466403507346
l2 norm of weights: 2.2755633759823475

Iteration Number: 9386
Loss: 10.725819178079675
l2 norm of gradients: 0.0542328631656321

```

12 norm of weights: 2.2756004961286935
-----
Iteration Number: 9387
Loss: 10.725677301278944
12 norm of gradients: 0.054226261922047776
12 norm of weights: 2.2756376191860914
-----
Iteration Number: 9388
Loss: 10.725535488079753
12 norm of gradients: 0.05421966267184321
12 norm of weights: 2.2756747451510595
-----
Iteration Number: 9389
Loss: 10.72539373844414
12 norm of gradients: 0.05421306541410772
12 norm of weights: 2.2757118740201188
-----
Iteration Number: 9390
Loss: 10.725252052334174
12 norm of gradients: 0.054206470147931
12 norm of weights: 2.2757490057897933
-----
Iteration Number: 9391
Loss: 10.72511042971194
12 norm of gradients: 0.05419987687240331
12 norm of weights: 2.27578614045661
-----
Iteration Number: 9392
Loss: 10.72496887053956
12 norm of gradients: 0.05419328558661534
12 norm of weights: 2.275823278017098
-----
Iteration Number: 9393
Loss: 10.724827374779174
12 norm of gradients: 0.05418669628965823
12 norm of weights: 2.2758604184677895
-----
Iteration Number: 9394
Loss: 10.724685942392949
12 norm of gradients: 0.05418010898062362
12 norm of weights: 2.275897561805219
-----
Iteration Number: 9395
Loss: 10.724544573343074
12 norm of gradients: 0.05417352365860356
12 norm of weights: 2.2759347080259236
-----
Iteration Number: 9396

```

```

Loss: 10.724403267591777
l2 norm of gradients: 0.054166940322690676
l2 norm of weights: 2.2759718571264442
-----
Iteration Number: 9397
Loss: 10.724262025101295
l2 norm of gradients: 0.05416035897197786
l2 norm of weights: 2.2760090091033236
-----
Iteration Number: 9398
Loss: 10.724120845833898
l2 norm of gradients: 0.054153779605558706
l2 norm of weights: 2.2760461639531075
-----
Iteration Number: 9399
Loss: 10.723979729751878
l2 norm of gradients: 0.05414720222252708
l2 norm of weights: 2.2760833216723433
-----
Iteration Number: 9400
Loss: 10.723838676817559
l2 norm of gradients: 0.05414062682197739
l2 norm of weights: 2.276120482257583
-----
Iteration Number: 9401
Loss: 10.723697686993287
l2 norm of gradients: 0.05413405340300459
l2 norm of weights: 2.27615764570538
-----
Iteration Number: 9402
Loss: 10.723556760241422
l2 norm of gradients: 0.05412748196470393
l2 norm of weights: 2.27619481201229
-----
Iteration Number: 9403
Loss: 10.723415896524374
l2 norm of gradients: 0.05412091250617121
l2 norm of weights: 2.276231981174873
-----
Iteration Number: 9404
Loss: 10.723275095804556
l2 norm of gradients: 0.054114345026502746
l2 norm of weights: 2.27626915318969
-----
Iteration Number: 9405
Loss: 10.723134358044412
l2 norm of gradients: 0.054107779524795216
l2 norm of weights: 2.2763063280533062

```

```
-----  
Iteration Number: 9406  
Loss: 10.722993683206415  
l2 norm of gradients: 0.05410121600014578  
l2 norm of weights: 2.276343505762288  
-----  
Iteration Number: 9407  
Loss: 10.722853071253068  
l2 norm of gradients: 0.0540946544516521  
l2 norm of weights: 2.2763806863132054  
-----  
Iteration Number: 9408  
Loss: 10.72271252214688  
l2 norm of gradients: 0.05408809487841231  
l2 norm of weights: 2.2764178697026307  
-----  
Iteration Number: 9409  
Loss: 10.722572035850412  
l2 norm of gradients: 0.05408153727952495  
l2 norm of weights: 2.2764550559271397  
-----  
Iteration Number: 9410  
Loss: 10.722431612326226  
l2 norm of gradients: 0.054074981654089097  
l2 norm of weights: 2.2764922449833094  
-----  
Iteration Number: 9411  
Loss: 10.72229125153692  
l2 norm of gradients: 0.05406842800120418  
l2 norm of weights: 2.2765294368677207  
-----  
Iteration Number: 9412  
Loss: 10.722150953445123  
l2 norm of gradients: 0.05406187631997018  
l2 norm of weights: 2.276566631576957  
-----  
Iteration Number: 9413  
Loss: 10.722010718013477  
l2 norm of gradients: 0.054055326609487496  
l2 norm of weights: 2.276603829107603  
-----  
Iteration Number: 9414  
Loss: 10.721870545204656  
l2 norm of gradients: 0.054048778868857034  
l2 norm of weights: 2.2766410294562487  
-----  
Iteration Number: 9415  
Loss: 10.721730434981357
```

l2 norm of gradients: 0.054042233097180085
l2 norm of weights: 2.2766782326194845

Iteration Number: 9416
Loss: 10.721590387306298
l2 norm of gradients: 0.054035689293558446
l2 norm of weights: 2.2767154385939046

Iteration Number: 9417
Loss: 10.721450402142239
l2 norm of gradients: 0.0540291474570944
l2 norm of weights: 2.276752647376105

Iteration Number: 9418
Loss: 10.72131047945194
l2 norm of gradients: 0.05402260758689059
l2 norm of weights: 2.276789858962685

Iteration Number: 9419
Loss: 10.721170619198205
l2 norm of gradients: 0.05401606968205027
l2 norm of weights: 2.2768270733502463

Iteration Number: 9420
Loss: 10.721030821343854
l2 norm of gradients: 0.054009533741677
l2 norm of weights: 2.2768642905353937

Iteration Number: 9421
Loss: 10.72089108585174
l2 norm of gradients: 0.05400299976487489
l2 norm of weights: 2.276901510514734

Iteration Number: 9422
Loss: 10.720751412684727
l2 norm of gradients: 0.05399646775074851
l2 norm of weights: 2.2769387332848776

Iteration Number: 9423
Loss: 10.720611801805719
l2 norm of gradients: 0.05398993769840282
l2 norm of weights: 2.2769759588424354

Iteration Number: 9424
Loss: 10.720472253177634
l2 norm of gradients: 0.0539834096069433
l2 norm of weights: 2.277013187184024

Iteration Number: 9425
Loss: 10.720332766763427
l2 norm of gradients: 0.05397688347547587
l2 norm of weights: 2.2770504183062608

Iteration Number: 9426
Loss: 10.72019334252606
l2 norm of gradients: 0.0539703593031069
l2 norm of weights: 2.2770876522057653

Iteration Number: 9427
Loss: 10.720053980428538
l2 norm of gradients: 0.05396383708894322
l2 norm of weights: 2.277124888879161

Iteration Number: 9428
Loss: 10.719914680433876
l2 norm of gradients: 0.05395731683209212
l2 norm of weights: 2.277162128323074

Iteration Number: 9429
Loss: 10.719775442505126
l2 norm of gradients: 0.053950798531661355
l2 norm of weights: 2.2771993705341322

Iteration Number: 9430
Loss: 10.71963626660536
l2 norm of gradients: 0.0539442821867591
l2 norm of weights: 2.277236615508966

Iteration Number: 9431
Loss: 10.719497152697674
l2 norm of gradients: 0.05393776779649402
l2 norm of weights: 2.2772738632442096

Iteration Number: 9432
Loss: 10.719358100745183
l2 norm of gradients: 0.053931255359975214
l2 norm of weights: 2.2773111137364985

Iteration Number: 9433
Loss: 10.71921911071104
l2 norm of gradients: 0.053924744876312276
l2 norm of weights: 2.277348366982472

Iteration Number: 9434
Loss: 10.719080182558411
l2 norm of gradients: 0.05391823634461522

```

12 norm of weights: 2.2773856229787715
-----
Iteration Number: 9435
Loss: 10.718941316250493
12 norm of gradients: 0.05391172976399453
12 norm of weights: 2.2774228817220408
-----
Iteration Number: 9436
Loss: 10.718802511750509
12 norm of gradients: 0.05390522513356112
12 norm of weights: 2.2774601432089265
-----
Iteration Number: 9437
Loss: 10.718663769021699
12 norm of gradients: 0.05389872245242637
12 norm of weights: 2.277497407436078
-----
Iteration Number: 9438
Loss: 10.718525088027336
12 norm of gradients: 0.053892221719702156
12 norm of weights: 2.2775346744001475
-----
Iteration Number: 9439
Loss: 10.718386468730712
12 norm of gradients: 0.05388572293450075
12 norm of weights: 2.277571944097789
-----
Iteration Number: 9440
Loss: 10.718247911095144
12 norm of gradients: 0.0538792260959349
12 norm of weights: 2.277609216525659
-----
Iteration Number: 9441
Loss: 10.71810941508398
12 norm of gradients: 0.053872731203117794
12 norm of weights: 2.2776464916804193
-----
Iteration Number: 9442
Loss: 10.717970980660585
12 norm of gradients: 0.05386623825516312
12 norm of weights: 2.27768376955873
-----
Iteration Number: 9443
Loss: 10.71783260778835
12 norm of gradients: 0.053859747251184954
12 norm of weights: 2.2777210501572576
-----
Iteration Number: 9444

```

Loss: 10.717694296430697
l2 norm of gradients: 0.05385325819029791
l2 norm of weights: 2.2777583334726685

Iteration Number: 9445
Loss: 10.717556046551062
l2 norm of gradients: 0.05384677107161692
l2 norm of weights: 2.277795619501634

Iteration Number: 9446
Loss: 10.717417858112917
l2 norm of gradients: 0.053840285894257496
l2 norm of weights: 2.2778329082408253

Iteration Number: 9447
Loss: 10.717279731079746
l2 norm of gradients: 0.053833802657335554
l2 norm of weights: 2.277870199686919

Iteration Number: 9448
Loss: 10.71714166541507
l2 norm of gradients: 0.05382732135996744
l2 norm of weights: 2.277907493836593

Iteration Number: 9449
Loss: 10.717003661082426
l2 norm of gradients: 0.05382084200127001
l2 norm of weights: 2.2779447906865276

Iteration Number: 9450
Loss: 10.716865718045378
l2 norm of gradients: 0.05381436458036051
l2 norm of weights: 2.277982090233406

Iteration Number: 9451
Loss: 10.716727836267516
l2 norm of gradients: 0.05380788909635665
l2 norm of weights: 2.2780193924739134

Iteration Number: 9452
Loss: 10.716590015712454
l2 norm of gradients: 0.0538014155483766
l2 norm of weights: 2.2780566974047387

Iteration Number: 9453
Loss: 10.716452256343825
l2 norm of gradients: 0.053794943935539054
l2 norm of weights: 2.278094005022573


```

-----
Iteration Number: 9454
Loss: 10.716314558125292
l2 norm of gradients: 0.053788474256962995
l2 norm of weights: 2.278131315324109
-----
Iteration Number: 9455
Loss: 10.716176921020546
l2 norm of gradients: 0.05378200651176797
l2 norm of weights: 2.2781686283060436
-----
Iteration Number: 9456
Loss: 10.716039344993293
l2 norm of gradients: 0.05377554069907399
l2 norm of weights: 2.2782059439650753
-----
Iteration Number: 9457
Loss: 10.715901830007267
l2 norm of gradients: 0.05376907681800143
l2 norm of weights: 2.278243262297905
-----
Iteration Number: 9458
Loss: 10.715764376026234
l2 norm of gradients: 0.05376261486767119
l2 norm of weights: 2.278280583301236
-----
Iteration Number: 9459
Loss: 10.71562698301397
l2 norm of gradients: 0.053756154847204606
l2 norm of weights: 2.2783179069717763
-----
Iteration Number: 9460
Loss: 10.715489650934288
l2 norm of gradients: 0.0537496967557234
l2 norm of weights: 2.2783552333062334
-----
Iteration Number: 9461
Loss: 10.71535237975102
l2 norm of gradients: 0.0537432405923498
l2 norm of weights: 2.2783925623013195
-----
Iteration Number: 9462
Loss: 10.715215169428017
l2 norm of gradients: 0.0537367863562065
l2 norm of weights: 2.2784298939537484
-----
Iteration Number: 9463
Loss: 10.715078019929168

