

logistic-regression

February 25, 2022

Name: Rahul Govindkumar

```
[1]: 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.

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

We can convert a list into numpy array (tensor)

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

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

We can check the dimensions of the array

```
[4]: a.shape
```

```
[4]: (2, 3)
```

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

```
[5]: a * 3
```

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

You can transpose a tensor

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

(3, 2)

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

You can apply aggregate functions on the whole tensor

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

```
[7]: 24
```

or on one dimension of it

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

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

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

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

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

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

```
[10]: 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.

```
[11]: 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

```
[12]: # 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

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

(2, 3)

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

Now we can multiply each column of a by b:

```
[14]: a * b_repeat
```

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

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

```
[15]: 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

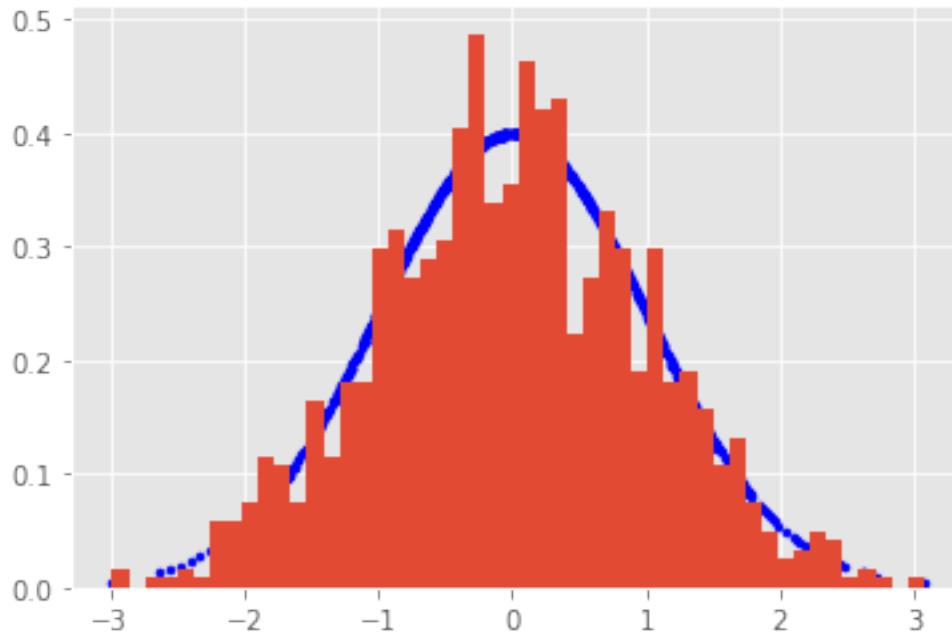
```
[16]: #implementation of Normal distribution
      def normal(x, mu, sigma):
          return np.exp( -0.5 * ((x-mu)/sigma)**2)/np.sqrt(2.0*np.pi*sigma**2)

      #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.

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

```
[17]: <matplotlib.collections.PathCollection at 0x21d39eb53a0>
```



```
[18]: 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

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

```
[20]: 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)

```
[21]: 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.

```
[22]: 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

```
[23]: 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 $N(0,1)$

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

```
[25]: #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

```
[26]: #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

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

```
[27]: 7.464679606069855
```

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

```
[28]: #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.
↪0000000000000000001)) + (1-y_true)*np.log(1-y_prob + 0.000000000000000001))
↪+ lambda_*0.5*np.linalg.norm(w)

    return cross_entropy_Loss
```

```
[29]: #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)
```

```
[30]: cost_list=[]
```

```
[31]: #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_)

    weights = weights - learning_rate * dw
    bias = bias - learning_rate * db

    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)

```

```

[32]: 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]

```

```

[33]: 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: 165.08992278766098
l2 norm of gradients: 1.5712841571143472
l2 norm of weights: 6.749898914526341
-----
Iteration Number: 1
Loss: 164.96958129222804
l2 norm of gradients: 1.5706966434456426
l2 norm of weights: 6.749433510757873

```

```
-----  
Iteration Number: 2  
Loss: 164.8493290889799  
l2 norm of gradients: 1.5701090226134353  
l2 norm of weights: 6.74896845133604  
-----  
Iteration Number: 3  
Loss: 164.7291661691766  
l2 norm of gradients: 1.5695212957976998  
l2 norm of weights: 6.748503736038181  
-----  
Iteration Number: 4  
Loss: 164.6090925240257  
l2 norm of gradients: 1.5689334641718993  
l2 norm of weights: 6.748039364641625  
-----  
Iteration Number: 5  
Loss: 164.48910814463196  
l2 norm of gradients: 1.5683455289030148  
l2 norm of weights: 6.747575336923698  
-----  
Iteration Number: 6  
Loss: 164.36921302180806  
l2 norm of gradients: 1.5677574911515728  
l2 norm of weights: 6.747111652661718  
-----  
Iteration Number: 7  
Loss: 164.24940714573742  
l2 norm of gradients: 1.567169352071679  
l2 norm of weights: 6.746648311633001  
-----  
Iteration Number: 8  
Loss: 164.12969050669403  
l2 norm of gradients: 1.5665811128110474  
l2 norm of weights: 6.746185313614855  
-----  
Iteration Number: 9  
Loss: 164.01006309459927  
l2 norm of gradients: 1.565992774511037  
l2 norm of weights: 6.745722658384585  
-----  
Iteration Number: 10  
Loss: 163.8905248989357  
l2 norm of gradients: 1.5654043383066818  
l2 norm of weights: 6.745260345719499  
-----  
Iteration Number: 11  
Loss: 163.77107590929776
```


l2 norm of gradients: 1.5648158053267311
l2 norm of weights: 6.7447983753969

Iteration Number: 12
Loss: 163.65171611449006
l2 norm of gradients: 1.5642271766936815
l2 norm of weights: 6.744336747194092

Iteration Number: 13
Loss: 163.53244550359398
l2 norm of gradients: 1.563638453523817
l2 norm of weights: 6.743875460888381

Iteration Number: 14
Loss: 163.41326406514648
l2 norm of gradients: 1.563049636927246
l2 norm of weights: 6.743414516257074

Iteration Number: 15
Loss: 163.29417178735827
l2 norm of gradients: 1.5624607280079403
l2 norm of weights: 6.742953913077481

Iteration Number: 16
Loss: 163.17516865843484
l2 norm of gradients: 1.5618717278637757
l2 norm of weights: 6.742493651126916

Iteration Number: 17
Loss: 163.05625466623042
l2 norm of gradients: 1.5612826375865716
l2 norm of weights: 6.7420337301826985

Iteration Number: 18
Loss: 162.93742979803687
l2 norm of gradients: 1.5606934582621332
l2 norm of weights: 6.741574150022154

Iteration Number: 19
Loss: 162.81869404145257
l2 norm of gradients: 1.5601041909702935
l2 norm of weights: 6.741114910422614

Iteration Number: 20
Loss: 162.7000473834082
l2 norm of gradients: 1.5595148367849556
l2 norm of weights: 6.740656011161419

Iteration Number: 21
Loss: 162.581489810618
l2 norm of gradients: 1.5589253967741357
l2 norm of weights: 6.740197452015916

Iteration Number: 22
Loss: 162.46302130988988
l2 norm of gradients: 1.5583358720000084
l2 norm of weights: 6.739739232763462

Iteration Number: 23
Loss: 162.3446418675239
l2 norm of gradients: 1.5577462635189494
l2 norm of weights: 6.739281353181427

Iteration Number: 24
Loss: 162.226351469332
l2 norm of gradients: 1.557156572381581
l2 norm of weights: 6.738823813047192

Iteration Number: 25
Loss: 162.1081501013587
l2 norm of gradients: 1.5565667996328192
l2 norm of weights: 6.7383666121381465

Iteration Number: 26
Loss: 161.99003774907868
l2 norm of gradients: 1.555976946311917
l2 norm of weights: 6.7379097502316965

Iteration Number: 27
Loss: 161.87201439798503
l2 norm of gradients: 1.5553870134525123
l2 norm of weights: 6.737453227105262

Iteration Number: 28
Loss: 161.7540800330982
l2 norm of gradients: 1.5547970020826742
l2 norm of weights: 6.7369970425362755

Iteration Number: 29
Loss: 161.63623463927613
l2 norm of gradients: 1.5542069132249503
l2 norm of weights: 6.7365411963021895

Iteration Number: 30
Loss: 161.51847820119144
l2 norm of gradients: 1.553616747896414

```

12 norm of weights: 6.736085688180468
-----
Iteration Number: 31
Loss: 161.40081070342282
12 norm of gradients: 1.5530265071087113
12 norm of weights: 6.735630517948597
-----
Iteration Number: 32
Loss: 161.2832321298334
12 norm of gradients: 1.5524361918681109
12 norm of weights: 6.735175685384079
-----
Iteration Number: 33
Loss: 161.16574246470014
12 norm of gradients: 1.5518458031755504
12 norm of weights: 6.734721190264436
-----
Iteration Number: 34
Loss: 161.04834169143888
12 norm of gradients: 1.5512553420266861
12 norm of weights: 6.7342670323672085
-----
Iteration Number: 35
Loss: 160.93102979359915
12 norm of gradients: 1.550664809411942
12 norm of weights: 6.733813211469962
-----
Iteration Number: 36
Loss: 160.81380675469694
12 norm of gradients: 1.5500742063165565
12 norm of weights: 6.733359727350279
-----
Iteration Number: 37
Loss: 160.69667255736056
12 norm of gradients: 1.549483533720635
12 norm of weights: 6.732906579785767
-----
Iteration Number: 38
Loss: 160.5796271847534
12 norm of gradients: 1.548892792599196
12 norm of weights: 6.732453768554058
-----
Iteration Number: 39
Loss: 160.4626706193951
12 norm of gradients: 1.5483019839222238
12 norm of weights: 6.732001293432806
-----
Iteration Number: 40

```

Loss: 160.34580284343997
l2 norm of gradients: 1.5477111086547144
l2 norm of weights: 6.731549154199692

Iteration Number: 41
Loss: 160.22902383932086
l2 norm of gradients: 1.5471201677567281
l2 norm of weights: 6.731097350632423

Iteration Number: 42
Loss: 160.11233358868444
l2 norm of gradients: 1.546529162183439
l2 norm of weights: 6.730645882508728

Iteration Number: 43
Loss: 159.99573207357665
l2 norm of gradients: 1.5459380928851834
l2 norm of weights: 6.730194749606371

Iteration Number: 44
Loss: 159.87921927517792
l2 norm of gradients: 1.5453469608075114
l2 norm of weights: 6.729743951703138

Iteration Number: 45
Loss: 159.7627951749452
l2 norm of gradients: 1.5447557668912364
l2 norm of weights: 6.729293488576847

Iteration Number: 46
Loss: 159.64645975398778
l2 norm of gradients: 1.5441645120724852
l2 norm of weights: 6.728843360005346

Iteration Number: 47
Loss: 159.53021299301133
l2 norm of gradients: 1.5435731972827478
l2 norm of weights: 6.728393565766512

Iteration Number: 48
Loss: 159.41405487270887
l2 norm of gradients: 1.542981823448929
l2 norm of weights: 6.727944105638256

Iteration Number: 49
Loss: 159.29798537362254
l2 norm of gradients: 1.542390391493397
l2 norm of weights: 6.727494979398519