```

```

12 norm of gradients: 0.0537303340464166
12 norm of weights: 2.278467228260237
-----
Iteration Number: 9464
Loss: 10.714940931218369
12 norm of gradients: 0.05372388366210366
12 norm of weights: 2.2785045652175047
-----
Iteration Number: 9465
Loss: 10.714803903259556
12 norm of gradients: 0.053717435202391674
12 norm of weights: 2.2785419048222724
-----
Iteration Number: 9466
Loss: 10.71466693601668
12 norm of gradients: 0.0537109886664051
12 norm of weights: 2.2785792470712654
-----
Iteration Number: 9467
Loss: 10.71453002945372
12 norm of gradients: 0.05370454405326887
12 norm of weights: 2.2786165919612107
-----
Iteration Number: 9468
Loss: 10.714393183534675
12 norm of gradients: 0.05369810136210833
12 norm of weights: 2.278653939488837
-----
Iteration Number: 9469
Loss: 10.71425639822357
12 norm of gradients: 0.05369166059204921
12 norm of weights: 2.278691289650877
-----
Iteration Number: 9470
Loss: 10.714119673484461
12 norm of gradients: 0.05368522174221784
12 norm of weights: 2.278728642444065
-----
Iteration Number: 9471
Loss: 10.713983009281417
12 norm of gradients: 0.05367878481174088
12 norm of weights: 2.2787659978651384
-----
Iteration Number: 9472
Loss: 10.713846405578536
12 norm of gradients: 0.05367234979974544
12 norm of weights: 2.2788033559108363
-----

```

Iteration Number: 9473
Loss: 10.713709862339943
l2 norm of gradients: 0.05366591670535912
l2 norm of weights: 2.2788407165779017

Iteration Number: 9474
Loss: 10.713573379529787
l2 norm of gradients: 0.05365948552770993
l2 norm of weights: 2.278878079863079

Iteration Number: 9475
Loss: 10.71343695711223
l2 norm of gradients: 0.053653056265926334
l2 norm of weights: 2.2789154457631153

Iteration Number: 9476
Loss: 10.713300595051471
l2 norm of gradients: 0.05364662891913728
l2 norm of weights: 2.2789528142747613

Iteration Number: 9477
Loss: 10.713164293311731
l2 norm of gradients: 0.053640203486472124
l2 norm of weights: 2.2789901853947683

Iteration Number: 9478
Loss: 10.713028051857249
l2 norm of gradients: 0.053633779967060646
l2 norm of weights: 2.279027559119892

Iteration Number: 9479
Loss: 10.712891870652289
l2 norm of gradients: 0.05362735836003311
l2 norm of weights: 2.2790649354468893

Iteration Number: 9480
Loss: 10.712755749661149
l2 norm of gradients: 0.053620938664520175
l2 norm of weights: 2.279102314372521

Iteration Number: 9481
Loss: 10.712619688848138
l2 norm of gradients: 0.05361452087965303
l2 norm of weights: 2.279139695893549

Iteration Number: 9482
Loss: 10.712483688177596
l2 norm of gradients: 0.05360810500456323

```

12 norm of weights: 2.279177080006739
-----
Iteration Number: 9483
Loss: 10.712347747613883
12 norm of gradients: 0.053601691038382776
12 norm of weights: 2.2792144667088574
-----
Iteration Number: 9484
Loss: 10.712211867121386
12 norm of gradients: 0.05359527898024417
12 norm of weights: 2.2792518559966752
-----
Iteration Number: 9485
Loss: 10.712076046664519
12 norm of gradients: 0.05358886882928028
12 norm of weights: 2.2792892478669655
-----
Iteration Number: 9486
Loss: 10.71194028620771
12 norm of gradients: 0.05358246058462449
12 norm of weights: 2.2793266423165033
-----
Iteration Number: 9487
Loss: 10.711804585715424
12 norm of gradients: 0.05357605424541056
12 norm of weights: 2.2793640393420653
-----
Iteration Number: 9488
Loss: 10.711668945152134
12 norm of gradients: 0.05356964981077276
12 norm of weights: 2.279401438940433
-----
Iteration Number: 9489
Loss: 10.71153336448235
12 norm of gradients: 0.053563247279845724
12 norm of weights: 2.279438841108388
-----
Iteration Number: 9490
Loss: 10.711397843670598
12 norm of gradients: 0.05355684665176458
12 norm of weights: 2.2794762458427162
-----
Iteration Number: 9491
Loss: 10.711262382681436
12 norm of gradients: 0.0535504479256649
12 norm of weights: 2.2795136531402056
-----
Iteration Number: 9492

```

```

Loss: 10.711126981479435
l2 norm of gradients: 0.05354405110068269
l2 norm of weights: 2.279551062997646
-----
Iteration Number: 9493
Loss: 10.710991640029203
l2 norm of gradients: 0.05353765617595439
l2 norm of weights: 2.2795884754118303
-----
Iteration Number: 9494
Loss: 10.710856358295356
l2 norm of gradients: 0.05353126315061686
l2 norm of weights: 2.279625890379554
-----
Iteration Number: 9495
Loss: 10.710721136242546
l2 norm of gradients: 0.053524872023807415
l2 norm of weights: 2.2796633078976143
-----
Iteration Number: 9496
Loss: 10.710585973835444
l2 norm of gradients: 0.05351848279466384
l2 norm of weights: 2.279700727962812
-----
Iteration Number: 9497
Loss: 10.710450871038745
l2 norm of gradients: 0.053512095462324334
l2 norm of weights: 2.2797381505719496
-----
Iteration Number: 9498
Loss: 10.71031582781717
l2 norm of gradients: 0.05350571002592749
l2 norm of weights: 2.279775575721833
-----
Iteration Number: 9499
Loss: 10.71018084413546
l2 norm of gradients: 0.05349932648461246
l2 norm of weights: 2.2798130034092696
-----
Iteration Number: 9500
Loss: 10.71004591995838
l2 norm of gradients: 0.053492944837518694
l2 norm of weights: 2.2798504336310694
-----
Iteration Number: 9501
Loss: 10.709911055250725
l2 norm of gradients: 0.05348656508378619
l2 norm of weights: 2.279887866384046

```

```

-----
Iteration Number: 9502
Loss: 10.709776249977299
l2 norm of gradients: 0.05348018722255533
l2 norm of weights: 2.2799253016650134
-----
Iteration Number: 9503
Loss: 10.709641504102947
l2 norm of gradients: 0.05347381125296698
l2 norm of weights: 2.2799627394707906
-----
Iteration Number: 9504
Loss: 10.70950681759253
l2 norm of gradients: 0.053467437174162344
l2 norm of weights: 2.280000179798197
-----
Iteration Number: 9505
Loss: 10.709372190410921
l2 norm of gradients: 0.05346106498528318
l2 norm of weights: 2.280037622644056
-----
Iteration Number: 9506
Loss: 10.709237622523045
l2 norm of gradients: 0.053454694685471635
l2 norm of weights: 2.280075068005192
-----
Iteration Number: 9507
Loss: 10.70910311389382
l2 norm of gradients: 0.053448326273870264
l2 norm of weights: 2.2801125158784337
-----
Iteration Number: 9508
Loss: 10.708968664488207
l2 norm of gradients: 0.05344195974962214
l2 norm of weights: 2.280149966260611
-----
Iteration Number: 9509
Loss: 10.708834274271181
l2 norm of gradients: 0.05343559511187066
l2 norm of weights: 2.280187419148556
-----
Iteration Number: 9510
Loss: 10.708699943207746
l2 norm of gradients: 0.053429232359759764
l2 norm of weights: 2.280224874539104
-----
Iteration Number: 9511
Loss: 10.708565671262924

```

```

12 norm of gradients: 0.05342287149243377
12 norm of weights: 2.2802623324290927
-----
Iteration Number: 9512
Loss: 10.708431458401767
12 norm of gradients: 0.05341651250903745
12 norm of weights: 2.2802997928153625
-----
Iteration Number: 9513
Loss: 10.708297304589347
12 norm of gradients: 0.053410155408715985
12 norm of weights: 2.2803372556947554
-----
Iteration Number: 9514
Loss: 10.708163209790758
12 norm of gradients: 0.05340380019061503
12 norm of weights: 2.280374721064117
-----
Iteration Number: 9515
Loss: 10.708029173971118
12 norm of gradients: 0.05339744685388069
12 norm of weights: 2.2804121889202946
-----
Iteration Number: 9516
Loss: 10.70789519709557
12 norm of gradients: 0.053391095397659405
12 norm of weights: 2.280449659260138
-----
Iteration Number: 9517
Loss: 10.707761279129283
12 norm of gradients: 0.05338474582109818
12 norm of weights: 2.2804871320804994
-----
Iteration Number: 9518
Loss: 10.707627420037438
12 norm of gradients: 0.05337839812334439
12 norm of weights: 2.280524607378234
-----
Iteration Number: 9519
Loss: 10.707493619785257
12 norm of gradients: 0.05337205230354586
12 norm of weights: 2.2805620851501995
-----
Iteration Number: 9520
Loss: 10.70735987833797
12 norm of gradients: 0.053365708360850785
12 norm of weights: 2.280599565393255
-----

```

Iteration Number: 9521
Loss: 10.707226195660834
l2 norm of gradients: 0.053359366294407864
l2 norm of weights: 2.2806370481042624

Iteration Number: 9522
Loss: 10.707092571719132
l2 norm of gradients: 0.05335302610336628
l2 norm of weights: 2.2806745332800875

Iteration Number: 9523
Loss: 10.706959006478172
l2 norm of gradients: 0.05334668778687552
l2 norm of weights: 2.280712020917597

Iteration Number: 9524
Loss: 10.70682549990328
l2 norm of gradients: 0.05334035134408559
l2 norm of weights: 2.28074951101366

Iteration Number: 9525
Loss: 10.706692051959813
l2 norm of gradients: 0.053334016774146877
l2 norm of weights: 2.280787003565149

Iteration Number: 9526
Loss: 10.70655866261314
l2 norm of gradients: 0.05332768407621026
l2 norm of weights: 2.280824498568939

Iteration Number: 9527
Loss: 10.706425331828662
l2 norm of gradients: 0.053321353249427016
l2 norm of weights: 2.2808619960219056

Iteration Number: 9528
Loss: 10.7062920595718
l2 norm of gradients: 0.053315024292948855
l2 norm of weights: 2.2808994959209294

Iteration Number: 9529
Loss: 10.706158845807996
l2 norm of gradients: 0.05330869720592796
l2 norm of weights: 2.280936998262892

Iteration Number: 9530
Loss: 10.706025690502724
l2 norm of gradients: 0.05330237198751684

l2 norm of weights: 2.2809745030446766

Iteration Number: 9531
Loss: 10.70589259362147
l2 norm of gradients: 0.05329604863686856
l2 norm of weights: 2.2810120102631712

Iteration Number: 9532
Loss: 10.705759555129749
l2 norm of gradients: 0.053289727153136526
l2 norm of weights: 2.2810495199152645

Iteration Number: 9533
Loss: 10.7056265749931
l2 norm of gradients: 0.05328340753547467
l2 norm of weights: 2.281087031997848

Iteration Number: 9534
Loss: 10.70549365317708
l2 norm of gradients: 0.05327708978303729
l2 norm of weights: 2.281124546507815

Iteration Number: 9535
Loss: 10.705360789647273
l2 norm of gradients: 0.05327077389497905
l2 norm of weights: 2.2811620634420633

Iteration Number: 9536
Loss: 10.705227984369287
l2 norm of gradients: 0.05326445987045518
l2 norm of weights: 2.2811995827974907

Iteration Number: 9537
Loss: 10.705095237308749
l2 norm of gradients: 0.053258147708621244
l2 norm of weights: 2.2812371045709985

Iteration Number: 9538
Loss: 10.704962548431313
l2 norm of gradients: 0.05325183740863329
l2 norm of weights: 2.281274628759491

Iteration Number: 9539
Loss: 10.704829917702657
l2 norm of gradients: 0.053245528969647786
l2 norm of weights: 2.2813121553598736

Iteration Number: 9540

Loss: 10.704697345088475
l2 norm of gradients: 0.05323922239082158
l2 norm of weights: 2.2813496843690553

Iteration Number: 9541
Loss: 10.704564830554489
l2 norm of gradients: 0.05323291767131205
l2 norm of weights: 2.2813872157839468

Iteration Number: 9542
Loss: 10.704432374066442
l2 norm of gradients: 0.053226614810276865
l2 norm of weights: 2.2814247496014612

Iteration Number: 9543
Loss: 10.704299975590107
l2 norm of gradients: 0.05322031380687422
l2 norm of weights: 2.2814622858185154

Iteration Number: 9544
Loss: 10.704167635091268
l2 norm of gradients: 0.0532140146602628
l2 norm of weights: 2.281499824432026

Iteration Number: 9545
Loss: 10.704035352535742
l2 norm of gradients: 0.05320771736960153
l2 norm of weights: 2.2815373654389144

Iteration Number: 9546
Loss: 10.703903127889362
l2 norm of gradients: 0.0532014219340499
l2 norm of weights: 2.281574908836104

Iteration Number: 9547
Loss: 10.70377096111799
l2 norm of gradients: 0.05319512835276784
l2 norm of weights: 2.28161245462052

Iteration Number: 9548
Loss: 10.703638852187506
l2 norm of gradients: 0.0531888366249156
l2 norm of weights: 2.281650002789089

Iteration Number: 9549
Loss: 10.703506801063813
l2 norm of gradients: 0.053182546749653996
l2 norm of weights: 2.281687553338743

```
-----  
Iteration Number: 9550  
Loss: 10.70337480771284  
l2 norm of gradients: 0.053176258726144135  
l2 norm of weights: 2.2817251062664132  
-----  
Iteration Number: 9551  
Loss: 10.70324287210054  
l2 norm of gradients: 0.05316997255354767  
l2 norm of weights: 2.2817626615690356  
-----  
Iteration Number: 9552  
Loss: 10.703110994192881  
l2 norm of gradients: 0.05316368823102659  
l2 norm of weights: 2.281800219243547  
-----  
Iteration Number: 9553  
Loss: 10.70297917395586  
l2 norm of gradients: 0.053157405757743324  
l2 norm of weights: 2.2818377792868874  
-----  
Iteration Number: 9554  
Loss: 10.7028474113555  
l2 norm of gradients: 0.05315112513286083  
l2 norm of weights: 2.2818753416959994  
-----  
Iteration Number: 9555  
Loss: 10.702715706357834  
l2 norm of gradients: 0.05314484635554233  
l2 norm of weights: 2.2819129064678267  
-----  
Iteration Number: 9556  
Loss: 10.702584058928931  
l2 norm of gradients: 0.05313856942495159  
l2 norm of weights: 2.2819504735993164  
-----  
Iteration Number: 9557  
Loss: 10.70245246903488  
l2 norm of gradients: 0.05313229434025274  
l2 norm of weights: 2.281988043087419  
-----  
Iteration Number: 9558  
Loss: 10.702320936641788  
l2 norm of gradients: 0.053126021100610446  
l2 norm of weights: 2.282025614929085  
-----  
Iteration Number: 9559  
Loss: 10.702189461715783
```

```

12 norm of gradients: 0.05311974970518963
12 norm of weights: 2.282063189121269
-----
Iteration Number: 9560
Loss: 10.702058044223028
12 norm of gradients: 0.05311348015315577
12 norm of weights: 2.2821007656609273
-----
Iteration Number: 9561
Loss: 10.701926684129694
12 norm of gradients: 0.05310721244367466
12 norm of weights: 2.282138344545019
-----
Iteration Number: 9562
Loss: 10.701795381401986
12 norm of gradients: 0.05310094657591265
12 norm of weights: 2.282175925770505
-----
Iteration Number: 9563
Loss: 10.701664136006121
12 norm of gradients: 0.05309468254903641
12 norm of weights: 2.28221350933435
-----
Iteration Number: 9564
Loss: 10.701532947908348
12 norm of gradients: 0.05308842036221309
12 norm of weights: 2.282251095233519
-----
Iteration Number: 9565
Loss: 10.701401817074935
12 norm of gradients: 0.053082160014610226
12 norm of weights: 2.28228868346498
-----
Iteration Number: 9566
Loss: 10.701270743472172
12 norm of gradients: 0.053075901505395824
12 norm of weights: 2.282326274025705
-----
Iteration Number: 9567
Loss: 10.701139727066371
12 norm of gradients: 0.05306964483373827
12 norm of weights: 2.2823638669126667
-----
Iteration Number: 9568
Loss: 10.70100876782387
12 norm of gradients: 0.053063389998806394
12 norm of weights: 2.28240146212284
-----

```

Iteration Number: 9569
Loss: 10.700877865711025
l2 norm of gradients: 0.053057136999769404
l2 norm of weights: 2.282439059653204

Iteration Number: 9570
Loss: 10.700747020694216
l2 norm of gradients: 0.05305088583579704
l2 norm of weights: 2.2824766595007375

Iteration Number: 9571
Loss: 10.700616232739849
l2 norm of gradients: 0.053044636506059326
l2 norm of weights: 2.282514261662423

Iteration Number: 9572
Loss: 10.700485501814347
l2 norm of gradients: 0.053038389009726815
l2 norm of weights: 2.2825518661352473

Iteration Number: 9573
Loss: 10.700354827884158
l2 norm of gradients: 0.05303214334597045
l2 norm of weights: 2.282589472916196

Iteration Number: 9574
Loss: 10.700224210915758
l2 norm of gradients: 0.053025899513961576
l2 norm of weights: 2.28262708200226

Iteration Number: 9575
Loss: 10.700093650875633
l2 norm of gradients: 0.05301965751287199
l2 norm of weights: 2.28266469339043

Iteration Number: 9576
Loss: 10.699963147730301
l2 norm of gradients: 0.05301341734187386
l2 norm of weights: 2.282702307077701

Iteration Number: 9577
Loss: 10.6998327014463
l2 norm of gradients: 0.053007179000139897
l2 norm of weights: 2.2827399230610697

Iteration Number: 9578
Loss: 10.699702311990192
l2 norm of gradients: 0.05300094248684304

l2 norm of weights: 2.282777541337535

Iteration Number: 9579
Loss: 10.699571979328557
l2 norm of gradients: 0.052994707801156796
l2 norm of weights: 2.282815161904099

Iteration Number: 9580
Loss: 10.699441703427999
l2 norm of gradients: 0.052988474942255066
l2 norm of weights: 2.2828527847577647

Iteration Number: 9581
Loss: 10.699311484255151
l2 norm of gradients: 0.052982243909312166
l2 norm of weights: 2.282890409895538

Iteration Number: 9582
Loss: 10.699181321776653
l2 norm of gradients: 0.0529760147015028
l2 norm of weights: 2.2829280373144285

Iteration Number: 9583
Loss: 10.699051215959187
l2 norm of gradients: 0.052969787318002094
l2 norm of weights: 2.2829656670114455

Iteration Number: 9584
Loss: 10.698921166769443
l2 norm of gradients: 0.05296356175798566
l2 norm of weights: 2.2830032989836035

Iteration Number: 9585
Loss: 10.698791174174136
l2 norm of gradients: 0.05295733802062947
l2 norm of weights: 2.283040933227917

Iteration Number: 9586
Loss: 10.698661238140005
l2 norm of gradients: 0.05295111610510989
l2 norm of weights: 2.283078569741404

Iteration Number: 9587
Loss: 10.698531358633815
l2 norm of gradients: 0.05294489601060383
l2 norm of weights: 2.283116208521085

Iteration Number: 9588

Loss: 10.698401535622345
l2 norm of gradients: 0.052938677736288464
l2 norm of weights: 2.2831538495639823

Iteration Number: 9589
Loss: 10.698271769072402
l2 norm of gradients: 0.05293246128134147
l2 norm of weights: 2.28319149286712

Iteration Number: 9590
Loss: 10.698142058950813
l2 norm of gradients: 0.052926246644940944
l2 norm of weights: 2.283229138427526

Iteration Number: 9591
Loss: 10.69801240522443
l2 norm of gradients: 0.05292003382626535
l2 norm of weights: 2.2832667862422293

Iteration Number: 9592
Loss: 10.697882807860125
l2 norm of gradients: 0.052913822824493674
l2 norm of weights: 2.283304436308262

Iteration Number: 9593
Loss: 10.697753266824787
l2 norm of gradients: 0.05290761363880518
l2 norm of weights: 2.2833420886226583

Iteration Number: 9594
Loss: 10.697623782085339
l2 norm of gradients: 0.05290140626837965
l2 norm of weights: 2.283379743182454

Iteration Number: 9595
Loss: 10.697494353608718
l2 norm of gradients: 0.05289520071239722
l2 norm of weights: 2.283417399984688

Iteration Number: 9596
Loss: 10.697364981361883
l2 norm of gradients: 0.052888996970038536
l2 norm of weights: 2.2834550590264016

Iteration Number: 9597
Loss: 10.697235665311815
l2 norm of gradients: 0.05288279504048453
l2 norm of weights: 2.283492720304638

```

-----
Iteration Number: 9598
Loss: 10.69710640542552
l2 norm of gradients: 0.05287659492291668
l2 norm of weights: 2.283530383816443
-----
Iteration Number: 9599
Loss: 10.696977201670027
l2 norm of gradients: 0.0528703966165168
l2 norm of weights: 2.2835680495588644
-----
Iteration Number: 9600
Loss: 10.696848054012387
l2 norm of gradients: 0.05286420012046714
l2 norm of weights: 2.283605717528952
-----
Iteration Number: 9601
Loss: 10.696718962419665
l2 norm of gradients: 0.052858005433950396
l2 norm of weights: 2.2836433877237594
-----
Iteration Number: 9602
Loss: 10.696589926858955
l2 norm of gradients: 0.0528518125561496
l2 norm of weights: 2.283681060140341
-----
Iteration Number: 9603
Loss: 10.696460947297371
l2 norm of gradients: 0.05284562148624825
l2 norm of weights: 2.2837187347757535
-----
Iteration Number: 9604
Loss: 10.696332023702059
l2 norm of gradients: 0.05283943222343032
l2 norm of weights: 2.2837564116270572
-----
Iteration Number: 9605
Loss: 10.696203156040168
l2 norm of gradients: 0.05283324476688014
l2 norm of weights: 2.283794090691313
-----
Iteration Number: 9606
Loss: 10.696074344278882
l2 norm of gradients: 0.052827059115782386
l2 norm of weights: 2.2838317719655863
-----
Iteration Number: 9607
Loss: 10.695945588385406

```


l2 norm of gradients: 0.052820875269322234
l2 norm of weights: 2.283869455446942

Iteration Number: 9608
Loss: 10.695816888326961
l2 norm of gradients: 0.052814693226685284
l2 norm of weights: 2.2839071411324503

Iteration Number: 9609
Loss: 10.695688244070794
l2 norm of gradients: 0.052808512987057535
l2 norm of weights: 2.283944829019181

Iteration Number: 9610
Loss: 10.695559655584178
l2 norm of gradients: 0.052802334549625386
l2 norm of weights: 2.283982519104207

Iteration Number: 9611
Loss: 10.695431122834398
l2 norm of gradients: 0.0527961579135756
l2 norm of weights: 2.2840202113846053

Iteration Number: 9612
Loss: 10.695302645788773
l2 norm of gradients: 0.05278998307809546
l2 norm of weights: 2.2840579058574524

Iteration Number: 9613
Loss: 10.69517422441463
l2 norm of gradients: 0.05278381004237258
l2 norm of weights: 2.28409560251983

Iteration Number: 9614
Loss: 10.695045858679329
l2 norm of gradients: 0.052777638805595016
l2 norm of weights: 2.284133301368819

Iteration Number: 9615
Loss: 10.694917548550245
l2 norm of gradients: 0.05277146936695121
l2 norm of weights: 2.2841710024015045

Iteration Number: 9616
Loss: 10.694789293994784
l2 norm of gradients: 0.05276530172563014
l2 norm of weights: 2.284208705614974

Iteration Number: 9617
Loss: 10.69466109498036
l2 norm of gradients: 0.05275913588082101
l2 norm of weights: 2.284246411006316

Iteration Number: 9618
Loss: 10.694532951474422
l2 norm of gradients: 0.05275297183171355
l2 norm of weights: 2.284284118572623

Iteration Number: 9619
Loss: 10.69440486344443
l2 norm of gradients: 0.052746809577497854
l2 norm of weights: 2.284321828310988

Iteration Number: 9620
Loss: 10.694276830857874
l2 norm of gradients: 0.05274064911736448
l2 norm of weights: 2.2843595402185066

Iteration Number: 9621
Loss: 10.694148853682261
l2 norm of gradients: 0.05273449045050436
l2 norm of weights: 2.2843972542922786

Iteration Number: 9622
Loss: 10.69402093188512
l2 norm of gradients: 0.05272833357610888
l2 norm of weights: 2.2844349705294036

Iteration Number: 9623
Loss: 10.693893065434008
l2 norm of gradients: 0.05272217849336974
l2 norm of weights: 2.284472688926985

Iteration Number: 9624
Loss: 10.693765254296489
l2 norm of gradients: 0.052716025201479125
l2 norm of weights: 2.2845104094821274

Iteration Number: 9625
Loss: 10.693637498440172
l2 norm of gradients: 0.052709873699629685
l2 norm of weights: 2.2845481321919388