```
-----
Iteration Number: 50
Loss: 159.18200447603593
l2 norm of gradients: 1.5417989023340344
l2 norm of weights: 6.727046186825276
-----
Iteration Number: 51
Loss: 159.0661121596922
l2 norm of gradients: 1.5412073568842892
l2 norm of weights: 6.726597727696533
-----
Iteration Number: 52
Loss: 158.95030840460035
l2 norm of gradients: 1.5406157560532239
l2 norm of weights: 6.726149601790336
-----
Iteration Number: 53
Loss: 158.83459319045713
l2 norm of gradients: 1.5400241007455662
l2 norm of weights: 6.725701808884763
-----
Iteration Number: 54
Loss: 158.7189664965225
l2 norm of gradients: 1.539432391861759
l2 norm of weights: 6.725254348757925
-----
Iteration Number: 55
Loss: 158.60342830204382
l2 norm of gradients: 1.538840630298012
l2 norm of weights: 6.7248072211879775
-----
Iteration Number: 56
Loss: 158.48797858613568
l2 norm of gradients: 1.5382488169463493
l2 norm of weights: 6.724360425953105
-----
Iteration Number: 57
Loss: 158.3726173274307
l2 norm of gradients: 1.5376569526946626
l2 norm of weights: 6.723913962831536
-----
Iteration Number: 58
Loss: 158.2573445047843
l2 norm of gradients: 1.5370650384267583
l2 norm of weights: 6.723467831601536
-----
Iteration Number: 59
Loss: 158.14216009638704
```

```

12 norm of gradients: 1.53647307502241
12 norm of weights: 6.723022032041412
-----
Iteration Number: 60
Loss: 158.0270640806605
12 norm of gradients: 1.535881063357407
12 norm of weights: 6.722576563929508
-----
Iteration Number: 61
Loss: 157.9120564354215
12 norm of gradients: 1.5352890043036045
12 norm of weights: 6.7221314270442125
-----
Iteration Number: 62
Loss: 157.79713713874997
12 norm of gradients: 1.5346968987289742
12 norm of weights: 6.721686621163953
-----
Iteration Number: 63
Loss: 157.68230616818164
12 norm of gradients: 1.5341047474976524
12 norm of weights: 6.721242146067203
-----
Iteration Number: 64
Loss: 157.5675635012241
12 norm of gradients: 1.533512551469992
12 norm of weights: 6.7207980015324775
-----
Iteration Number: 65
Loss: 157.45290911507098
12 norm of gradients: 1.5329203115026098
12 norm of weights: 6.720354187338336
-----
Iteration Number: 66
Loss: 157.33834298693617
12 norm of gradients: 1.5323280284484364
12 norm of weights: 6.719910703263382
-----
Iteration Number: 67
Loss: 157.2238650935418
12 norm of gradients: 1.5317357031567673
12 norm of weights: 6.719467549086265
-----
Iteration Number: 68
Loss: 157.10947541171623
12 norm of gradients: 1.531143336473309
12 norm of weights: 6.719024724585682
-----

```

Iteration Number: 69
Loss: 156.99517391808175
l2 norm of gradients: 1.5305509292402308
l2 norm of weights: 6.718582229540376

Iteration Number: 70
Loss: 156.88096058878114
l2 norm of gradients: 1.5299584822962125
l2 norm of weights: 6.718140063729136

Iteration Number: 71
Loss: 156.76683540005627
l2 norm of gradients: 1.5293659964764923
l2 norm of weights: 6.717698226930802

Iteration Number: 72
Loss: 156.65279832786504
l2 norm of gradients: 1.5287734726129176
l2 norm of weights: 6.717256718924261

Iteration Number: 73
Loss: 156.53884934816153
l2 norm of gradients: 1.5281809115339906
l2 norm of weights: 6.716815539488453

Iteration Number: 74
Loss: 156.42498843648198
l2 norm of gradients: 1.5275883140649191
l2 norm of weights: 6.716374688402364

Iteration Number: 75
Loss: 156.31121556812403
l2 norm of gradients: 1.5269956810276615
l2 norm of weights: 6.7159341654450335

Iteration Number: 76
Loss: 156.19753071864884
l2 norm of gradients: 1.5264030132409776
l2 norm of weights: 6.715493970395554

Iteration Number: 77
Loss: 156.0839338629171
l2 norm of gradients: 1.5258103115204742
l2 norm of weights: 6.715054103033068

Iteration Number: 78
Loss: 155.97042497611227
l2 norm of gradients: 1.5252175766786524

```

12 norm of weights: 6.714614563136772
-----
Iteration Number: 79
Loss: 155.85700403260807
12 norm of gradients: 1.5246248095249555
12 norm of weights: 6.714175350485918
-----
Iteration Number: 80
Loss: 155.7436710074066
12 norm of gradients: 1.524032010865814
12 norm of weights: 6.713736464859809
-----
Iteration Number: 81
Loss: 155.63042587451366
12 norm of gradients: 1.5234391815046946
12 norm of weights: 6.713297906037808
-----
Iteration Number: 82
Loss: 155.51726860860754
12 norm of gradients: 1.5228463222421427
12 norm of weights: 6.712859673799331
-----
Iteration Number: 83
Loss: 155.40419918328695
12 norm of gradients: 1.5222534338758313
12 norm of weights: 6.712421767923849
-----
Iteration Number: 84
Loss: 155.29121757285998
12 norm of gradients: 1.5216605172006052
12 norm of weights: 6.711984188190894
-----
Iteration Number: 85
Loss: 155.1783237507417
12 norm of gradients: 1.5210675730085241
12 norm of weights: 6.711546934380054
-----
Iteration Number: 86
Loss: 155.06551769075284
12 norm of gradients: 1.5204746020889104
12 norm of weights: 6.711110006270974
-----
Iteration Number: 87
Loss: 154.95279936626457
12 norm of gradients: 1.5198816052283903
12 norm of weights: 6.710673403643361
-----
Iteration Number: 88

```



```
Loss: 154.84016875044284
l2 norm of gradients: 1.5192885832109397
l2 norm of weights: 6.710237126276981
-----
Iteration Number: 89
Loss: 154.72762581630093
l2 norm of gradients: 1.5186955368179251
l2 norm of weights: 6.709801173951661
-----
Iteration Number: 90
Loss: 154.61517053687675
l2 norm of gradients: 1.518102466828149
l2 norm of weights: 6.709365546447288
-----
Iteration Number: 91
Loss: 154.5028028849822
l2 norm of gradients: 1.5175093740178895
l2 norm of weights: 6.70893024354381
-----
Iteration Number: 92
Loss: 154.39052283306305
l2 norm of gradients: 1.516916259160943
l2 norm of weights: 6.708495265021241
-----
Iteration Number: 93
Loss: 154.27833035355707
l2 norm of gradients: 1.5163231230286642
l2 norm of weights: 6.708060610659655
-----
Iteration Number: 94
Loss: 154.16622541892693
l2 norm of gradients: 1.5157299663900086
l2 norm of weights: 6.707626280239191
-----
Iteration Number: 95
Loss: 154.05420800114223
l2 norm of gradients: 1.515136790011569
l2 norm of weights: 6.707192273540053
-----
Iteration Number: 96
Loss: 153.94227807236902
l2 norm of gradients: 1.5145435946576176
l2 norm of weights: 6.706758590342511
-----
Iteration Number: 97
Loss: 153.8304356040997
l2 norm of gradients: 1.5139503810901418
l2 norm of weights: 6.706325230426897
```

```

-----
Iteration Number: 98
Loss: 153.71868056830277
l2 norm of gradients: 1.5133571500688832
l2 norm of weights: 6.705892193573613
-----
Iteration Number: 99
Loss: 153.60701293625243
l2 norm of gradients: 1.512763902351373
l2 norm of weights: 6.705459479563127
-----
Iteration Number: 100
Loss: 153.49543267942067
l2 norm of gradients: 1.51217063869297
l2 norm of weights: 6.7050270881759735
-----
Iteration Number: 101
Loss: 153.38393976904862
l2 norm of gradients: 1.5115773598468942
l2 norm of weights: 6.70459501919276
-----
Iteration Number: 102
Loss: 153.27253417613974
l2 norm of gradients: 1.5109840665642595
l2 norm of weights: 6.704163272394155
-----
Iteration Number: 103
Loss: 153.16121587154998
l2 norm of gradients: 1.5103907595941104
l2 norm of weights: 6.7037318475609045
-----
Iteration Number: 104
Loss: 153.04998482615528
l2 norm of gradients: 1.509797439683451
l2 norm of weights: 6.703300744473821
-----
Iteration Number: 105
Loss: 152.93884101052765
l2 norm of gradients: 1.5092041075772789
l2 norm of weights: 6.702869962913788
-----
Iteration Number: 106
Loss: 152.82778439492748
l2 norm of gradients: 1.508610764018613
l2 norm of weights: 6.702439502661761
-----
Iteration Number: 107
Loss: 152.7168149499206

```

```

12 norm of gradients: 1.5080174097485237
12 norm of weights: 6.702009363498768
-----
Iteration Number: 108
Loss: 152.60593264551278
12 norm of gradients: 1.5074240455061614
12 norm of weights: 6.70157954520591
-----
Iteration Number: 109
Loss: 152.49513745181252
12 norm of gradients: 1.506830672028783
12 norm of weights: 6.701150047564361
-----
Iteration Number: 110
Loss: 152.3844293388084
12 norm of gradients: 1.5062372900517758
12 norm of weights: 6.70072087035537
-----
Iteration Number: 111
Loss: 152.27380827597787
12 norm of gradients: 1.5056439003086843
12 norm of weights: 6.700292013360259
-----
Iteration Number: 112
Loss: 152.1632742330538
12 norm of gradients: 1.5050505035312318
12 norm of weights: 6.699863476360427
-----
Iteration Number: 113
Loss: 152.05282717948558
12 norm of gradients: 1.504457100449343
12 norm of weights: 6.699435259137348
-----
Iteration Number: 114
Loss: 151.9424670844681
12 norm of gradients: 1.5038636917911616
12 norm of weights: 6.699007361472575
-----
Iteration Number: 115
Loss: 151.83219391732646
12 norm of gradients: 1.5032702782830738
12 norm of weights: 6.6985797831477365
-----
Iteration Number: 116
Loss: 151.72200764712008
12 norm of gradients: 1.502676860649721
12 norm of weights: 6.698152523944537
-----

```

Iteration Number: 117
Loss: 151.6119082425951
l2 norm of gradients: 1.5020834396140184
l2 norm of weights: 6.697725583644763

Iteration Number: 118
Loss: 151.5018956726456
l2 norm of gradients: 1.5014900158971685
l2 norm of weights: 6.6972989620302785

Iteration Number: 119
Loss: 151.3919699057718
l2 norm of gradients: 1.500896590218673
l2 norm of weights: 6.69687265888303

Iteration Number: 120
Loss: 151.28213091047013
l2 norm of gradients: 1.5003031632963455
l2 norm of weights: 6.69644667398504

Iteration Number: 121
Loss: 151.17237865508346
l2 norm of gradients: 1.499709735846319
l2 norm of weights: 6.696021007118414

Iteration Number: 122
Loss: 151.06271310800173
l2 norm of gradients: 1.4991163085830541
l2 norm of weights: 6.6955956580653435

Iteration Number: 123
Loss: 150.95313423698718
l2 norm of gradients: 1.4985228822193446
l2 norm of weights: 6.695170626608095

Iteration Number: 124
Loss: 150.84364201020088
l2 norm of gradients: 1.4979294574663218
l2 norm of weights: 6.694745912529023

Iteration Number: 125
Loss: 150.73423639539737
l2 norm of gradients: 1.497336035033456
l2 norm of weights: 6.694321515610566

Iteration Number: 126
Loss: 150.62491736032356
l2 norm of gradients: 1.496742615628556