Iteration Number: 9626
Loss: 10.69350979783266
l2 norm of gradients: 0.052703723987014325

l2 norm of weights: 2.284585857053528

Iteration Number: 9627
Loss: 10.693382152441602
l2 norm of gradients: 0.05269757606282651
l2 norm of weights: 2.284623584064008

Iteration Number: 9628
Loss: 10.693254562234651
l2 norm of gradients: 0.05269142992626002
l2 norm of weights: 2.284661313220493

Iteration Number: 9629
Loss: 10.693127027179495
l2 norm of gradients: 0.05268528557650912
l2 norm of weights: 2.2846990445200985

Iteration Number: 9630
Loss: 10.692999547243831
l2 norm of gradients: 0.0526791430127684
l2 norm of weights: 2.2847367779599437

Iteration Number: 9631
Loss: 10.69287212239539
l2 norm of gradients: 0.05267300223423294
l2 norm of weights: 2.2847745135371493

Iteration Number: 9632
Loss: 10.69274475260191
l2 norm of gradients: 0.05266686324009816
l2 norm of weights: 2.284812251248839

Iteration Number: 9633
Loss: 10.692617437831169
l2 norm of gradients: 0.052660726029559915
l2 norm of weights: 2.2848499910921385

Iteration Number: 9634
Loss: 10.692490178050955
l2 norm of gradients: 0.052654590601814565
l2 norm of weights: 2.2848877330641746

Iteration Number: 9635
Loss: 10.692362973229073
l2 norm of gradients: 0.052648456956058694
l2 norm of weights: 2.284925477162078

Iteration Number: 9636

Loss: 10.69223582333336
l2 norm of gradients: 0.05264232509148943
l2 norm of weights: 2.284963223382981

Iteration Number: 9637
Loss: 10.69210872833167
l2 norm of gradients: 0.05263619500730422
l2 norm of weights: 2.285000971724017

Iteration Number: 9638
Loss: 10.691981688191877
l2 norm of gradients: 0.052630066702701046
l2 norm of weights: 2.285038722182324

Iteration Number: 9639
Loss: 10.691854702881878
l2 norm of gradients: 0.05262394017687818
l2 norm of weights: 2.28507647475504

Iteration Number: 9640
Loss: 10.691727772369592
l2 norm of gradients: 0.05261781542903435
l2 norm of weights: 2.2851142294393068

Iteration Number: 9641
Loss: 10.69160089662296
l2 norm of gradients: 0.05261169245836863
l2 norm of weights: 2.285151986232267

Iteration Number: 9642
Loss: 10.691474075609944
l2 norm of gradients: 0.05260557126408063
l2 norm of weights: 2.2851897451310674

Iteration Number: 9643
Loss: 10.691347309298527
l2 norm of gradients: 0.05259945184537024
l2 norm of weights: 2.285227506132855

Iteration Number: 9644
Loss: 10.691220597656708
l2 norm of gradients: 0.05259333420143781
l2 norm of weights: 2.28526526923478

Iteration Number: 9645
Loss: 10.691093940652518
l2 norm of gradients: 0.052587218331484144
l2 norm of weights: 2.285303034433995

```
-----  
Iteration Number: 9646  
Loss: 10.690967338254003  
l2 norm of gradients: 0.052581104234710324  
l2 norm of weights: 2.285340801727655  
-----  
Iteration Number: 9647  
Loss: 10.690840790429228  
l2 norm of gradients: 0.05257499191031795  
l2 norm of weights: 2.2853785711129158  
-----  
Iteration Number: 9648  
Loss: 10.690714297146288  
l2 norm of gradients: 0.05256888135750898  
l2 norm of weights: 2.2854163425869367  
-----  
Iteration Number: 9649  
Loss: 10.69058785837329  
l2 norm of gradients: 0.05256277257548582  
l2 norm of weights: 2.2854541161468798  
-----  
Iteration Number: 9650  
Loss: 10.690461474078363  
l2 norm of gradients: 0.052556665563451246  
l2 norm of weights: 2.285491891789907  
-----  
Iteration Number: 9651  
Loss: 10.69033514422967  
l2 norm of gradients: 0.052550560320608404  
l2 norm of weights: 2.2855296695131857  
-----  
Iteration Number: 9652  
Loss: 10.690208868795377  
l2 norm of gradients: 0.05254445684616094  
l2 norm of weights: 2.2855674493138824  
-----  
Iteration Number: 9653  
Loss: 10.690082647743687  
l2 norm of gradients: 0.0525383551393128  
l2 norm of weights: 2.285605231189168  
-----  
Iteration Number: 9654  
Loss: 10.689956481042808  
l2 norm of gradients: 0.05253225519926844  
l2 norm of weights: 2.285643015136215  
-----  
Iteration Number: 9655  
Loss: 10.689830368660987
```

```

12 norm of gradients: 0.05252615702523259
12 norm of weights: 2.285680801152197
-----
Iteration Number: 9656
Loss: 10.689704310566484
12 norm of gradients: 0.05252006061641054
12 norm of weights: 2.2857185892342917
-----
Iteration Number: 9657
Loss: 10.68957830672757
12 norm of gradients: 0.05251396597200782
12 norm of weights: 2.2857563793796776
-----
Iteration Number: 9658
Loss: 10.689452357112557
12 norm of gradients: 0.052507873091230534
12 norm of weights: 2.2857941715855357
-----
Iteration Number: 9659
Loss: 10.689326461689763
12 norm of gradients: 0.05250178197328502
12 norm of weights: 2.28583196584905
-----
Iteration Number: 9660
Loss: 10.689200620427536
12 norm of gradients: 0.052495692617378145
12 norm of weights: 2.285869762167405
-----
Iteration Number: 9661
Loss: 10.689074833294239
12 norm of gradients: 0.05248960502271715
12 norm of weights: 2.28590756053779
-----
Iteration Number: 9662
Loss: 10.68894910025826
12 norm of gradients: 0.05248351918850962
12 norm of weights: 2.2859453609573936
-----
Iteration Number: 9663
Loss: 10.688823421288003
12 norm of gradients: 0.05247743511396358
12 norm of weights: 2.285983163423409
-----
Iteration Number: 9664
Loss: 10.688697796351903
12 norm of gradients: 0.052471352798287546
12 norm of weights: 2.28602096793303
-----

```

Iteration Number: 9665
Loss: 10.688572225418405
l2 norm of gradients: 0.05246527224069027
l2 norm of weights: 2.2860587744834526

Iteration Number: 9666
Loss: 10.688446708455979
l2 norm of gradients: 0.05245919344038099
l2 norm of weights: 2.2860965830718767

Iteration Number: 9667
Loss: 10.688321245433121
l2 norm of gradients: 0.05245311639656939
l2 norm of weights: 2.286134393695503

Iteration Number: 9668
Loss: 10.688195836318346
l2 norm of gradients: 0.052447041108465486
l2 norm of weights: 2.286172206351534

Iteration Number: 9669
Loss: 10.688070481080182
l2 norm of gradients: 0.05244096757527972
l2 norm of weights: 2.2862100210371756

Iteration Number: 9670
Loss: 10.687945179687187
l2 norm of gradients: 0.052434895796222965
l2 norm of weights: 2.2862478377496354

Iteration Number: 9671
Loss: 10.687819932107937
l2 norm of gradients: 0.05242882577050642
l2 norm of weights: 2.286285656486123

Iteration Number: 9672
Loss: 10.68769473831103
l2 norm of gradients: 0.05242275749734174
l2 norm of weights: 2.2863234772438497

Iteration Number: 9673
Loss: 10.687569598265082
l2 norm of gradients: 0.05241669097594099
l2 norm of weights: 2.2863613000200305

Iteration Number: 9674
Loss: 10.687444511938732
l2 norm of gradients: 0.0524106262055166

```
l2 norm of weights: 2.286399124811881
-----
Iteration Number: 9675
Loss: 10.687319479300644
l2 norm of gradients: 0.052404563185281396
l2 norm of weights: 2.28643695161662
-----
Iteration Number: 9676
Loss: 10.687194500319494
l2 norm of gradients: 0.05239850191444868
l2 norm of weights: 2.2864747804314676
-----
Iteration Number: 9677
Loss: 10.687069574963981
l2 norm of gradients: 0.052392442392232064
l2 norm of weights: 2.286512611253647
-----
Iteration Number: 9678
Loss: 10.686944703202837
l2 norm of gradients: 0.052386384617845554
l2 norm of weights: 2.2865504440803828
-----
Iteration Number: 9679
Loss: 10.686819885004798
l2 norm of gradients: 0.05238032859050366
l2 norm of weights: 2.2865882789089027
-----
Iteration Number: 9680
Loss: 10.686695120338632
l2 norm of gradients: 0.052374274309421154
l2 norm of weights: 2.2866261157364356
-----
Iteration Number: 9681
Loss: 10.686570409173122
l2 norm of gradients: 0.052368221773813346
l2 norm of weights: 2.286663954560213
-----
Iteration Number: 9682
Loss: 10.686445751477073
l2 norm of gradients: 0.05236217098289585
l2 norm of weights: 2.2867017953774686
-----
Iteration Number: 9683
Loss: 10.686321147219315
l2 norm of gradients: 0.05235612193588466
l2 norm of weights: 2.286739638185438
-----
Iteration Number: 9684
```


Loss: 10.686196596368696
l2 norm of gradients: 0.05235007463199627
l2 norm of weights: 2.2867774829813596

Iteration Number: 9685
Loss: 10.686072098894082
l2 norm of gradients: 0.052344029070447474
l2 norm of weights: 2.286815329762473

Iteration Number: 9686
Loss: 10.685947654764359
l2 norm of gradients: 0.05233798525045555
l2 norm of weights: 2.286853178526021

Iteration Number: 9687
Loss: 10.685823263948446
l2 norm of gradients: 0.0523319431712381
l2 norm of weights: 2.2868910292692473

Iteration Number: 9688
Loss: 10.685698926415263
l2 norm of gradients: 0.05232590283201313
l2 norm of weights: 2.286928881989399

Iteration Number: 9689
Loss: 10.68557464213377
l2 norm of gradients: 0.05231986423199911
l2 norm of weights: 2.286966736683725

Iteration Number: 9690
Loss: 10.685450411072935
l2 norm of gradients: 0.05231382737041477
l2 norm of weights: 2.287004593349476

Iteration Number: 9691
Loss: 10.685326233201753
l2 norm of gradients: 0.05230779224647943
l2 norm of weights: 2.2870424519839045

Iteration Number: 9692
Loss: 10.685202108489236
l2 norm of gradients: 0.05230175885941269
l2 norm of weights: 2.287080312584267

Iteration Number: 9693
Loss: 10.685078036904418
l2 norm of gradients: 0.05229572720843451
l2 norm of weights: 2.2871181751478193

```

-----
Iteration Number: 9694
Loss: 10.684954018416354
l2 norm of gradients: 0.05228969729276528
l2 norm of weights: 2.2871560396718222
-----
Iteration Number: 9695
Loss: 10.68483005299412
l2 norm of gradients: 0.052283669111625866
l2 norm of weights: 2.2871939061535365
-----
Iteration Number: 9696
Loss: 10.68470614060681
l2 norm of gradients: 0.052277642664237456
l2 norm of weights: 2.287231774590227
-----
Iteration Number: 9697
Loss: 10.684582281223546
l2 norm of gradients: 0.05227161794982161
l2 norm of weights: 2.287269644979158
-----
Iteration Number: 9698
Loss: 10.684458474813464
l2 norm of gradients: 0.05226559496760033
l2 norm of weights: 2.2873075173175987
-----
Iteration Number: 9699
Loss: 10.684334721345715
l2 norm of gradients: 0.052259573716796
l2 norm of weights: 2.287345391602819
-----
Iteration Number: 9700
Loss: 10.684211020789483
l2 norm of gradients: 0.052253554196631434
l2 norm of weights: 2.2873832678320913
-----
Iteration Number: 9701
Loss: 10.684087373113966
l2 norm of gradients: 0.05224753640632972
l2 norm of weights: 2.28742114600269
-----
Iteration Number: 9702
Loss: 10.683963778288387
l2 norm of gradients: 0.052241520345114485
l2 norm of weights: 2.2874590261118923
-----
Iteration Number: 9703
Loss: 10.68384023628198