```

12 norm of weights: 6.693897435635244
-----
Iteration Number: 127
Loss: 150.51568487250216
12 norm of gradients: 1.4961491999577679
12 norm of weights: 6.693473672385664
-----
Iteration Number: 128
Loss: 150.40653889923078
12 norm of gradients: 1.4955557887255715
12 norm of weights: 6.693050225644518
-----
Iteration Number: 129
Loss: 150.29747940794874
12 norm of gradients: 1.494962382634773
12 norm of weights: 6.692627095194583
-----
Iteration Number: 130
Loss: 150.1885063657048
12 norm of gradients: 1.4943689823864983
12 norm of weights: 6.692204280818726
-----
Iteration Number: 131
Loss: 150.07961973966354
12 norm of gradients: 1.4937755886801798
12 norm of weights: 6.691781782299898
-----
Iteration Number: 132
Loss: 149.97081949661083
12 norm of gradients: 1.4931822022135484
12 norm of weights: 6.691359599421143
-----
Iteration Number: 133
Loss: 149.86210560331264
12 norm of gradients: 1.4925888236826144
12 norm of weights: 6.690937731965587
-----
Iteration Number: 134
Loss: 149.75347802646283
12 norm of gradients: 1.4919954537816524
12 norm of weights: 6.690516179716453
-----
Iteration Number: 135
Loss: 149.644936732731
12 norm of gradients: 1.4914020932031822
12 norm of weights: 6.690094942457049
-----
Iteration Number: 136

```

```
Loss: 149.5364816882266
l2 norm of gradients: 1.4908087426379464
l2 norm of weights: 6.689674019970774
-----
Iteration Number: 137
Loss: 149.42811285955915
l2 norm of gradients: 1.490215402774886
l2 norm of weights: 6.689253412041121
-----
Iteration Number: 138
Loss: 149.31983021255914
l2 norm of gradients: 1.4896220743011146
l2 norm of weights: 6.688833118451675
-----
Iteration Number: 139
Loss: 149.21163371348624
l2 norm of gradients: 1.489028757901889
l2 norm of weights: 6.68841313898611
-----
Iteration Number: 140
Loss: 149.10352332816416
l2 norm of gradients: 1.4884354542605767
l2 norm of weights: 6.687993473428198
-----
Iteration Number: 141
Loss: 148.9954990223427
l2 norm of gradients: 1.4878421640586235
l2 norm of weights: 6.687574121561801
-----
Iteration Number: 142
Loss: 148.88756076177452
l2 norm of gradients: 1.4872488879755157
l2 norm of weights: 6.687155083170879
-----
Iteration Number: 143
Loss: 148.77970851207994
l2 norm of gradients: 1.4866556266887405
l2 norm of weights: 6.686736358039485
-----
Iteration Number: 144
Loss: 148.67194223853306
l2 norm of gradients: 1.4860623808737454
l2 norm of weights: 6.68631794595177
-----
Iteration Number: 145
Loss: 148.5642619064371
l2 norm of gradients: 1.4854691512038924
l2 norm of weights: 6.68589984669198
```

```

-----
Iteration Number: 146
Loss: 148.45666748117134
l2 norm of gradients: 1.4848759383504109
l2 norm of weights: 6.685482060044458
-----
Iteration Number: 147
Loss: 148.34915892766531
l2 norm of gradients: 1.4842827429823484
l2 norm of weights: 6.685064585793648
-----
Iteration Number: 148
Loss: 148.24173621088624
l2 norm of gradients: 1.483689565766517
l2 norm of weights: 6.684647423724088
-----
Iteration Number: 149
Loss: 148.13439929578453
l2 norm of gradients: 1.4830964073674382
l2 norm of weights: 6.684230573620419
-----
Iteration Number: 150
Loss: 148.02714814697754
l2 norm of gradients: 1.4825032684472847
l2 norm of weights: 6.68381403526738
-----
Iteration Number: 151
Loss: 147.91998272910615
l2 norm of gradients: 1.4819101496658198
l2 norm of weights: 6.683397808449811
-----
Iteration Number: 152
Loss: 147.8129030067472
l2 norm of gradients: 1.481317051680332
l2 norm of weights: 6.682981892952655
-----
Iteration Number: 153
Loss: 147.7059089441303
l2 norm of gradients: 1.4807239751455696
l2 norm of weights: 6.6825662885609525
-----
Iteration Number: 154
Loss: 147.5990005057137
l2 norm of gradients: 1.480130920713672
l2 norm of weights: 6.682150995059851
-----
Iteration Number: 155
Loss: 147.4921776556354

```

l2 norm of gradients: 1.4795378890340942
l2 norm of weights: 6.6817360122346

Iteration Number: 156
Loss: 147.3854403580364
l2 norm of gradients: 1.4789448807535353
l2 norm of weights: 6.681321339870551

Iteration Number: 157
Loss: 147.27878857670012
l2 norm of gradients: 1.478351896515857
l2 norm of weights: 6.680906977753161

Iteration Number: 158
Loss: 147.17222227545776
l2 norm of gradients: 1.477758936962006
l2 norm of weights: 6.680492925667995

Iteration Number: 159
Loss: 147.06574141824862
l2 norm of gradients: 1.4771660027299274
l2 norm of weights: 6.680079183400719

Iteration Number: 160
Loss: 146.9593459686978
l2 norm of gradients: 1.4765730944544782
l2 norm of weights: 6.679665750737109

Iteration Number: 161
Loss: 146.85303589004505
l2 norm of gradients: 1.4759802127673396
l2 norm of weights: 6.679252627463047

Iteration Number: 162
Loss: 146.74681114588756
l2 norm of gradients: 1.4753873582969232
l2 norm of weights: 6.678839813364523

Iteration Number: 163
Loss: 146.64067169967308
l2 norm of gradients: 1.4747945316682758
l2 norm of weights: 6.6784273082276355

Iteration Number: 164
Loss: 146.53461751449572
l2 norm of gradients: 1.4742017335029811
l2 norm of weights: 6.678015111838592

Iteration Number: 165
Loss: 146.4286485533279
l2 norm of gradients: 1.4736089644190595
l2 norm of weights: 6.677603223983709

Iteration Number: 166
Loss: 146.32276477945734
l2 norm of gradients: 1.4730162250308632
l2 norm of weights: 6.6771916444449416

Iteration Number: 167
Loss: 146.216966155558
l2 norm of gradients: 1.4724235159489703
l2 norm of weights: 6.67678037302225

Iteration Number: 168
Loss: 146.11125264459
l2 norm of gradients: 1.4718308377800755
l2 norm of weights: 6.676369409488865

Iteration Number: 169
Loss: 146.00562420921938
l2 norm of gradients: 1.4712381911268773
l2 norm of weights: 6.675958753636021

Iteration Number: 170
Loss: 145.90008081213105
l2 norm of gradients: 1.470645576587964
l2 norm of weights: 6.675548405250596

Iteration Number: 171
Loss: 145.79462241565577
l2 norm of gradients: 1.4700529947576952
l2 norm of weights: 6.6751383641195785

Iteration Number: 172
Loss: 145.6892489822811
l2 norm of gradients: 1.4694604462260825
l2 norm of weights: 6.674728630030074

Iteration Number: 173
Loss: 145.58396047444407
l2 norm of gradients: 1.4688679315786652
l2 norm of weights: 6.674319202769302

Iteration Number: 174
Loss: 145.47875685426402
l2 norm of gradients: 1.4682754513963858

```

12 norm of weights: 6.673910082124599
-----
Iteration Number: 175
Loss: 145.37363808400286
12 norm of gradients: 1.4676830062554616
12 norm of weights: 6.673501267883416
-----
Iteration Number: 176
Loss: 145.26860412559807
12 norm of gradients: 1.4670905967272536
12 norm of weights: 6.67309275983332
-----
Iteration Number: 177
Loss: 145.16365494100785
12 norm of gradients: 1.4664982233781345
12 norm of weights: 6.672684557761999
-----
Iteration Number: 178
Loss: 145.05879049209904
12 norm of gradients: 1.4659058867693509
12 norm of weights: 6.672276661457255
-----
Iteration Number: 179
Loss: 144.95401074069073
12 norm of gradients: 1.465313587456888
12 norm of weights: 6.671869070707015
-----
Iteration Number: 180
Loss: 144.84931564844734
12 norm of gradients: 1.4647213259913259
12 norm of weights: 6.671461785299321
-----
Iteration Number: 181
Loss: 144.74470517697338
12 norm of gradients: 1.4641291029177002
12 norm of weights: 6.671054805022336
-----
Iteration Number: 182
Loss: 144.6401792878341
12 norm of gradients: 1.463536918775354
12 norm of weights: 6.670648129664344
-----
Iteration Number: 183
Loss: 144.53573794243377
12 norm of gradients: 1.462944774097793
12 norm of weights: 6.67024175901375
-----
Iteration Number: 184

```

```
Loss: 144.43138110202025
l2 norm of gradients: 1.4623526694125344
l2 norm of weights: 6.669835692859084
-----
Iteration Number: 185
Loss: 144.327108727971
l2 norm of gradients: 1.4617606052409566
l2 norm of weights: 6.6694299309889935
-----
Iteration Number: 186
Loss: 144.2229207814986
l2 norm of gradients: 1.4611685820981446
l2 norm of weights: 6.6690244731922546
-----
Iteration Number: 187
Loss: 144.11881722357825
l2 norm of gradients: 1.4605766004927359
l2 norm of weights: 6.668619319257765
-----
Iteration Number: 188
Loss: 144.01479801526835
l2 norm of gradients: 1.4599846609267613
l2 norm of weights: 6.668214468974544
-----
Iteration Number: 189
Loss: 143.91086311762618
l2 norm of gradients: 1.4593927638954873
l2 norm of weights: 6.6678099221317435
-----
Iteration Number: 190
Loss: 143.80701249154265
l2 norm of gradients: 1.458800909887254
l2 norm of weights: 6.667405678518635
-----
Iteration Number: 191
Loss: 143.70324609757793
l2 norm of gradients: 1.458209099383314
l2 norm of weights: 6.6670017379246165
-----
Iteration Number: 192
Loss: 143.59956389669532
l2 norm of gradients: 1.457617332857665
l2 norm of weights: 6.666598100139217
-----
Iteration Number: 193
Loss: 143.4959658494459
l2 norm of gradients: 1.457025610776888
l2 norm of weights: 6.66619476495209
```

```
-----  
Iteration Number: 194  
Loss: 143.39245191641578  
l2 norm of gradients: 1.4564339335999765  
l2 norm of weights: 6.665791732153017  
-----  
Iteration Number: 195  
Loss: 143.2890220582479  
l2 norm of gradients: 1.4558423017781705  
l2 norm of weights: 6.665389001531911  
-----  
Iteration Number: 196  
Loss: 143.18567623514218  
l2 norm of gradients: 1.4552507157547832  
l2 norm of weights: 6.664986572878809  
-----  
Iteration Number: 197  
Loss: 143.0824144076084  
l2 norm of gradients: 1.4546591759650342  
l2 norm of weights: 6.664584445983885  
-----  
Iteration Number: 198  
Loss: 142.97923653596771  
l2 norm of gradients: 1.454067682835874  
l2 norm of weights: 6.664182620637437  
-----  
Iteration Number: 199  
Loss: 142.87614258032986  
l2 norm of gradients: 1.4534762367858107  
l2 norm of weights: 6.663781096629898  
-----  
Iteration Number: 200  
Loss: 142.7731325011353  
l2 norm of gradients: 1.4528848382247375  
l2 norm of weights: 6.663379873751829  
-----  
Iteration Number: 201  
Loss: 142.67020625828843  
l2 norm of gradients: 1.4522934875537554  
l2 norm of weights: 6.662978951793927  
-----  
Iteration Number: 202  
Loss: 142.5673638119869  
l2 norm of gradients: 1.451702185164999  
l2 norm of weights: 6.66257833054702  
-----  
Iteration Number: 203  
Loss: 142.46460512208114
```

l2 norm of gradients: 1.4511109314414583
l2 norm of weights: 6.662178009802067

Iteration Number: 204
Loss: 142.36193014867993
l2 norm of gradients: 1.450519726756803
l2 norm of weights: 6.661777989350163

Iteration Number: 205
Loss: 142.2593388515332
l2 norm of gradients: 1.4499285714752035
l2 norm of weights: 6.661378268982537

Iteration Number: 206
Loss: 142.15683119062058
l2 norm of gradients: 1.449337465951153
l2 norm of weights: 6.660978848490551

Iteration Number: 207
Loss: 142.0544071256231
l2 norm of gradients: 1.44874641052929
l2 norm of weights: 6.660579727665704

Iteration Number: 208
Loss: 141.9520666162557
l2 norm of gradients: 1.4481554055442174
l2 norm of weights: 6.6601809062996296

Iteration Number: 209
Loss: 141.84980962225634
l2 norm of gradients: 1.4475644513203265
l2 norm of weights: 6.659782384184097

Iteration Number: 210
Loss: 141.74763610333343
l2 norm of gradients: 1.446973548171616
l2 norm of weights: 6.659384161111012

Iteration Number: 211
Loss: 141.6455460188794
l2 norm of gradients: 1.446382696401515
l2 norm of weights: 6.658986236872419

Iteration Number: 212
Loss: 141.54353932871294
l2 norm of gradients: 1.4457918963027026
l2 norm of weights: 6.658588611260498

Iteration Number: 213
Loss: 141.44161599218012
l2 norm of gradients: 1.4452011481569313
l2 norm of weights: 6.658191284067567

Iteration Number: 214
Loss: 141.33977596863005
l2 norm of gradients: 1.4446104522348489
l2 norm of weights: 6.657794255086083

Iteration Number: 215
Loss: 141.23801921780333
l2 norm of gradients: 1.4440198087958205
l2 norm of weights: 6.6573975241086405

Iteration Number: 216
Loss: 141.13634569878903
l2 norm of gradients: 1.443429218087752
l2 norm of weights: 6.6570010909279755

Iteration Number: 217
Loss: 141.0347553710924
l2 norm of gradients: 1.4428386803469146
l2 norm of weights: 6.6566049553369595

Iteration Number: 218
Loss: 140.93324819390895
l2 norm of gradients: 1.442248195797769
l2 norm of weights: 6.656209117128606

Iteration Number: 219
Loss: 140.8318241266817
l2 norm of gradients: 1.4416577646527913
l2 norm of weights: 6.6558135760960715

Iteration Number: 220
Loss: 140.73048312865376
l2 norm of gradients: 1.4410673871123
l2 norm of weights: 6.655418332032647

Iteration Number: 221
Loss: 140.62922515907582
l2 norm of gradients: 1.440477063364283
l2 norm of weights: 6.655023384731769

Iteration Number: 222
Loss: 140.52805017689323
l2 norm of gradients: 1.4398867935842274

```

12 norm of weights: 6.654628733987012
-----
Iteration Number: 223
Loss: 140.42695814170497
12 norm of gradients: 1.439296577934949
12 norm of weights: 6.654234379592096
-----
Iteration Number: 224
Loss: 140.32594901225113
12 norm of gradients: 1.4387064165664254
12 norm of weights: 6.653840321340879
-----
Iteration Number: 225
Loss: 140.22502274795244
12 norm of gradients: 1.4381163096156273
12 norm of weights: 6.653446559027364
-----
Iteration Number: 226
Loss: 140.12417930780617
12 norm of gradients: 1.4375262572063559
12 norm of weights: 6.653053092445692
-----
Iteration Number: 227
Loss: 140.023418650938
12 norm of gradients: 1.4369362594490767
12 norm of weights: 6.652659921390151
-----
Iteration Number: 228
Loss: 139.92274073654005
12 norm of gradients: 1.4363463164407593
12 norm of weights: 6.6522670456551705
-----
Iteration Number: 229
Loss: 139.82214552345428
12 norm of gradients: 1.4357564282647188
12 norm of weights: 6.651874465035321
-----
Iteration Number: 230
Loss: 139.72163297100084
12 norm of gradients: 1.4351665949904557
12 norm of weights: 6.65148217932532
-----
Iteration Number: 231
Loss: 139.62120303800913
12 norm of gradients: 1.4345768166735027
12 norm of weights: 6.651090188320025
-----
Iteration Number: 232

```

```

Loss: 139.52085568362963
l2 norm of gradients: 1.433987093355269
l2 norm of weights: 6.65069849181444
-----
Iteration Number: 233
Loss: 139.4205908670787
l2 norm of gradients: 1.4333974250628914
l2 norm of weights: 6.65030708960371
-----
Iteration Number: 234
Loss: 139.32040854710579
l2 norm of gradients: 1.432807811809084
l2 norm of weights: 6.649915981483127
-----
Iteration Number: 235
Loss: 139.2203086827953
l2 norm of gradients: 1.4322182535919932
l2 norm of weights: 6.6495251672481235
-----
Iteration Number: 236
Loss: 139.12029123343143
l2 norm of gradients: 1.431628750395053
l2 norm of weights: 6.6491346466942804
-----
Iteration Number: 237
Loss: 139.02035615782953
l2 norm of gradients: 1.4310393021868444
l2 norm of weights: 6.648744419617318
-----
Iteration Number: 238
Loss: 138.9205034150339
l2 norm of gradients: 1.4304499089209577
l2 norm of weights: 6.648354485813106
-----
Iteration Number: 239
Loss: 138.82073296423988
l2 norm of gradients: 1.4298605705358556
l2 norm of weights: 6.647964845077656
-----
Iteration Number: 240
Loss: 138.7210447645651
l2 norm of gradients: 1.429271286954743
l2 norm of weights: 6.647575497207123
-----
Iteration Number: 241
Loss: 138.62143877491178
l2 norm of gradients: 1.4286820580854354
l2 norm of weights: 6.64718644199781

```



```
-----
Iteration Number: 242
Loss: 138.52191495441542
l2 norm of gradients: 1.4280928838202331
l2 norm of weights: 6.646797679246161
-----
Iteration Number: 243
Loss: 138.42247326233496
l2 norm of gradients: 1.4275037640357984
l2 norm of weights: 6.646409208748767
-----
Iteration Number: 244
Loss: 138.32311365767913
l2 norm of gradients: 1.4269146985930352
l2 norm of weights: 6.646021030302362
-----
Iteration Number: 245
Loss: 138.22383609957265
l2 norm of gradients: 1.426325687336972
l2 norm of weights: 6.645633143703825
-----
Iteration Number: 246
Loss: 138.12464054727056
l2 norm of gradients: 1.425736730096649
l2 norm of weights: 6.6452455487501805
-----
Iteration Number: 247
Loss: 138.02552696004923
l2 norm of gradients: 1.4251478266850077
l2 norm of weights: 6.644858245238597
-----
Iteration Number: 248
Loss: 137.9264952969406
l2 norm of gradients: 1.424558976898785
l2 norm of weights: 6.644471232966385
-----
Iteration Number: 249
Loss: 137.82754551742858
l2 norm of gradients: 1.4239701805184102
l2 norm of weights: 6.644084511731004
-----
Iteration Number: 250
Loss: 137.72867758056375
l2 norm of gradients: 1.423381437307905
l2 norm of weights: 6.643698081330051
-----
Iteration Number: 251
Loss: 137.6298914458573
```