```

```

12 norm of gradients: 0.0522355060122097
12 norm of weights: 2.287496908156976
-----
Iteration Number: 9704
Loss: 10.683716747064013
12 norm of gradients: 0.05222949340683969
12 norm of weights: 2.2875347921352223
-----
Iteration Number: 9705
Loss: 10.68359331060376
12 norm of gradients: 0.05222348252822921
12 norm of weights: 2.2875726780439143
-----
Iteration Number: 9706
Loss: 10.683469926870526
12 norm of gradients: 0.05221747337560342
12 norm of weights: 2.2876105658803367
-----
Iteration Number: 9707
Loss: 10.683346595833633
12 norm of gradients: 0.052211465948187825
12 norm of weights: 2.2876484556417775
-----
Iteration Number: 9708
Loss: 10.68322331746242
12 norm of gradients: 0.052205460245208356
12 norm of weights: 2.2876863473255256
-----
Iteration Number: 9709
Loss: 10.683100091726258
12 norm of gradients: 0.052199456265891384
12 norm of weights: 2.2877242409288727
-----
Iteration Number: 9710
Loss: 10.682976918594523
12 norm of gradients: 0.05219345400946358
12 norm of weights: 2.2877621364491123
-----
Iteration Number: 9711
Loss: 10.68285379803662
12 norm of gradients: 0.052187453475152065
12 norm of weights: 2.28780003388354
-----
Iteration Number: 9712
Loss: 10.682730730021976
12 norm of gradients: 0.05218145466218431
12 norm of weights: 2.2878379332294534
-----

```

Iteration Number: 9713
Loss: 10.682607714520035
l2 norm of gradients: 0.05217545756978823
l2 norm of weights: 2.287875834484153

Iteration Number: 9714
Loss: 10.682484751500262
l2 norm of gradients: 0.05216946219719213
l2 norm of weights: 2.2879137376449408

Iteration Number: 9715
Loss: 10.682361840932142
l2 norm of gradients: 0.05216346854362466
l2 norm of weights: 2.287951642709121

Iteration Number: 9716
Loss: 10.682238982785178
l2 norm of gradients: 0.0521574766083149
l2 norm of weights: 2.2879895496739997

Iteration Number: 9717
Loss: 10.6821161770289
l2 norm of gradients: 0.05215148639049231
l2 norm of weights: 2.288027458536885

Iteration Number: 9718
Loss: 10.681993423632854
l2 norm of gradients: 0.052145497889386695
l2 norm of weights: 2.288065369295088

Iteration Number: 9719
Loss: 10.681870722566602
l2 norm of gradients: 0.052139511104228375
l2 norm of weights: 2.288103281945921

Iteration Number: 9720
Loss: 10.681748073799737
l2 norm of gradients: 0.05213352603424791
l2 norm of weights: 2.288141196486698

Iteration Number: 9721
Loss: 10.681625477301862
l2 norm of gradients: 0.05212754267867636
l2 norm of weights: 2.2881791129147375

Iteration Number: 9722
Loss: 10.681502933042609
l2 norm of gradients: 0.05212156103674513

```
l2 norm of weights: 2.288217031227357
-----
Iteration Number: 9723
Loss: 10.68138044099162
l2 norm of gradients: 0.052115581107686014
l2 norm of weights: 2.2882549514218784
-----
Iteration Number: 9724
Loss: 10.681258001118566
l2 norm of gradients: 0.05210960289073118
l2 norm of weights: 2.288292873495624
-----
Iteration Number: 9725
Loss: 10.681135613393138
l2 norm of gradients: 0.052103626385113296
l2 norm of weights: 2.2883307974459197
-----
Iteration Number: 9726
Loss: 10.68101327778504
l2 norm of gradients: 0.05209765159006525
l2 norm of weights: 2.288368723270092
-----
Iteration Number: 9727
Loss: 10.680890994263999
l2 norm of gradients: 0.05209167850482043
l2 norm of weights: 2.2884066509654715
-----
Iteration Number: 9728
Loss: 10.68076876279977
l2 norm of gradients: 0.05208570712861261
l2 norm of weights: 2.2884445805293887
-----
Iteration Number: 9729
Loss: 10.680646583362117
l2 norm of gradients: 0.05207973746067588
l2 norm of weights: 2.2884825119591774
-----
Iteration Number: 9730
Loss: 10.680524455920832
l2 norm of gradients: 0.052073769500244824
l2 norm of weights: 2.2885204452521735
-----
Iteration Number: 9731
Loss: 10.68040238044572
l2 norm of gradients: 0.05206780324655427
l2 norm of weights: 2.2885583804057146
-----
Iteration Number: 9732
```

```

Loss: 10.680280356906616
l2 norm of gradients: 0.052061838698839645
l2 norm of weights: 2.28859631741714
-----
Iteration Number: 9733
Loss: 10.680158385273362
l2 norm of gradients: 0.052055875856336524
l2 norm of weights: 2.2886342562837925
-----
Iteration Number: 9734
Loss: 10.680036465515837
l2 norm of gradients: 0.05204991471828107
l2 norm of weights: 2.2886721970030157
-----
Iteration Number: 9735
Loss: 10.679914597603924
l2 norm of gradients: 0.05204395528390971
l2 norm of weights: 2.288710139572155
-----
Iteration Number: 9736
Loss: 10.679792781507532
l2 norm of gradients: 0.052037997552459296
l2 norm of weights: 2.2887480839885597
-----
Iteration Number: 9737
Loss: 10.679671017196595
l2 norm of gradients: 0.052032041523167115
l2 norm of weights: 2.2887860302495793
-----
Iteration Number: 9738
Loss: 10.679549304641062
l2 norm of gradients: 0.052026087195270736
l2 norm of weights: 2.2888239783525663
-----
Iteration Number: 9739
Loss: 10.679427643810902
l2 norm of gradients: 0.05202013456800825
l2 norm of weights: 2.2888619282948754
-----
Iteration Number: 9740
Loss: 10.6793060346761
l2 norm of gradients: 0.052014183640617974
l2 norm of weights: 2.2888998800738625
-----
Iteration Number: 9741
Loss: 10.679184477206674
l2 norm of gradients: 0.05200823441233876
l2 norm of weights: 2.288937833686886

```

```
-----
Iteration Number: 9742
Loss: 10.679062971372646
l2 norm of gradients: 0.052002286882409744
l2 norm of weights: 2.288975789131307
-----
Iteration Number: 9743
Loss: 10.678941517144072
l2 norm of gradients: 0.051996341050070546
l2 norm of weights: 2.289013746404488
-----
Iteration Number: 9744
Loss: 10.678820114491018
l2 norm of gradients: 0.05199039691456109
l2 norm of weights: 2.2890517055037938
-----
Iteration Number: 9745
Loss: 10.678698763383574
l2 norm of gradients: 0.051984454475121664
l2 norm of weights: 2.28908966642659
-----
Iteration Number: 9746
Loss: 10.67857746379185
l2 norm of gradients: 0.051978513730993006
l2 norm of weights: 2.2891276291702476
-----
Iteration Number: 9747
Loss: 10.678456215685978
l2 norm of gradients: 0.051972574681416264
l2 norm of weights: 2.2891655937321356
-----
Iteration Number: 9748
Loss: 10.678335019036103
l2 norm of gradients: 0.051966637325632906
l2 norm of weights: 2.2892035601096277
-----
Iteration Number: 9749
Loss: 10.678213873812394
l2 norm of gradients: 0.05196070166288477
l2 norm of weights: 2.289241528300099
-----
Iteration Number: 9750
Loss: 10.678092779985043
l2 norm of gradients: 0.051954767692414174
l2 norm of weights: 2.2892794983009264
-----
Iteration Number: 9751
Loss: 10.677971737524262
```

l2 norm of gradients: 0.051948835413463734
l2 norm of weights: 2.289317470109489

Iteration Number: 9752
Loss: 10.677850746400274
l2 norm of gradients: 0.05194290482527648
l2 norm of weights: 2.289355443723168

Iteration Number: 9753
Loss: 10.67772980658333
l2 norm of gradients: 0.05193697592709583
l2 norm of weights: 2.2893934191393464

Iteration Number: 9754
Loss: 10.677608918043695
l2 norm of gradients: 0.051931048718165555
l2 norm of weights: 2.28943139635541

Iteration Number: 9755
Loss: 10.677488080751662
l2 norm of gradients: 0.05192512319772987
l2 norm of weights: 2.2894693753687454

Iteration Number: 9756
Loss: 10.677367294677536
l2 norm of gradients: 0.051919199365033315
l2 norm of weights: 2.2895073561767423

Iteration Number: 9757
Loss: 10.677246559791648
l2 norm of gradients: 0.05191327721932085
l2 norm of weights: 2.289545338776792

Iteration Number: 9758
Loss: 10.677125876064341
l2 norm of gradients: 0.051907356759837826
l2 norm of weights: 2.289583323166288

Iteration Number: 9759
Loss: 10.677005243465986
l2 norm of gradients: 0.05190143798582987
l2 norm of weights: 2.2896213093426256

Iteration Number: 9760
Loss: 10.676884661966971
l2 norm of gradients: 0.0518955208965432
l2 norm of weights: 2.2896592973032033

Iteration Number: 9761
Loss: 10.676764131537698
l2 norm of gradients: 0.051889605491224206
l2 norm of weights: 2.2896972870454193

Iteration Number: 9762
Loss: 10.676643652148595
l2 norm of gradients: 0.051883691769119764
l2 norm of weights: 2.289735278566676

Iteration Number: 9763
Loss: 10.67652322377011
l2 norm of gradients: 0.05187777972947715
l2 norm of weights: 2.289773271864376

Iteration Number: 9764
Loss: 10.676402846372708
l2 norm of gradients: 0.05187186937154394
l2 norm of weights: 2.2898112669359265

Iteration Number: 9765
Loss: 10.676282519926874
l2 norm of gradients: 0.05186596069456817
l2 norm of weights: 2.289849263778734

Iteration Number: 9766
Loss: 10.67616224440311
l2 norm of gradients: 0.05186005369779824
l2 norm of weights: 2.2898872623902093

Iteration Number: 9767
Loss: 10.676042019771948
l2 norm of gradients: 0.051854148380482905
l2 norm of weights: 2.289925262767763

Iteration Number: 9768
Loss: 10.67592184600393
l2 norm of gradients: 0.0518482447418713
l2 norm of weights: 2.2899632649088093

Iteration Number: 9769
Loss: 10.675801723069617
l2 norm of gradients: 0.051842342781213
l2 norm of weights: 2.2900012688107636

Iteration Number: 9770
Loss: 10.675681650939593
l2 norm of gradients: 0.051836442497757824

l2 norm of weights: 2.2900392744710447

Iteration Number: 9771
Loss: 10.675561629584465
l2 norm of gradients: 0.0518305438907562
l2 norm of weights: 2.290077281887072

Iteration Number: 9772
Loss: 10.67544165897485
l2 norm of gradients: 0.05182464695945869
l2 norm of weights: 2.2901152910562668

Iteration Number: 9773
Loss: 10.675321739081394
l2 norm of gradients: 0.05181875170311637
l2 norm of weights: 2.2901533019760536

Iteration Number: 9774
Loss: 10.675201869874758
l2 norm of gradients: 0.051812858120980715
l2 norm of weights: 2.290191314643858

Iteration Number: 9775
Loss: 10.675082051325624
l2 norm of gradients: 0.051806966212303504
l2 norm of weights: 2.2902293290571083

Iteration Number: 9776
Loss: 10.674962283404692
l2 norm of gradients: 0.0518010759763369
l2 norm of weights: 2.290267345213234

Iteration Number: 9777
Loss: 10.674842566082685
l2 norm of gradients: 0.051795187412333535
l2 norm of weights: 2.2903053631096673

Iteration Number: 9778
Loss: 10.674722899330337
l2 norm of gradients: 0.05178930051954635
l2 norm of weights: 2.290343382743842

Iteration Number: 9779
Loss: 10.674603283118412
l2 norm of gradients: 0.051783415297228616
l2 norm of weights: 2.290381404113194

Iteration Number: 9780

Loss: 10.67448371741769
l2 norm of gradients: 0.05177753174463414
l2 norm of weights: 2.290419427215161

Iteration Number: 9781
Loss: 10.674364202198962
l2 norm of gradients: 0.0517716498610169
l2 norm of weights: 2.2904574520471837

Iteration Number: 9782
Loss: 10.674244737433055
l2 norm of gradients: 0.05176576964563144
l2 norm of weights: 2.2904954786067035

Iteration Number: 9783
Loss: 10.6741253230908
l2 norm of gradients: 0.051759891097732595
l2 norm of weights: 2.2905335068911645

Iteration Number: 9784
Loss: 10.674005959143056
l2 norm of gradients: 0.05175401421657554
l2 norm of weights: 2.2905715368980126

Iteration Number: 9785
Loss: 10.673886645560696
l2 norm of gradients: 0.05174813900141592
l2 norm of weights: 2.290609568624696

Iteration Number: 9786
Loss: 10.673767382314619
l2 norm of gradients: 0.05174226545150967
l2 norm of weights: 2.290647602068665