```

12 norm of gradients: 1.4227927470147903
12 norm of weights: 6.643311941561272
-----
Iteration Number: 252
Loss: 137.53118707262647
12 norm of gradients: 1.4222041093699918
12 norm of weights: 6.642926092222557
-----
Iteration Number: 253
Loss: 137.43256442030838
12 norm of gradients: 1.4216155240877537
12 norm of weights: 6.642540533111936
-----
Iteration Number: 254
Loss: 137.33402344835406
12 norm of gradients: 1.4210269908655542
12 norm of weights: 6.642155264027585
-----
Iteration Number: 255
Loss: 137.23556411614098
12 norm of gradients: 1.4204385093840264
12 norm of weights: 6.641770284767822
-----
Iteration Number: 256
Loss: 137.13718638331486
12 norm of gradients: 1.4198500793068807
12 norm of weights: 6.6413855951311085
-----
Iteration Number: 257
Loss: 137.0388902093598
12 norm of gradients: 1.4192617002808334
12 norm of weights: 6.641001194916049
-----
Iteration Number: 258
Loss: 136.94067555404922
12 norm of gradients: 1.4186733719355402
12 norm of weights: 6.640617083921388
-----
Iteration Number: 259
Loss: 136.8425423769394
12 norm of gradients: 1.4180850938835294
12 norm of weights: 6.640233261946015
-----
Iteration Number: 260
Loss: 136.74449063790433
12 norm of gradients: 1.4174968657201454
12 norm of weights: 6.6398497287889615
-----

```

Iteration Number: 261
Loss: 136.64652029664344
l2 norm of gradients: 1.4169086870234904
l2 norm of weights: 6.639466484249396

Iteration Number: 262
Loss: 136.54863131314812
l2 norm of gradients: 1.4163205573543751
l2 norm of weights: 6.6390835281266325

Iteration Number: 263
Loss: 136.4508236470777
l2 norm of gradients: 1.4157324762562717
l2 norm of weights: 6.638700860220124

Iteration Number: 264
Loss: 136.3530972586806
l2 norm of gradients: 1.4151444432552704
l2 norm of weights: 6.638318480329463

Iteration Number: 265
Loss: 136.25545210791768
l2 norm of gradients: 1.4145564578600407
l2 norm of weights: 6.637936388254384

Iteration Number: 266
Loss: 136.15788815490257
l2 norm of gradients: 1.4139685195617986
l2 norm of weights: 6.637554583794758

Iteration Number: 267
Loss: 136.06040535995757
l2 norm of gradients: 1.4133806278342758
l2 norm of weights: 6.637173066750598

Iteration Number: 268
Loss: 135.9630036831137
l2 norm of gradients: 1.4127927821336956
l2 norm of weights: 6.636791836922051

Iteration Number: 269
Loss: 135.86568308485977
l2 norm of gradients: 1.4122049818987514
l2 norm of weights: 6.6364108941094075

Iteration Number: 270
Loss: 135.76844352559792
l2 norm of gradients: 1.4116172265505902

```

l2 norm of weights: 6.636030238113092
-----
Iteration Number: 271
Loss: 135.67128496584587
l2 norm of gradients: 1.4110295154928008
l2 norm of weights: 6.635649868733665
-----
Iteration Number: 272
Loss: 135.57420736609802
l2 norm of gradients: 1.4104418481114063
l2 norm of weights: 6.635269785771827
-----
Iteration Number: 273
Loss: 135.47721068712298
l2 norm of gradients: 1.4098542237748604
l2 norm of weights: 6.6348899890284105
-----
Iteration Number: 274
Loss: 135.38029488970187
l2 norm of gradients: 1.4092666418340503
l2 norm of weights: 6.634510478304384
-----
Iteration Number: 275
Loss: 135.2834599347104
l2 norm of gradients: 1.4086791016223004
l2 norm of weights: 6.634131253400851
-----
Iteration Number: 276
Loss: 135.1867057829449
l2 norm of gradients: 1.4080916024553842
l2 norm of weights: 6.63375231411905
-----
Iteration Number: 277
Loss: 135.0900323955765
l2 norm of gradients: 1.4075041436315383
l2 norm of weights: 6.63337366026035
-----
Iteration Number: 278
Loss: 134.99343973363153
l2 norm of gradients: 1.406916724431481
l2 norm of weights: 6.632995291626256
-----
Iteration Number: 279
Loss: 134.89692775844472
l2 norm of gradients: 1.4063293441184372
l2 norm of weights: 6.6326172080184
-----
Iteration Number: 280

```

Loss: 134.80049643112986
l2 norm of gradients: 1.4057420019381652
l2 norm of weights: 6.632239409238549

Iteration Number: 281
Loss: 134.70414571351384
l2 norm of gradients: 1.4051546971189894
l2 norm of weights: 6.6318618950885995

Iteration Number: 282
Loss: 134.6078755667221
l2 norm of gradients: 1.4045674288718377
l2 norm of weights: 6.631484665370576

Iteration Number: 283
Loss: 134.5116859527165
l2 norm of gradients: 1.4039801963902812
l2 norm of weights: 6.631107719886633

Iteration Number: 284
Loss: 134.4155768330812
l2 norm of gradients: 1.403392998850581
l2 norm of weights: 6.6307310584390535

Iteration Number: 285
Loss: 134.31954816974653
l2 norm of gradients: 1.4028058354117376
l2 norm of weights: 6.630354680830244

Iteration Number: 286
Loss: 134.2235999245308
l2 norm of gradients: 1.4022187052155448
l2 norm of weights: 6.629978586862744

Iteration Number: 287
Loss: 134.12773205988398
l2 norm of gradients: 1.4016316073866482
l2 norm of weights: 6.629602776339211

Iteration Number: 288
Loss: 134.03194453777962
l2 norm of gradients: 1.4010445410326071
l2 norm of weights: 6.629227249062432

Iteration Number: 289
Loss: 133.936237320611
l2 norm of gradients: 1.4004575052439632
l2 norm of weights: 6.628852004835316

```
-----  
Iteration Number: 290  
Loss: 133.84061037086002  
l2 norm of gradients: 1.3998704990943098  
l2 norm of weights: 6.628477043460895  
-----  
Iteration Number: 291  
Loss: 133.74506365098202  
l2 norm of gradients: 1.3992835216403667  
l2 norm of weights: 6.62810236474232  
-----  
Iteration Number: 292  
Loss: 133.64959712383114  
l2 norm of gradients: 1.3986965719220619  
l2 norm of weights: 6.627727968482869  
-----  
Iteration Number: 293  
Loss: 133.55421075227528  
l2 norm of gradients: 1.3981096489626108  
l2 norm of weights: 6.627353854485933  
-----  
Iteration Number: 294  
Loss: 133.45890449915188  
l2 norm of gradients: 1.397522751768608  
l2 norm of weights: 6.626980022555026  
-----  
Iteration Number: 295  
Loss: 133.36367832768528  
l2 norm of gradients: 1.396935879330115  
l2 norm of weights: 6.626606472493779  
-----  
Iteration Number: 296  
Loss: 133.26853220116087  
l2 norm of gradients: 1.3963490306207575  
l2 norm of weights: 6.6262332041059375  
-----  
Iteration Number: 297  
Loss: 133.17346608280513  
l2 norm of gradients: 1.3957622045978224  
l2 norm of weights: 6.625860217195368  
-----  
Iteration Number: 298  
Loss: 133.07847993616474  
l2 norm of gradients: 1.3951754002023637  
l2 norm of weights: 6.625487511566045  
-----  
Iteration Number: 299  
Loss: 132.98357372502173
```

l2 norm of gradients: 1.3945886163593062
l2 norm of weights: 6.625115087022062

Iteration Number: 300
Loss: 132.88874741306657
l2 norm of gradients: 1.3940018519775568
l2 norm of weights: 6.624742943367622

Iteration Number: 301
Loss: 132.79400096425542
l2 norm of gradients: 1.3934151059501194
l2 norm of weights: 6.624371080407038

Iteration Number: 302
Loss: 132.69933434280557
l2 norm of gradients: 1.3928283771542114
l2 norm of weights: 6.623999497944739

Iteration Number: 303
Loss: 132.60474751271687
l2 norm of gradients: 1.3922416644513855
l2 norm of weights: 6.623628195785257

Iteration Number: 304
Loss: 132.51024043864663
l2 norm of gradients: 1.391654966687654
l2 norm of weights: 6.6232571737332355

Iteration Number: 305
Loss: 132.41581308503686
l2 norm of gradients: 1.3910682826936178
l2 norm of weights: 6.622886431593423

Iteration Number: 306
Loss: 132.32146541656917
l2 norm of gradients: 1.390481611284596
l2 norm of weights: 6.6225159691706725

Iteration Number: 307
Loss: 132.22719739797398
l2 norm of gradients: 1.389894951260762
l2 norm of weights: 6.622145786269945

Iteration Number: 308
Loss: 132.1330089945625
l2 norm of gradients: 1.3893083014072825
l2 norm of weights: 6.6217758826963005

Iteration Number: 309
Loss: 132.0389001712315
l2 norm of gradients: 1.388721660494456
l2 norm of weights: 6.621406258254903

Iteration Number: 310
Loss: 131.94487089345122
l2 norm of gradients: 1.3881350272778596
l2 norm of weights: 6.621036912751015

Iteration Number: 311
Loss: 131.85092112675
l2 norm of gradients: 1.387548400498495
l2 norm of weights: 6.6206678459900035

Iteration Number: 312
Loss: 131.75705083677383
l2 norm of gradients: 1.3869617788829394
l2 norm of weights: 6.620299057777326

Iteration Number: 313
Loss: 131.6632599892206
l2 norm of gradients: 1.386375161143498
l2 norm of weights: 6.619930547918542

Iteration Number: 314
Loss: 131.56954855013763
l2 norm of gradients: 1.385788545978361
l2 norm of weights: 6.619562316219303

Iteration Number: 315
Loss: 131.47591648579598
l2 norm of gradients: 1.3852019320717621
l2 norm of weights: 6.619194362485358

Iteration Number: 316
Loss: 131.3823637625181
l2 norm of gradients: 1.3846153180941394
l2 norm of weights: 6.618826686522546

Iteration Number: 317
Loss: 131.28889034675436
l2 norm of gradients: 1.3840287027023
l2 norm of weights: 6.618459288136798

Iteration Number: 318
Loss: 131.19549620513413
l2 norm of gradients: 1.3834420845395874


```

12 norm of weights: 6.618092167134133
-----
Iteration Number: 319
Loss: 131.10218130468473
12 norm of gradients: 1.3828554622360496
12 norm of weights: 6.617725323320662
-----
Iteration Number: 320
Loss: 131.00894561218755
12 norm of gradients: 1.3822688344086114
12 norm of weights: 6.617358756502582
-----
Iteration Number: 321
Loss: 130.915789095117
12 norm of gradients: 1.381682199661247
12 norm of weights: 6.616992466486172
-----
Iteration Number: 322
Loss: 130.82271172076443
12 norm of gradients: 1.3810955565851581
12 norm of weights: 6.616626453077801
-----
Iteration Number: 323
Loss: 130.72971345669063
12 norm of gradients: 1.3805089037589513
12 norm of weights: 6.616260716083917
-----
Iteration Number: 324
Loss: 130.63679427067987
12 norm of gradients: 1.3799222397488182
12 norm of weights: 6.61589525531105
-----
Iteration Number: 325
Loss: 130.54395413064856
12 norm of gradients: 1.3793355631087179
12 norm of weights: 6.61553007056581
-----
Iteration Number: 326
Loss: 130.45119300478981
12 norm of gradients: 1.3787488723805625
12 norm of weights: 6.615165161654885
-----
Iteration Number: 327
Loss: 130.35851086125376
12 norm of gradients: 1.3781621660944026
12 norm of weights: 6.614800528385042
-----
Iteration Number: 328

```