Iteration Number: 9787
Loss: 10.673648169375737
l2 norm of gradients: 0.051736393566113224
l2 norm of weights: 2.2906856372273707

Iteration Number: 9788
Loss: 10.673529006714986
l2 norm of gradients: 0.051730523344483206
l2 norm of weights: 2.2907236740982673

Iteration Number: 9789
Loss: 10.673409894303315
l2 norm of gradients: 0.05172465478587679
l2 norm of weights: 2.290761712678811

```
-----  
Iteration Number: 9790  
Loss: 10.673290832111705  
l2 norm of gradients: 0.051718787889551456  
l2 norm of weights: 2.29079975296646  
-----  
Iteration Number: 9791  
Loss: 10.67317182011114  
l2 norm of gradients: 0.051712922654765084  
l2 norm of weights: 2.2908377949586733  
-----  
Iteration Number: 9792  
Loss: 10.673052858272634  
l2 norm of gradients: 0.05170705908077585  
l2 norm of weights: 2.290875838652914  
-----  
Iteration Number: 9793  
Loss: 10.672933946567218  
l2 norm of gradients: 0.05170119716684245  
l2 norm of weights: 2.2909138840466454  
-----  
Iteration Number: 9794  
Loss: 10.672815084965938  
l2 norm of gradients: 0.05169533691222382  
l2 norm of weights: 2.290951931137333  
-----  
Iteration Number: 9795  
Loss: 10.672696273439874  
l2 norm of gradients: 0.05168947831617933  
l2 norm of weights: 2.290989979922446  
-----  
Iteration Number: 9796  
Loss: 10.672577511960101  
l2 norm of gradients: 0.051683621377968725  
l2 norm of weights: 2.2910280303994526  
-----  
Iteration Number: 9797  
Loss: 10.672458800497733  
l2 norm of gradients: 0.051677766096852124  
l2 norm of weights: 2.2910660825658256  
-----  
Iteration Number: 9798  
Loss: 10.672340139023895  
l2 norm of gradients: 0.05167191247209004  
l2 norm of weights: 2.2911041364190385  
-----  
Iteration Number: 9799  
Loss: 10.672221527509734
```

```

12 norm of gradients: 0.05166606050294331
12 norm of weights: 2.2911421919565678
-----
Iteration Number: 9800
Loss: 10.672102965926419
12 norm of gradients: 0.05166021018867322
12 norm of weights: 2.2911802491758904
-----
Iteration Number: 9801
Loss: 10.671984454245127
12 norm of gradients: 0.051654361528541325
12 norm of weights: 2.2912183080744866
-----
Iteration Number: 9802
Loss: 10.671865992437066
12 norm of gradients: 0.05164851452180968
12 norm of weights: 2.2912563686498375
-----
Iteration Number: 9803
Loss: 10.671747580473456
12 norm of gradients: 0.05164266916774061
12 norm of weights: 2.291294430899428
-----
Iteration Number: 9804
Loss: 10.67162921832554
12 norm of gradients: 0.05163682546559687
12 norm of weights: 2.2913324948207423
-----
Iteration Number: 9805
Loss: 10.671510905964583
12 norm of gradients: 0.05163098341464158
12 norm of weights: 2.291370560411269
-----
Iteration Number: 9806
Loss: 10.671392643361859
12 norm of gradients: 0.05162514301413824
12 norm of weights: 2.2914086276684973
-----
Iteration Number: 9807
Loss: 10.671274430488669
12 norm of gradients: 0.05161930426335069
12 norm of weights: 2.2914466965899187
-----
Iteration Number: 9808
Loss: 10.67115626731633
12 norm of gradients: 0.05161346716154318
12 norm of weights: 2.2914847671730274
-----

```

Iteration Number: 9809
Loss: 10.671038153816184
l2 norm of gradients: 0.051607631707980266
l2 norm of weights: 2.291522839415318

Iteration Number: 9810
Loss: 10.670920089959582
l2 norm of gradients: 0.05160179790192701
l2 norm of weights: 2.2915609133142887

Iteration Number: 9811
Loss: 10.670802075717903
l2 norm of gradients: 0.05159596574264872
l2 norm of weights: 2.291598988867438

Iteration Number: 9812
Loss: 10.670684111062538
l2 norm of gradients: 0.05159013522941112
l2 norm of weights: 2.2916370660722687

Iteration Number: 9813
Loss: 10.670566195964904
l2 norm of gradients: 0.05158430636148033
l2 norm of weights: 2.291675144926283

Iteration Number: 9814
Loss: 10.670448330396429
l2 norm of gradients: 0.051578479138122815
l2 norm of weights: 2.291713225426987

Iteration Number: 9815
Loss: 10.670330514328569
l2 norm of gradients: 0.051572653558605394
l2 norm of weights: 2.2917513075718867

Iteration Number: 9816
Loss: 10.670212747732792
l2 norm of gradients: 0.05156682962219531
l2 norm of weights: 2.2917893913584924

Iteration Number: 9817
Loss: 10.67009503058059
l2 norm of gradients: 0.051561007328160186
l2 norm of weights: 2.2918274767843148

Iteration Number: 9818
Loss: 10.669977362843467
l2 norm of gradients: 0.05155518667576791

```

12 norm of weights: 2.2918655638468666
-----
Iteration Number: 9819
Loss: 10.669859744492953
12 norm of gradients: 0.05154936766428682
12 norm of weights: 2.291903652543664
-----
Iteration Number: 9820
Loss: 10.669742175500595
12 norm of gradients: 0.05154355029298566
12 norm of weights: 2.291941742872223
-----
Iteration Number: 9821
Loss: 10.669624655837955
12 norm of gradients: 0.051537734561133475
12 norm of weights: 2.291979834830063
-----
Iteration Number: 9822
Loss: 10.669507185476618
12 norm of gradients: 0.05153192046799966
12 norm of weights: 2.2920179284147038
-----
Iteration Number: 9823
Loss: 10.669389764388187
12 norm of gradients: 0.051526108012854135
12 norm of weights: 2.29205602362367
-----
Iteration Number: 9824
Loss: 10.669272392544286
12 norm of gradients: 0.05152029719496701
12 norm of weights: 2.292094120454485
-----
Iteration Number: 9825
Loss: 10.669155069916554
12 norm of gradients: 0.05151448801360883
12 norm of weights: 2.292132218904676
-----
Iteration Number: 9826
Loss: 10.66903779647665
12 norm of gradients: 0.05150868046805057
12 norm of weights: 2.2921703189717717
-----
Iteration Number: 9827
Loss: 10.668920572196251
12 norm of gradients: 0.051502874557563465
12 norm of weights: 2.2922084206533024
-----
Iteration Number: 9828

```

Loss: 10.668803397047057
l2 norm of gradients: 0.05149707028141921
l2 norm of weights: 2.292246523946801

Iteration Number: 9829
Loss: 10.66868627100078
l2 norm of gradients: 0.05149126763888983
l2 norm of weights: 2.292284628849801

Iteration Number: 9830
Loss: 10.668569194029162
l2 norm of gradients: 0.05148546662924774
l2 norm of weights: 2.29232273535984

Iteration Number: 9831
Loss: 10.66845216610395
l2 norm of gradients: 0.05147966725176567
l2 norm of weights: 2.2923608434744556

Iteration Number: 9832
Loss: 10.668335187196918
l2 norm of gradients: 0.051473869505716786
l2 norm of weights: 2.2923989531911877

Iteration Number: 9833
Loss: 10.668218257279857
l2 norm of gradients: 0.05146807339037459
l2 norm of weights: 2.292437064507579

Iteration Number: 9834
Loss: 10.668101376324575
l2 norm of gradients: 0.051462278905012974
l2 norm of weights: 2.2924751774211733

Iteration Number: 9835
Loss: 10.667984544302906
l2 norm of gradients: 0.051456486048906167
l2 norm of weights: 2.292513291929517

Iteration Number: 9836
Loss: 10.66786776118669
l2 norm of gradients: 0.05145069482132878
l2 norm of weights: 2.2925514080301572

Iteration Number: 9837
Loss: 10.667751026947798
l2 norm of gradients: 0.05144490522155578
l2 norm of weights: 2.2925895257206443


```
-----  
Iteration Number: 9838  
Loss: 10.667634341558113  
l2 norm of gradients: 0.051439117248862555  
l2 norm of weights: 2.29262764499853  
-----  
Iteration Number: 9839  
Loss: 10.667517704989537  
l2 norm of gradients: 0.05143333090252476  
l2 norm of weights: 2.292665765861368  
-----  
Iteration Number: 9840  
Loss: 10.667401117213995  
l2 norm of gradients: 0.05142754618181853  
l2 norm of weights: 2.292703888306714  
-----  
Iteration Number: 9841  
Loss: 10.667284578203425  
l2 norm of gradients: 0.051421763086020277  
l2 norm of weights: 2.292742012332125  
-----  
Iteration Number: 9842  
Loss: 10.667168087929785  
l2 norm of gradients: 0.051415981614406824  
l2 norm of weights: 2.2927801379351607  
-----  
Iteration Number: 9843  
Loss: 10.667051646365058  
l2 norm of gradients: 0.05141020176625538  
l2 norm of weights: 2.2928182651133824  
-----  
Iteration Number: 9844  
Loss: 10.666935253481235  
l2 norm of gradients: 0.05140442354084347  
l2 norm of weights: 2.2928563938643536  
-----  
Iteration Number: 9845  
Loss: 10.66681890925033  
l2 norm of gradients: 0.05139864693744899  
l2 norm of weights: 2.292894524185639  
-----  
Iteration Number: 9846  
Loss: 10.666702613644386  
l2 norm of gradients: 0.05139287195535026  
l2 norm of weights: 2.2929326560748056  
-----  
Iteration Number: 9847  
Loss: 10.666586366635446
```

l2 norm of gradients: 0.051387098593825904
l2 norm of weights: 2.2929707895294227

Iteration Number: 9848
Loss: 10.66647016819559
l2 norm of gradients: 0.051381326852154977
l2 norm of weights: 2.293008924547061

Iteration Number: 9849
Loss: 10.666354018296895
l2 norm of gradients: 0.05137555672961678
l2 norm of weights: 2.2930470611252933

Iteration Number: 9850
Loss: 10.666237916911475
l2 norm of gradients: 0.05136978822549111
l2 norm of weights: 2.2930851992616943

Iteration Number: 9851
Loss: 10.66612186401146
l2 norm of gradients: 0.051364021339058076
l2 norm of weights: 2.2931233389538406

Iteration Number: 9852
Loss: 10.666005859568987
l2 norm of gradients: 0.05135825606959811
l2 norm of weights: 2.29316148019931

Iteration Number: 9853
Loss: 10.665889903556227
l2 norm of gradients: 0.05135249241639213
l2 norm of weights: 2.293199622995684

Iteration Number: 9854
Loss: 10.665773995945358
l2 norm of gradients: 0.05134673037872126
l2 norm of weights: 2.293237767340544

Iteration Number: 9855
Loss: 10.665658136708583
l2 norm of gradients: 0.051340969955867094
l2 norm of weights: 2.2932759132314744

Iteration Number: 9856
Loss: 10.665542325818116
l2 norm of gradients: 0.05133521114711157
l2 norm of weights: 2.293314060666061

Iteration Number: 9857
Loss: 10.665426563246198
l2 norm of gradients: 0.05132945395173702
l2 norm of weights: 2.2933522096418923

Iteration Number: 9858
Loss: 10.665310848965083
l2 norm of gradients: 0.05132369836902602
l2 norm of weights: 2.2933903601565575

Iteration Number: 9859
Loss: 10.665195182947047
l2 norm of gradients: 0.051317944398261664
l2 norm of weights: 2.2934285122076483

Iteration Number: 9860
Loss: 10.665079565164383
l2 norm of gradients: 0.05131219203872733
l2 norm of weights: 2.2934666657927587

Iteration Number: 9861
Loss: 10.664963995589398
l2 norm of gradients: 0.05130644128970675
l2 norm of weights: 2.293504820909484

Iteration Number: 9862
Loss: 10.664848474194425
l2 norm of gradients: 0.05130069215048404
l2 norm of weights: 2.293542977555421

Iteration Number: 9863
Loss: 10.664733000951808
l2 norm of gradients: 0.05129494462034368
l2 norm of weights: 2.29358113572817

Iteration Number: 9864
Loss: 10.664617575833915
l2 norm of gradients: 0.051289198698570515
l2 norm of weights: 2.293619295425331

Iteration Number: 9865
Loss: 10.664502198813132
l2 norm of gradients: 0.051283454384449735
l2 norm of weights: 2.293657456644508