```
Loss: 130.26590766882333
l2 norm of gradients: 1.377575442768615
l2 norm of weights: 6.6144361705631205
-----
Iteration Number: 329
Loss: 130.1733833959368
l2 norm of gradients: 1.3769887009100936
l2 norm of weights: 6.6140720879960355
-----
Iteration Number: 330
Loss: 130.08093801161976
l2 norm of gradients: 1.3764019390144404
l2 norm of weights: 6.6137082804907745
-----
Iteration Number: 331
Loss: 129.98857148482097
l2 norm of gradients: 1.3758151555661586
l2 norm of weights: 6.613344747854394
-----
Iteration Number: 332
Loss: 129.89628378494214
l2 norm of gradients: 1.3752283490388475
l2 norm of weights: 6.612981489894022
-----
Iteration Number: 333
Loss: 129.8040748813243
l2 norm of gradients: 1.374641517895398
l2 norm of weights: 6.612618506416854
-----
Iteration Number: 334
Loss: 129.71194474374542
l2 norm of gradients: 1.374054660588191
l2 norm of weights: 6.6122557972301506
-----
Iteration Number: 335
Loss: 129.6198933419641
l2 norm of gradients: 1.3734677755592952
l2 norm of weights: 6.611893362141237
-----
Iteration Number: 336
Loss: 129.52792064601744
l2 norm of gradients: 1.3728808612406678
l2 norm of weights: 6.611531200957503
-----
Iteration Number: 337
Loss: 129.4360266262761
l2 norm of gradients: 1.3722939160543555
l2 norm of weights: 6.611169313486399
```

```

-----
Iteration Number: 338
Loss: 129.34421125309575
l2 norm of gradients: 1.3717069384126956
l2 norm of weights: 6.6108076995354335
-----
Iteration Number: 339
Loss: 129.2524744971162
l2 norm of gradients: 1.371119926718521
l2 norm of weights: 6.610446358912177
-----
Iteration Number: 340
Loss: 129.1608163291742
l2 norm of gradients: 1.370532879365363
l2 norm of weights: 6.610085291424255
-----
Iteration Number: 341
Loss: 129.06923672032815
l2 norm of gradients: 1.3699457947376572
l2 norm of weights: 6.609724496879346
-----
Iteration Number: 342
Loss: 128.97773564185465
l2 norm of gradients: 1.3693586712109485
l2 norm of weights: 6.609363975085184
-----
Iteration Number: 343
Loss: 128.8863130650649
l2 norm of gradients: 1.368771507152098
l2 norm of weights: 6.609003725849557
-----
Iteration Number: 344
Loss: 128.7949689618897
l2 norm of gradients: 1.3681843009194907
l2 norm of weights: 6.608643748980297
-----
Iteration Number: 345
Loss: 128.703703304072
l2 norm of gradients: 1.367597050863242
l2 norm of weights: 6.608284044285291
-----
Iteration Number: 346
Loss: 128.61251606353557
l2 norm of gradients: 1.3670097553254086
l2 norm of weights: 6.60792461157247
-----
Iteration Number: 347
Loss: 128.52140721258812

```

l2 norm of gradients: 1.3664224126401938
l2 norm of weights: 6.607565450649811

Iteration Number: 348
Loss: 128.43037672395437
l2 norm of gradients: 1.3658350211341603
l2 norm of weights: 6.607206561325333

Iteration Number: 349
Loss: 128.3394245700893
l2 norm of gradients: 1.3652475791264371
l2 norm of weights: 6.606847943407098

Iteration Number: 350
Loss: 128.2485507238765
l2 norm of gradients: 1.3646600849289303
l2 norm of weights: 6.606489596703211

Iteration Number: 351
Loss: 128.15775515858024
l2 norm of gradients: 1.3640725368465345
l2 norm of weights: 6.606131521021811

Iteration Number: 352
Loss: 128.067037847361
l2 norm of gradients: 1.3634849331773407
l2 norm of weights: 6.605773716171076

Iteration Number: 353
Loss: 127.97639876377339
l2 norm of gradients: 1.3628972722128483
l2 norm of weights: 6.605416181959221

Iteration Number: 354
Loss: 127.8858378816714
l2 norm of gradients: 1.3623095522381747
l2 norm of weights: 6.6050589181944925

Iteration Number: 355
Loss: 127.79535517475313
l2 norm of gradients: 1.361721771532268
l2 norm of weights: 6.604701924685169

Iteration Number: 356
Loss: 127.70495061742962
l2 norm of gradients: 1.3611339283681128
l2 norm of weights: 6.6043452012395605

Iteration Number: 357
Loss: 127.61462418396215
l2 norm of gradients: 1.360546021012945
l2 norm of weights: 6.603988747666004

Iteration Number: 358
Loss: 127.52437584893111
l2 norm of gradients: 1.3599580477284599
l2 norm of weights: 6.603632563772865

Iteration Number: 359
Loss: 127.4342055870017
l2 norm of gradients: 1.359370006771022
l2 norm of weights: 6.603276649368534

Iteration Number: 360
Loss: 127.34411337344358
l2 norm of gradients: 1.358781896391875
l2 norm of weights: 6.6029210042614235

Iteration Number: 361
Loss: 127.25409918320503
l2 norm of gradients: 1.3581937148373513
l2 norm of weights: 6.602565628259968

Iteration Number: 362
Loss: 127.16416299186832
l2 norm of gradients: 1.3576054603490815
l2 norm of weights: 6.6022105211726245

Iteration Number: 363
Loss: 127.07430477508171
l2 norm of gradients: 1.3570171311642003
l2 norm of weights: 6.601855682807866

Iteration Number: 364
Loss: 126.98452450857347
l2 norm of gradients: 1.3564287255155587
l2 norm of weights: 6.601501112974183

Iteration Number: 365
Loss: 126.89482216854206
l2 norm of gradients: 1.3558402416319288
l2 norm of weights: 6.60114681148008

Iteration Number: 366
Loss: 126.80519773115913
l2 norm of gradients: 1.3552516777382115

```

12 norm of weights: 6.600792778134078
-----
Iteration Number: 367
Loss: 126.71565117291671
12 norm of gradients: 1.3546630320556439
12 norm of weights: 6.600439012744706
-----
Iteration Number: 368
Loss: 126.62618247066905
12 norm of gradients: 1.3540743028020044
12 norm of weights: 6.6000855151205045
-----
Iteration Number: 369
Loss: 126.5367916012025
12 norm of gradients: 1.3534854881918181
12 norm of weights: 6.599732285070021
-----
Iteration Number: 370
Loss: 126.44747854181159
12 norm of gradients: 1.3528965864365623
12 norm of weights: 6.599379322401812
-----
Iteration Number: 371
Loss: 126.35824326959242
12 norm of gradients: 1.352307595744869
12 norm of weights: 6.599026626924436
-----
Iteration Number: 372
Loss: 126.26908576236515
12 norm of gradients: 1.35171851432273
12 norm of weights: 6.5986741984464565
-----
Iteration Number: 373
Loss: 126.1800059976557
12 norm of gradients: 1.3511293403736977
12 norm of weights: 6.598322036776437
-----
Iteration Number: 374
Loss: 126.09100395374684
12 norm of gradients: 1.3505400720990874
12 norm of weights: 6.597970141722941
-----
Iteration Number: 375
Loss: 126.00207960856442
12 norm of gradients: 1.3499507076981787
12 norm of weights: 6.597618513094532
-----
Iteration Number: 376

```

Loss: 125.91323294068079
l2 norm of gradients: 1.349361245368414
l2 norm of weights: 6.597267150699767

Iteration Number: 377
Loss: 125.82446392872734
l2 norm of gradients: 1.3487716833055994
l2 norm of weights: 6.596916054347198

Iteration Number: 378
Loss: 125.73577255144814
l2 norm of gradients: 1.348182019704101
l2 norm of weights: 6.5965652238453725

Iteration Number: 379
Loss: 125.64715878802465
l2 norm of gradients: 1.347592252757044
l2 norm of weights: 6.596214659002825

Iteration Number: 380
Loss: 125.55862261772738
l2 norm of gradients: 1.3470023806565081
l2 norm of weights: 6.595864359628084

Iteration Number: 381
Loss: 125.47016401994641
l2 norm of gradients: 1.3464124015937227
l2 norm of weights: 6.5955143255296615

Iteration Number: 382
Loss: 125.38178297436681
l2 norm of gradients: 1.3458223137592615
l2 norm of weights: 6.595164556516059

Iteration Number: 383
Loss: 125.29347946102297
l2 norm of gradients: 1.345232115343237
l2 norm of weights: 6.594815052395761

Iteration Number: 384
Loss: 125.20525345994066
l2 norm of gradients: 1.3446418045354906
l2 norm of weights: 6.594465812977234

Iteration Number: 385
Loss: 125.1171049514457
l2 norm of gradients: 1.3440513795257858
l2 norm of weights: 6.594116838068929

```

-----
Iteration Number: 386
Loss: 125.02903391606908
l2 norm of gradients: 1.3434608385039977
l2 norm of weights: 6.59376812747927
-----
Iteration Number: 387
Loss: 124.94104033472581
l2 norm of gradients: 1.3428701796603035
l2 norm of weights: 6.593419681016665
-----
Iteration Number: 388
Loss: 124.8531241881239
l2 norm of gradients: 1.3422794011853687
l2 norm of weights: 6.593071498489495
-----
Iteration Number: 389
Loss: 124.76528545756122
l2 norm of gradients: 1.341688501270537
l2 norm of weights: 6.592723579706117
-----
Iteration Number: 390
Loss: 124.67752412443528
l2 norm of gradients: 1.3410974781080125
l2 norm of weights: 6.592375924474858
-----
Iteration Number: 391
Loss: 124.58984017033141
l2 norm of gradients: 1.3405063298910496
l2 norm of weights: 6.5920285326040196
-----
Iteration Number: 392
Loss: 124.50223357693888
l2 norm of gradients: 1.3399150548141323
l2 norm of weights: 6.59168140390187
-----
Iteration Number: 393
Loss: 124.41470432637489
l2 norm of gradients: 1.3393236510731612
l2 norm of weights: 6.591334538176646
-----
Iteration Number: 394
Loss: 124.32725240076145
l2 norm of gradients: 1.3387321168656316
l2 norm of weights: 6.590987935236551
-----
Iteration Number: 395
Loss: 124.23987778263044

```


l2 norm of gradients: 1.3381404503908176
l2 norm of weights: 6.590641594889753

Iteration Number: 396
Loss: 124.15258045459666
l2 norm of gradients: 1.337548649849949
l2 norm of weights: 6.590295516944383

Iteration Number: 397
Loss: 124.06536039933563
l2 norm of gradients: 1.3369567134463913
l2 norm of weights: 6.58994970120853

Iteration Number: 398
Loss: 123.97821759999503
l2 norm of gradients: 1.3363646393858235
l2 norm of weights: 6.589604147490247

Iteration Number: 399
Loss: 123.89115203975805
l2 norm of gradients: 1.3357724258764136
l2 norm of weights: 6.589258855597544

Iteration Number: 400
Loss: 123.80416370217903
l2 norm of gradients: 1.3351800711289943
l2 norm of weights: 6.588913825338383

Iteration Number: 401
Loss: 123.7172525706785
l2 norm of gradients: 1.3345875733572383
l2 norm of weights: 6.588569056520687

Iteration Number: 402
Loss: 123.63041862922039
l2 norm of gradients: 1.3339949307778292
l2 norm of weights: 6.588224548952328

Iteration Number: 403
Loss: 123.5436618618362
l2 norm of gradients: 1.333402141610637
l2 norm of weights: 6.587880302441131

Iteration Number: 404
Loss: 123.45698225279791
l2 norm of gradients: 1.3328092040788866
l2 norm of weights: 6.58753631679487

Iteration Number: 405
Loss: 123.37037978653078
l2 norm of gradients: 1.3322161164093298
l2 norm of weights: 6.5871925918212675

Iteration Number: 406
Loss: 123.28385444757768
l2 norm of gradients: 1.3316228768324134
l2 norm of weights: 6.586849127327993

Iteration Number: 407
Loss: 123.19740622086935
l2 norm of gradients: 1.3310294835824479
l2 norm of weights: 6.586505923122662

Iteration Number: 408
Loss: 123.11103509135906
l2 norm of gradients: 1.3304359348977757
l2 norm of weights: 6.586162979012831

Iteration Number: 409
Loss: 123.0247410443746
l2 norm of gradients: 1.329842229020935
l2 norm of weights: 6.585820294805999

Iteration Number: 410
Loss: 122.93852406522649
l2 norm of gradients: 1.3292483641988286
l2 norm of weights: 6.585477870309608

Iteration Number: 411
Loss: 122.8523841395443
l2 norm of gradients: 1.328654338682885
l2 norm of weights: 6.585135705331037

Iteration Number: 412
Loss: 122.76632125323998
l2 norm of gradients: 1.3280601507292247
l2 norm of weights: 6.584793799677603

Iteration Number: 413
Loss: 122.68033539209894
l2 norm of gradients: 1.3274657985988205
l2 norm of weights: 6.584452153156555

Iteration Number: 414
Loss: 122.5944265425463
l2 norm of gradients: 1.3268712805576621

```

12 norm of weights: 6.58411076557508
-----
Iteration Number: 415
Loss: 122.50859469084128
12 norm of gradients: 1.326276594876914
12 norm of weights: 6.583769636740298
-----
Iteration Number: 416
Loss: 122.42283982342906
12 norm of gradients: 1.325681739833079
12 norm of weights: 6.583428766459257
-----
Iteration Number: 417
Loss: 122.33716192733124
12 norm of gradients: 1.325086713708154
12 norm of weights: 6.583088154538937
-----
Iteration Number: 418
Loss: 122.25156098922783
12 norm of gradients: 1.324491514789792
12 norm of weights: 6.582747800786245
-----
Iteration Number: 419
Loss: 122.16603699638635
12 norm of gradients: 1.323896141371458
12 norm of weights: 6.582407705008014
-----
Iteration Number: 420
Loss: 122.08058993598965
12 norm of gradients: 1.3233005917525862
12 norm of weights: 6.582067867011003
-----
Iteration Number: 421
Loss: 121.99521979565668
12 norm of gradients: 1.3227048642387371
12 norm of weights: 6.581728286601894
-----
Iteration Number: 422
Loss: 121.90992656295067
12 norm of gradients: 1.322108957141753
12 norm of weights: 6.581388963587291
-----
Iteration Number: 423
Loss: 121.82471022584559
12 norm of gradients: 1.3215128687799136
12 norm of weights: 6.581049897773718
-----
Iteration Number: 424

```

```

Loss: 121.73957077226261
l2 norm of gradients: 1.3209165974780885
l2 norm of weights: 6.580711088967617
-----
Iteration Number: 425
Loss: 121.65450819041021
l2 norm of gradients: 1.3203201415678942
l2 norm of weights: 6.580372536975351
-----
Iteration Number: 426
Loss: 121.56952246867442
l2 norm of gradients: 1.3197234993878448
l2 norm of weights: 6.580034241603196
-----
Iteration Number: 427
Loss: 121.4846135956468
l2 norm of gradients: 1.3191266692835062
l2 norm of weights: 6.579696202657342
-----
Iteration Number: 428
Loss: 121.399781560049
l2 norm of gradients: 1.3185296496076475
l2 norm of weights: 6.579358419943894
-----
Iteration Number: 429
Loss: 121.31502635080452
l2 norm of gradients: 1.3179324387203935
l2 norm of weights: 6.579020893268869
-----
Iteration Number: 430
Loss: 121.23034795698332
l2 norm of gradients: 1.317335034989375
l2 norm of weights: 6.578683622438192
-----
Iteration Number: 431
Loss: 121.14574636774124
l2 norm of gradients: 1.3167374367898816
l2 norm of weights: 6.578346607257698
-----
Iteration Number: 432
Loss: 121.06122157262713
l2 norm of gradients: 1.3161396425050096
l2 norm of weights: 6.578009847533128
-----
Iteration Number: 433
Loss: 120.97677356121348
l2 norm of gradients: 1.3155416505258137
l2 norm of weights: 6.577673343070131

```

```

-----
Iteration Number: 434
Loss: 120.89240232315659
l2 norm of gradients: 1.3149434592514568
l2 norm of weights: 6.57733709367426
-----
Iteration Number: 435
Loss: 120.8081078483861
l2 norm of gradients: 1.314345067089359
l2 norm of weights: 6.577001099150969
-----
Iteration Number: 436
Loss: 120.7238901270842
l2 norm of gradients: 1.3137464724553456
l2 norm of weights: 6.576665359305614
-----
Iteration Number: 437
Loss: 120.63974914929065
l2 norm of gradients: 1.313147673773799
l2 norm of weights: 6.576329873943454
-----
Iteration Number: 438
Loss: 120.55568490563057
l2 norm of gradients: 1.3125486694778037
l2 norm of weights: 6.575994642869643
-----
Iteration Number: 439
Loss: 120.47169738657341
l2 norm of gradients: 1.311949458009298
l2 norm of weights: 6.575659665889235
-----
Iteration Number: 440
Loss: 120.38778658282875
l2 norm of gradients: 1.3113500378192198
l2 norm of weights: 6.57532494280718
-----
Iteration Number: 441
Loss: 120.30395248529457
l2 norm of gradients: 1.3107504073676555
l2 norm of weights: 6.57499047342832
-----
Iteration Number: 442
Loss: 120.22019508488529
l2 norm of gradients: 1.3101505651239889
l2 norm of weights: 6.574656257557392
-----
Iteration Number: 443
Loss: 120.13651437301571

```

l2 norm of gradients: 1.3095505095670485
l2 norm of weights: 6.574322294999028

Iteration Number: 444
Loss: 120.05291034069592
l2 norm of gradients: 1.3089502391852545
l2 norm of weights: 6.573988585557744

Iteration Number: 445
Loss: 119.96938297975048
l2 norm of gradients: 1.3083497524767682
l2 norm of weights: 6.573655129037949

Iteration Number: 446
Loss: 119.88593228148028
l2 norm of gradients: 1.3077490479496379
l2 norm of weights: 6.57332192524394

Iteration Number: 447
Loss: 119.80255823802948
l2 norm of gradients: 1.3071481241219491
l2 norm of weights: 6.572988973979901

Iteration Number: 448
Loss: 119.71926084097636
l2 norm of gradients: 1.3065469795219706
l2 norm of weights: 6.572656275049898

Iteration Number: 449
Loss: 119.63604008261629
l2 norm of gradients: 1.3059456126883033
l2 norm of weights: 6.5723238282578835

Iteration Number: 450
Loss: 119.55289595494864
l2 norm of gradients: 1.3053440221700274
l2 norm of weights: 6.571991633407692

Iteration Number: 451
Loss: 119.4698284505677
l2 norm of gradients: 1.3047422065268517
l2 norm of weights: 6.57165969030304

Iteration Number: 452
Loss: 119.38683756173465
l2 norm of gradients: 1.3041401643292616
l2 norm of weights: 6.571327998747521

Iteration Number: 453
Loss: 119.30392328131175
l2 norm of gradients: 1.3035378941586673
l2 norm of weights: 6.5709965585446115

Iteration Number: 454
Loss: 119.22108560182859
l2 norm of gradients: 1.3029353946075526
l2 norm of weights: 6.57066536949766

Iteration Number: 455
Loss: 119.1383245163314
l2 norm of gradients: 1.302332664279625
l2 norm of weights: 6.570334431409898

Iteration Number: 456
Loss: 119.05564001772281
l2 norm of gradients: 1.3017297017899623
l2 norm of weights: 6.570003744084424

Iteration Number: 457
Loss: 118.97303209928371
l2 norm of gradients: 1.301126505765166
l2 norm of weights: 6.569673307324215

Iteration Number: 458
Loss: 118.89050075419553
l2 norm of gradients: 1.300523074843507
l2 norm of weights: 6.5693431209321185

Iteration Number: 459
Loss: 118.80804597598691
l2 norm of gradients: 1.2999194076750789
l2 norm of weights: 6.569013184710854

Iteration Number: 460
Loss: 118.72566775798533
l2 norm of gradients: 1.2993155029219465
l2 norm of weights: 6.56868349846301

Iteration Number: 461
Loss: 118.64336609407093
l2 norm of gradients: 1.2987113592582975
l2 norm of weights: 6.568354061991044

Iteration Number: 462
Loss: 118.56114097793437
l2 norm of gradients: 1.2981069753705938

```

l2 norm of weights: 6.568024875097279
-----
Iteration Number: 463
Loss: 118.47899240347216
l2 norm of gradients: 1.2975023499577232
l2 norm of weights: 6.567695937583907
-----
Iteration Number: 464
Loss: 118.39692036472587
l2 norm of gradients: 1.2968974817311498
l2 norm of weights: 6.5673672492529835
-----
Iteration Number: 465
Loss: 118.31492485571972
l2 norm of gradients: 1.2962923694150692
l2 norm of weights: 6.5670388099064265
-----
Iteration Number: 466
Loss: 118.23300587080381
l2 norm of gradients: 1.2956870117465589
l2 norm of weights: 6.5667106193460185
-----
Iteration Number: 467
Loss: 118.15116340436671
l2 norm of gradients: 1.2950814074757333
l2 norm of weights: 6.5663826773734
-----
Iteration Number: 468
Loss: 118.06939745075725
l2 norm of gradients: 1.2944755553658969
l2 norm of weights: 6.5660549837900755
-----
Iteration Number: 469
Loss: 117.987708004684
l2 norm of gradients: 1.2938694541936988
l2 norm of weights: 6.565727538397405
-----
Iteration Number: 470
Loss: 117.90609506071182
l2 norm of gradients: 1.293263102749287
l2 norm of weights: 6.56540034099661
-----
Iteration Number: 471
Loss: 117.8245586136544
l2 norm of gradients: 1.2926564998364645
l2 norm of weights: 6.565073391388764
-----
Iteration Number: 472

```