Iteration Number: 9866
Loss: 10.664386869861861
l2 norm of gradients: 0.051277711677266934

```

12 norm of weights: 2.293695619383305
-----
Iteration Number: 9867
Loss: 10.66427158895252
12 norm of gradients: 0.051271970576307994
12 norm of weights: 2.29373378363933
-----
Iteration Number: 9868
Loss: 10.664156356057552
12 norm of gradients: 0.05126623108085925
12 norm of weights: 2.29377194941019
-----
Iteration Number: 9869
Loss: 10.664041171149414
12 norm of gradients: 0.05126049319020732
12 norm of weights: 2.2938101166934963
-----
Iteration Number: 9870
Loss: 10.663926034200577
12 norm of gradients: 0.051254756903639204
12 norm of weights: 2.293848285486861
-----
Iteration Number: 9871
Loss: 10.663810945183542
12 norm of gradients: 0.05124902222044232
12 norm of weights: 2.293886455787899
-----
Iteration Number: 9872
Loss: 10.663695904070819
12 norm of gradients: 0.051243289139904384
12 norm of weights: 2.2939246275942256
-----
Iteration Number: 9873
Loss: 10.663580910834938
12 norm of gradients: 0.05123755766131348
12 norm of weights: 2.293962800903459
-----
Iteration Number: 9874
Loss: 10.663465965448449
12 norm of gradients: 0.05123182778395808
12 norm of weights: 2.294000975713219
-----
Iteration Number: 9875
Loss: 10.663351067883912
12 norm of gradients: 0.05122609950712698
12 norm of weights: 2.2940391520211283
-----
Iteration Number: 9876

```

Loss: 10.663236218113923
l2 norm of gradients: 0.05122037283010941
l2 norm of weights: 2.2940773298248085

Iteration Number: 9877
Loss: 10.663121416111078
l2 norm of gradients: 0.051214647752194836
l2 norm of weights: 2.2941155091218866

Iteration Number: 9878
Loss: 10.663006661848005
l2 norm of gradients: 0.05120892427267322
l2 norm of weights: 2.2941536899099892

Iteration Number: 9879
Loss: 10.662891955297336
l2 norm of gradients: 0.051203202390834784
l2 norm of weights: 2.2941918721867456

Iteration Number: 9880
Loss: 10.662777296431736
l2 norm of gradients: 0.05119748210597017
l2 norm of weights: 2.2942300559497872

Iteration Number: 9881
Loss: 10.662662685223879
l2 norm of gradients: 0.051191763417370334
l2 norm of weights: 2.2942682411967463

Iteration Number: 9882
Loss: 10.662548121646456
l2 norm of gradients: 0.051186046324326634
l2 norm of weights: 2.294306427925257

Iteration Number: 9883
Loss: 10.662433605672181
l2 norm of gradients: 0.051180330826130736
l2 norm of weights: 2.2943446161329573

Iteration Number: 9884
Loss: 10.662319137273787
l2 norm of gradients: 0.05117461692207475
l2 norm of weights: 2.2943828058174844

Iteration Number: 9885
Loss: 10.662204716424021
l2 norm of gradients: 0.05116890461145104
l2 norm of weights: 2.29442099697648

```
-----
Iteration Number: 9886
Loss: 10.662090343095645
l2 norm of gradients: 0.0511631938935524
l2 norm of weights: 2.2944591896075845
-----
Iteration Number: 9887
Loss: 10.661976017261454
l2 norm of gradients: 0.05115748476767198
l2 norm of weights: 2.294497383708443
-----
Iteration Number: 9888
Loss: 10.661861738894244
l2 norm of gradients: 0.05115177723310323
l2 norm of weights: 2.2945355792767
-----
Iteration Number: 9889
Loss: 10.661747507966833
l2 norm of gradients: 0.05114607128914002
l2 norm of weights: 2.294573776310005
-----
Iteration Number: 9890
Loss: 10.661633324452064
l2 norm of gradients: 0.051140366935076564
l2 norm of weights: 2.2946119748060063
-----
Iteration Number: 9891
Loss: 10.661519188322794
l2 norm of gradients: 0.05113466417020744
l2 norm of weights: 2.2946501747623556
-----
Iteration Number: 9892
Loss: 10.661405099551896
l2 norm of gradients: 0.051128962993827534
l2 norm of weights: 2.2946883761767056
-----
Iteration Number: 9893
Loss: 10.661291058112266
l2 norm of gradients: 0.05112326340523214
l2 norm of weights: 2.2947265790467117
-----
Iteration Number: 9894
Loss: 10.661177063976808
l2 norm of gradients: 0.05111756540371686
l2 norm of weights: 2.294764783370031
-----
Iteration Number: 9895
Loss: 10.66106311711846
```

```
12 norm of gradients: 0.05111186898857779
12 norm of weights: 2.2948029891443213
-----
Iteration Number: 9896
Loss: 10.660949217510158
12 norm of gradients: 0.051106174159111194
12 norm of weights: 2.2948411963672437
-----
Iteration Number: 9897
Loss: 10.660835365124878
12 norm of gradients: 0.051100480914613805
12 norm of weights: 2.294879405036461
-----
Iteration Number: 9898
Loss: 10.660721559935594
12 norm of gradients: 0.05109478925438264
12 norm of weights: 2.2949176151496364
-----
Iteration Number: 9899
Loss: 10.660607801915308
12 norm of gradients: 0.05108909917771521
12 norm of weights: 2.2949558267044363
-----
Iteration Number: 9900
Loss: 10.660494091037041
12 norm of gradients: 0.051083410683909244
12 norm of weights: 2.2949940396985284
-----
Iteration Number: 9901
Loss: 10.660380427273829
12 norm of gradients: 0.05107772377226285
12 norm of weights: 2.295032254129583
-----
Iteration Number: 9902
Loss: 10.660266810598722
12 norm of gradients: 0.05107203844207456
12 norm of weights: 2.295070469995271
-----
Iteration Number: 9903
Loss: 10.660153240984796
12 norm of gradients: 0.05106635469264321
12 norm of weights: 2.2951086872932653
-----
Iteration Number: 9904
Loss: 10.660039718405141
12 norm of gradients: 0.05106067252326796
12 norm of weights: 2.2951469060212424
-----
```

Iteration Number: 9905
Loss: 10.659926242832865
l2 norm of gradients: 0.05105499193324844
l2 norm of weights: 2.2951851261768774

Iteration Number: 9906
Loss: 10.659812814241088
l2 norm of gradients: 0.051049312921884464
l2 norm of weights: 2.2952233477578505

Iteration Number: 9907
Loss: 10.659699432602958
l2 norm of gradients: 0.05104363548847639
l2 norm of weights: 2.295261570761842

Iteration Number: 9908
Loss: 10.659586097891639
l2 norm of gradients: 0.05103795963232477
l2 norm of weights: 2.295299795186534

Iteration Number: 9909
Loss: 10.659472810080304
l2 norm of gradients: 0.051032285352730626
l2 norm of weights: 2.2953380210296106

Iteration Number: 9910
Loss: 10.659359569142152
l2 norm of gradients: 0.05102661264899523
l2 norm of weights: 2.295376248288758

Iteration Number: 9911
Loss: 10.659246375050396
l2 norm of gradients: 0.05102094152042032
l2 norm of weights: 2.2954144769616645

Iteration Number: 9912
Loss: 10.659133227778273
l2 norm of gradients: 0.05101527196630788
l2 norm of weights: 2.295452707046019

Iteration Number: 9913
Loss: 10.659020127299026
l2 norm of gradients: 0.051009603985960356
l2 norm of weights: 2.2954909385395137

Iteration Number: 9914
Loss: 10.658907073585928
l2 norm of gradients: 0.051003937578680464


```
l2 norm of weights: 2.295529171439841
-----
Iteration Number: 9915
Loss: 10.658794066612263
l2 norm of gradients: 0.05099827274377127
l2 norm of weights: 2.2955674057446966
-----
Iteration Number: 9916
Loss: 10.658681106351331
l2 norm of gradients: 0.05099260948053628
l2 norm of weights: 2.2956056414517776
-----
Iteration Number: 9917
Loss: 10.658568192776459
l2 norm of gradients: 0.050986947788279294
l2 norm of weights: 2.2956438785587823
-----
Iteration Number: 9918
Loss: 10.65845532586098
l2 norm of gradients: 0.050981287666304415
l2 norm of weights: 2.295682117063411
-----
Iteration Number: 9919
Loss: 10.65834250557825
l2 norm of gradients: 0.05097562911391619
l2 norm of weights: 2.2957203569633666
-----
Iteration Number: 9920
Loss: 10.658229731901645
l2 norm of gradients: 0.050969972130419504
l2 norm of weights: 2.295758598256352
-----
Iteration Number: 9921
Loss: 10.65811700480456
l2 norm of gradients: 0.05096431671511951
l2 norm of weights: 2.2957968409400746
-----
Iteration Number: 9922
Loss: 10.6580043242604
l2 norm of gradients: 0.05095866286732183
l2 norm of weights: 2.295835085012241
-----
Iteration Number: 9923
Loss: 10.657891690242591
l2 norm of gradients: 0.05095301058633234
l2 norm of weights: 2.295873330470562
-----
Iteration Number: 9924
```

Loss: 10.657779102724573
l2 norm of gradients: 0.050947359871457355
l2 norm of weights: 2.2959115773127468

Iteration Number: 9925
Loss: 10.65766656167982
l2 norm of gradients: 0.05094171072200347
l2 norm of weights: 2.29594982553651

Iteration Number: 9926
Loss: 10.6575540670818
l2 norm of gradients: 0.05093606313727766
l2 norm of weights: 2.295988075139566

Iteration Number: 9927
Loss: 10.657441618904018
l2 norm of gradients: 0.05093041711658723
l2 norm of weights: 2.2960263261196316

Iteration Number: 9928
Loss: 10.657329217119985
l2 norm of gradients: 0.05092477265923989
l2 norm of weights: 2.296064578474425

Iteration Number: 9929
Loss: 10.65721686170323
l2 norm of gradients: 0.050919129764543665
l2 norm of weights: 2.296102832201667

Iteration Number: 9930
Loss: 10.657104552627306
l2 norm of gradients: 0.05091348843180691
l2 norm of weights: 2.2961410872990786

Iteration Number: 9931
Loss: 10.65699228986578
l2 norm of gradients: 0.050907848660338366
l2 norm of weights: 2.2961793437643845

Iteration Number: 9932
Loss: 10.656880073392239
l2 norm of gradients: 0.050902210449447094
l2 norm of weights: 2.29621760159531

Iteration Number: 9933
Loss: 10.656767903180281
l2 norm of gradients: 0.050896573798442535
l2 norm of weights: 2.296255860789582

```

-----
Iteration Number: 9934
Loss: 10.656655779203527
l2 norm of gradients: 0.0508909387066345
l2 norm of weights: 2.296294121344931
-----
Iteration Number: 9935
Loss: 10.656543701435613
l2 norm of gradients: 0.050885305173333055
l2 norm of weights: 2.2963323832590863
-----
Iteration Number: 9936
Loss: 10.656431669850198
l2 norm of gradients: 0.050879673197848724
l2 norm of weights: 2.2963706465297817
-----
Iteration Number: 9937
Loss: 10.65631968442095
l2 norm of gradients: 0.05087404277949232
l2 norm of weights: 2.2964089111547508
-----
Iteration Number: 9938
Loss: 10.65620774512156
l2 norm of gradients: 0.05086841391757503
l2 norm of weights: 2.2964471771317307
-----
Iteration Number: 9939
Loss: 10.656095851925732
l2 norm of gradients: 0.050862786611408375
l2 norm of weights: 2.2964854444584586
-----
Iteration Number: 9940
Loss: 10.655984004807193
l2 norm of gradients: 0.05085716086030419
l2 norm of weights: 2.2965237131326752
-----
Iteration Number: 9941
Loss: 10.655872203739683
l2 norm of gradients: 0.050851536663574746
l2 norm of weights: 2.296561983152121
-----
Iteration Number: 9942
Loss: 10.655760448696967
l2 norm of gradients: 0.05084591402053264
l2 norm of weights: 2.2966002545145403
-----
Iteration Number: 9943
Loss: 10.655648739652813