```

Loss: 117.74309865840208
l2 norm of gradients: 1.2920496442728457
l2 norm of weights: 6.5647466893748
-----
Iteration Number: 473
Loss: 117.6617151899411
l2 norm of gradients: 1.291442534890011
l2 norm of weights: 6.564420234755502
-----
Iteration Number: 474
Loss: 117.58040820349002
l2 norm of gradients: 1.2908351705336656
l2 norm of weights: 6.56409402733151
-----
Iteration Number: 475
Loss: 117.49917769400768
l2 norm of gradients: 1.2902275500637963
l2 norm of weights: 6.563768066903313
-----
Iteration Number: 476
Loss: 117.41802365689705
l2 norm of gradients: 1.2896196723548297
l2 norm of weights: 6.563442353271253
-----
Iteration Number: 477
Loss: 117.33694608742509
l2 norm of gradients: 1.2890115362957906
l2 norm of weights: 6.563116886235522
-----
Iteration Number: 478
Loss: 117.2559449811625
l2 norm of gradients: 1.2884031407904617
l2 norm of weights: 6.5627916655961585
-----
Iteration Number: 479
Loss: 117.17502033366534
l2 norm of gradients: 1.287794484757544
l2 norm of weights: 6.562466691153051
-----
Iteration Number: 480
Loss: 117.09417214047429
l2 norm of gradients: 1.2871855671308159
l2 norm of weights: 6.562141962705932
-----
Iteration Number: 481
Loss: 117.01340039729979
l2 norm of gradients: 1.2865763868592945
l2 norm of weights: 6.5618174800543825

```

```
-----
Iteration Number: 482
Loss: 116.93270510007345
l2 norm of gradients: 1.2859669429073983
l2 norm of weights: 6.561493242997824
-----
Iteration Number: 483
Loss: 116.852086244423
l2 norm of gradients: 1.2853572342551092
l2 norm of weights: 6.561169251335525
-----
Iteration Number: 484
Loss: 116.77154382658189
l2 norm of gradients: 1.2847472598981333
l2 norm of weights: 6.560845504866594
-----
Iteration Number: 485
Loss: 116.69107784232541
l2 norm of gradients: 1.284137018848067
l2 norm of weights: 6.5605220033899805
-----
Iteration Number: 486
Loss: 116.61068828800947
l2 norm of gradients: 1.2835265101325593
l2 norm of weights: 6.560198746704477
-----
Iteration Number: 487
Loss: 116.53037515956541
l2 norm of gradients: 1.2829157327954765
l2 norm of weights: 6.55987573460871
-----
Iteration Number: 488
Loss: 116.45013845327075
l2 norm of gradients: 1.2823046858970677
l2 norm of weights: 6.559552966901149
-----
Iteration Number: 489
Loss: 116.36997816570339
l2 norm of gradients: 1.281693368514131
l2 norm of weights: 6.559230443380099
-----
Iteration Number: 490
Loss: 116.28989429279228
l2 norm of gradients: 1.2810817797401775
l2 norm of weights: 6.558908163843699
-----
Iteration Number: 491
Loss: 116.20988683138046
```

```

12 norm of gradients: 1.2804699186856012
12 norm of weights: 6.558586128089928
-----
Iteration Number: 492
Loss: 116.12995577761605
12 norm of gradients: 1.2798577844778434
12 norm of weights: 6.558264335916592
-----
Iteration Number: 493
Loss: 116.05010112817617
12 norm of gradients: 1.279245376261562
12 norm of weights: 6.557942787121338
-----
Iteration Number: 494
Loss: 115.97032287977407
12 norm of gradients: 1.2786326931988001
12 norm of weights: 6.557621481501638
-----
Iteration Number: 495
Loss: 115.8906210288589
12 norm of gradients: 1.278019734469154
12 norm of weights: 6.557300418854802
-----
Iteration Number: 496
Loss: 115.81099557235964
12 norm of gradients: 1.2774064992699423
12 norm of weights: 6.556979598977967
-----
Iteration Number: 497
Loss: 115.73144650682389
12 norm of gradients: 1.276792986816377
12 norm of weights: 6.556659021668096
-----
Iteration Number: 498
Loss: 115.65197382915225
12 norm of gradients: 1.2761791963417324
12 norm of weights: 6.556338686721989
-----
Iteration Number: 499
Loss: 115.57257753608222
12 norm of gradients: 1.2755651270975163
12 norm of weights: 6.556018593936265
-----
Iteration Number: 500
Loss: 115.49325762463089
12 norm of gradients: 1.2749507783536411
12 norm of weights: 6.555698743107374
-----

```

Iteration Number: 501
Loss: 115.41401409170957
l2 norm of gradients: 1.274336149398595
l2 norm of weights: 6.555379134031594

Iteration Number: 502
Loss: 115.33484693405424
l2 norm of gradients: 1.2737212395396136
l2 norm of weights: 6.555059766505022

Iteration Number: 503
Loss: 115.25575614889762
l2 norm of gradients: 1.2731060481028529
l2 norm of weights: 6.554740640323584

Iteration Number: 504
Loss: 115.1767417331283
l2 norm of gradients: 1.2724905744335604
l2 norm of weights: 6.554421755283028

Iteration Number: 505
Loss: 115.09780368381367
l2 norm of gradients: 1.2718748178962502
l2 norm of weights: 6.5541031111789225

Iteration Number: 506
Loss: 115.01894199803742
l2 norm of gradients: 1.271258777874873
l2 norm of weights: 6.553784707806661

Iteration Number: 507
Loss: 114.94015667289958
l2 norm of gradients: 1.2706424537729915
l2 norm of weights: 6.553466544961454

Iteration Number: 508
Loss: 114.86144770538405
l2 norm of gradients: 1.2700258450139528
l2 norm of weights: 6.553148622438335

Iteration Number: 509
Loss: 114.78281509278136
l2 norm of gradients: 1.2694089510410629
l2 norm of weights: 6.552830940032156

Iteration Number: 510
Loss: 114.70425883221341
l2 norm of gradients: 1.26879177131776

```

12 norm of weights: 6.552513497537588
-----
Iteration Number: 511
Loss: 114.62577892063668
12 norm of gradients: 1.268174305327788
12 norm of weights: 6.552196294749117
-----
Iteration Number: 512
Loss: 114.54737535557291
12 norm of gradients: 1.2675565525753716
12 norm of weights: 6.551879331461049
-----
Iteration Number: 513
Loss: 114.46904813385538
12 norm of gradients: 1.2669385125853891
12 norm of weights: 6.551562607467504
-----
Iteration Number: 514
Loss: 114.39079725290635
12 norm of gradients: 1.266320184903547
12 norm of weights: 6.55124612256242
-----
Iteration Number: 515
Loss: 114.31262270968588
12 norm of gradients: 1.2657015690965543
12 norm of weights: 6.550929876539546
-----
Iteration Number: 516
Loss: 114.23452450161028
12 norm of gradients: 1.265082664752295
12 norm of weights: 6.55061386919245
-----
Iteration Number: 517
Loss: 114.15650262565214
12 norm of gradients: 1.2644634714800047
12 norm of weights: 6.550298100314508
-----
Iteration Number: 518
Loss: 114.07855707909036
12 norm of gradients: 1.263843988910441
12 norm of weights: 6.549982569698914
-----
Iteration Number: 519
Loss: 114.00068785912075
12 norm of gradients: 1.2632242166960597
12 norm of weights: 6.549667277138668
-----
Iteration Number: 520

```

```

Loss: 113.92289496287555
l2 norm of gradients: 1.2626041545111861
l2 norm of weights: 6.549352222426589
-----
Iteration Number: 521
Loss: 113.84517838730237
l2 norm of gradients: 1.2619838020521892
l2 norm of weights: 6.5490374053553
-----
Iteration Number: 522
Loss: 113.76753812974891
l2 norm of gradients: 1.261363159037654
l2 norm of weights: 6.548722825717238
-----
Iteration Number: 523
Loss: 113.68997418731955
l2 norm of gradients: 1.2607422252085532
l2 norm of weights: 6.548408483304647
-----
Iteration Number: 524
Loss: 113.61248655689157
l2 norm of gradients: 1.2601210003284202
l2 norm of weights: 6.548094377909583
-----
Iteration Number: 525
Loss: 113.53507523562783
l2 norm of gradients: 1.2594994841835196
l2 norm of weights: 6.547780509323908
-----
Iteration Number: 526
Loss: 113.45774022068817
l2 norm of gradients: 1.2588776765830183
l2 norm of weights: 6.547466877339294
-----
Iteration Number: 527
Loss: 113.38048150883542
l2 norm of gradients: 1.2582555773591562
l2 norm of weights: 6.547153481747219
-----
Iteration Number: 528
Loss: 113.3032990971678
l2 norm of gradients: 1.2576331863674157
l2 norm of weights: 6.546840322338966
-----
Iteration Number: 529
Loss: 113.22619298262678
l2 norm of gradients: 1.2570105034866907
l2 norm of weights: 6.546527398905626

```

```

-----
Iteration Number: 530
Loss: 113.14916316197981
l2 norm of gradients: 1.2563875286194548
l2 norm of weights: 6.546214711238101
-----
Iteration Number: 531
Loss: 113.0722096320828
l2 norm of gradients: 1.2557642616919296
l2 norm of weights: 6.5459022591270895
-----
Iteration Number: 532
Loss: 112.99533238985494
l2 norm of gradients: 1.2551407026542505
l2 norm of weights: 6.5455900423631
-----
Iteration Number: 533
Loss: 112.91853143206261
l2 norm of gradients: 1.2545168514806337
l2 norm of weights: 6.545278060736445
-----
Iteration Number: 534
Loss: 112.84180675519659
l2 norm of gradients: 1.2538927081695401
l2 norm of weights: 6.544966314037239
-----
Iteration Number: 535
Loss: 112.76515835600597
l2 norm of gradients: 1.2532682727438398
l2 norm of weights: 6.544654802055404
-----
Iteration Number: 536
Loss: 112.6885862310945
l2 norm of gradients: 1.252643545250975
l2 norm of weights: 6.544343524580661
-----
Iteration Number: 537
Loss: 112.61209037712435
l2 norm of gradients: 1.2520185257631218
l2 norm of weights: 6.544032481402537
-----
Iteration Number: 538
Loss: 112.53567079021532
l2 norm of gradients: 1.2513932143773514
l2 norm of weights: 6.543721672310361
-----
Iteration Number: 539
Loss: 112.45932746713177

```

l2 norm of gradients: 1.2507676112157882
l2 norm of weights: 6.543411097093262

Iteration Number: 540
Loss: 112.38306040398247
l2 norm of gradients: 1.2501417164257695
l2 norm of weights: 6.543100755540174

Iteration Number: 541
Loss: 112.30686959703279
l2 norm of gradients: 1.249515530180001
l2 norm of weights: 6.542790647439831

Iteration Number: 542
Loss: 112.23075504258841
l2 norm of gradients: 1.2488890526767131
l2 norm of weights: 6.542480772580768

Iteration Number: 543
Loss: 112.15471673666103
l2 norm of gradients: 1.248262284139815
l2 norm of weights: 6.542171130751322

Iteration Number: 544
Loss: 112.07875467529226
l2 norm of gradients: 1.2476352248190459
l2 norm of weights: 