```

```

12 norm of gradients: 0.05084029293049074
12 norm of weights: 2.296638527217678
-----
Iteration Number: 9944
Loss: 10.65553707658102
12 norm of gradients: 0.05083467339276231
12 norm of weights: 2.296676801259281
-----
Iteration Number: 9945
Loss: 10.655425459455397
12 norm of gradients: 0.050829055406661015
12 norm of weights: 2.296715076637097
-----
Iteration Number: 9946
Loss: 10.655313888249776
12 norm of gradients: 0.05082343897150081
12 norm of weights: 2.2967533533488775
-----
Iteration Number: 9947
Loss: 10.655202362937997
12 norm of gradients: 0.050817824086595964
12 norm of weights: 2.296791631392374
-----
Iteration Number: 9948
Loss: 10.655090883493926
12 norm of gradients: 0.05081221075126119
12 norm of weights: 2.296829910765341
-----
Iteration Number: 9949
Loss: 10.654979449891444
12 norm of gradients: 0.050806598964811434
12 norm of weights: 2.296868191465534
-----
Iteration Number: 9950
Loss: 10.654868062104448
12 norm of gradients: 0.05080098872656212
12 norm of weights: 2.29690647349071
-----
Iteration Number: 9951
Loss: 10.654756720106853
12 norm of gradients: 0.05079538003582889
12 norm of weights: 2.2969447568386285
-----
Iteration Number: 9952
Loss: 10.65464542387259
12 norm of gradients: 0.05078977289192786
12 norm of weights: 2.2969830415070507
-----

```

Iteration Number: 9953
Loss: 10.654534173375612
l2 norm of gradients: 0.05078416729417535
l2 norm of weights: 2.2970213274937388

Iteration Number: 9954
Loss: 10.654422968589882
l2 norm of gradients: 0.05077856324188818
l2 norm of weights: 2.2970596147964577

Iteration Number: 9955
Loss: 10.654311809489391
l2 norm of gradients: 0.050772960734383366
l2 norm of weights: 2.2970979034129733

Iteration Number: 9956
Loss: 10.654200696048129
l2 norm of gradients: 0.05076735977097839
l2 norm of weights: 2.297136193341054

Iteration Number: 9957
Loss: 10.654089628240124
l2 norm of gradients: 0.05076176035099104
l2 norm of weights: 2.2971744845784685

Iteration Number: 9958
Loss: 10.653978606039404
l2 norm of gradients: 0.05075616247373939
l2 norm of weights: 2.297212777122989

Iteration Number: 9959
Loss: 10.653867629420027
l2 norm of gradients: 0.05075056613854195
l2 norm of weights: 2.297251070972389

Iteration Number: 9960
Loss: 10.65375669835606
l2 norm of gradients: 0.05074497134471755
l2 norm of weights: 2.2972893661244433

Iteration Number: 9961
Loss: 10.653645812821594
l2 norm of gradients: 0.05073937809158532
l2 norm of weights: 2.2973276625769277

Iteration Number: 9962
Loss: 10.653534972790728
l2 norm of gradients: 0.05073378637846479

```
12 norm of weights: 2.297365960327622
-----
Iteration Number: 9963
Loss: 10.65342417823759
12 norm of gradients: 0.050728196204675766
12 norm of weights: 2.2974042593743054
-----
Iteration Number: 9964
Loss: 10.653313429136313
12 norm of gradients: 0.050722607569538516
12 norm of weights: 2.2974425597147596
-----
Iteration Number: 9965
Loss: 10.653202725461053
12 norm of gradients: 0.05071702047237355
12 norm of weights: 2.297480861346769
-----
Iteration Number: 9966
Loss: 10.653092067185986
12 norm of gradients: 0.05071143491250173
12 norm of weights: 2.297519164268119
-----
Iteration Number: 9967
Loss: 10.652981454285301
12 norm of gradients: 0.05070585088924432
12 norm of weights: 2.297557468476596
-----
Iteration Number: 9968
Loss: 10.652870886733199
12 norm of gradients: 0.05070026840192289
12 norm of weights: 2.2975957739699893
-----
Iteration Number: 9969
Loss: 10.65276036450391
12 norm of gradients: 0.05069468744985933
12 norm of weights: 2.2976340807460898
-----
Iteration Number: 9970
Loss: 10.652649887571677
12 norm of gradients: 0.05068910803237597
12 norm of weights: 2.297672388802689
-----
Iteration Number: 9971
Loss: 10.65253945591075
12 norm of gradients: 0.050683530148795335
12 norm of weights: 2.297710698137582
-----
Iteration Number: 9972
```

Loss: 10.652429069495412
l2 norm of gradients: 0.05067795379844041
l2 norm of weights: 2.2977490087485632

Iteration Number: 9973
Loss: 10.65231872829995
l2 norm of gradients: 0.05067237898063451
l2 norm of weights: 2.2977873206334314

Iteration Number: 9974
Loss: 10.652208432298673
l2 norm of gradients: 0.05066680569470123
l2 norm of weights: 2.297825633789985

Iteration Number: 9975
Loss: 10.65209818146591
l2 norm of gradients: 0.05066123393996456
l2 norm of weights: 2.2978639482160252

Iteration Number: 9976
Loss: 10.651987975776004
l2 norm of gradients: 0.05065566371574883
l2 norm of weights: 2.297902263909355

Iteration Number: 9977
Loss: 10.651877815203312
l2 norm of gradients: 0.05065009502137872
l2 norm of weights: 2.2979405808677784

Iteration Number: 9978
Loss: 10.651767699722214
l2 norm of gradients: 0.050644527856179206
l2 norm of weights: 2.2979788990891015

Iteration Number: 9979
Loss: 10.6516576293071
l2 norm of gradients: 0.05063896221947564
l2 norm of weights: 2.298017218571132

Iteration Number: 9980
Loss: 10.651547603932384
l2 norm of gradients: 0.050633398110593704
l2 norm of weights: 2.2980555393116804

Iteration Number: 9981
Loss: 10.651437623572495
l2 norm of gradients: 0.050627835528859445
l2 norm of weights: 2.2980938613085566

```
-----  
Iteration Number: 9982  
Loss: 10.651327688201874  
l2 norm of gradients: 0.05062227447359924  
l2 norm of weights: 2.298132184559574  
-----  
Iteration Number: 9983  
Loss: 10.651217797794986  
l2 norm of gradients: 0.050616714944139767  
l2 norm of weights: 2.2981705090625484  
-----  
Iteration Number: 9984  
Loss: 10.651107952326305  
l2 norm of gradients: 0.05061115693980814  
l2 norm of weights: 2.298208834815295  
-----  
Iteration Number: 9985  
Loss: 10.650998151770331  
l2 norm of gradients: 0.050605600459931725  
l2 norm of weights: 2.298247161815632  
-----  
Iteration Number: 9986  
Loss: 10.650888396101575  
l2 norm of gradients: 0.050600045503838245  
l2 norm of weights: 2.29828549006138  
-----  
Iteration Number: 9987  
Loss: 10.650778685294561  
l2 norm of gradients: 0.05059449207085581  
l2 norm of weights: 2.298323819550359  
-----  
Iteration Number: 9988  
Loss: 10.650669019323841  
l2 norm of gradients: 0.050588940160312824  
l2 norm of weights: 2.2983621502803944  
-----  
Iteration Number: 9989  
Loss: 10.650559398163976  
l2 norm of gradients: 0.050583389771538034  
l2 norm of weights: 2.2984004822493094  
-----  
Iteration Number: 9990  
Loss: 10.650449821789547  
l2 norm of gradients: 0.05057784090386057  
l2 norm of weights: 2.298438815454931  
-----  
Iteration Number: 9991  
Loss: 10.650340290175144
```



```

l2 norm of gradients: 0.050572293556609825
l2 norm of weights: 2.2984771498950884
-----
Iteration Number: 9992
Loss: 10.650230803295386
l2 norm of gradients: 0.05056674772911564
l2 norm of weights: 2.2985154855676107
-----
Iteration Number: 9993
Loss: 10.650121361124903
l2 norm of gradients: 0.05056120342070805
l2 norm of weights: 2.29855382247033
-----
Iteration Number: 9994
Loss: 10.650011963638338
l2 norm of gradients: 0.05055566063071759
l2 norm of weights: 2.2985921606010797
-----
Iteration Number: 9995
Loss: 10.649902610810358
l2 norm of gradients: 0.05055011935847504
l2 norm of weights: 2.298630499957695
-----
Iteration Number: 9996
Loss: 10.64979330261564
l2 norm of gradients: 0.05054457960331151
l2 norm of weights: 2.298668840538013
-----
Iteration Number: 9997
Loss: 10.649684039028884
l2 norm of gradients: 0.050539041364558494
l2 norm of weights: 2.298707182339872
-----
Iteration Number: 9998
Loss: 10.649574820024798
l2 norm of gradients: 0.0505335046415478
l2 norm of weights: 2.2987455253611118
-----
Iteration Number: 9999
Loss: 10.64946564557812
l2 norm of gradients: 0.050527969433611625
l2 norm of weights: 2.2987838695995753
-----

```

```
[108]: print("Train accuracy: ", accuracy(xtrain_normal, ytrain, model))
```

```
Train accuracy: 0.9265384615384615
```

```
[109]: #grid search for finding the best hyperparams and model
```

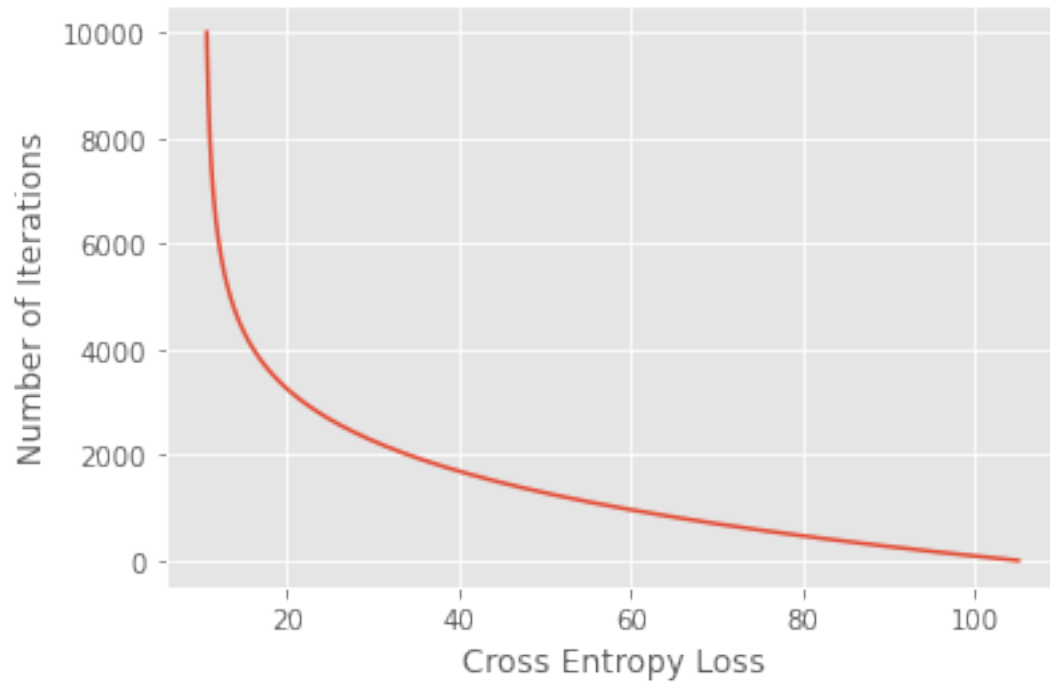
```
best_model = None
best_val = -1
for lr in [0.01, 0.001, 0.0001, 0.00001]:
    for la in [5, 2, 1, 0.1, 0.01]:
        model = fit(xtrain_normal, ytrain, lr, la, 10000, verbose=0)
        val_acc = accuracy(xval_normal, yval, model)
        print(lr, la, val_acc)
        if val_acc > best_val:
            best_val = val_acc
            best_model = model
```

```
0.01 5 0.9050949050949051
0.01 2 0.9200799200799201
0.01 1 0.9300699300699301
0.01 0.1 0.9140859140859141
0.01 0.01 0.9240759240759241
0.001 5 0.9210789210789211
0.001 2 0.9250749250749251
0.001 1 0.8751248751248751
0.001 0.1 0.9210789210789211
0.001 0.01 0.8981018981018981
0.0001 5 0.9210789210789211
0.0001 2 0.9120879120879121
0.0001 1 0.922077922077922
0.0001 0.1 0.8721278721278721
0.0001 0.01 0.9210789210789211
1e-05 5 0.8891108891108891
1e-05 2 0.8551448551448552
1e-05 1 0.926073926073926
1e-05 0.1 0.9140859140859141
1e-05 0.01 0.9180819180819181
```

```
[110]: print("Test accuracy: ", accuracy(xtest_normal, ytest, best_model))
```

```
Test accuracy: 0.937
```

```
[112]: fig = plt.figure()
ax = fig.add_subplot()
ax.set_xlabel('Cross Entropy Loss')
ax.set_ylabel('Number of Iterations')
ax.plot(cost_list, np.arange(max_iter))
plt.show()
```



[]: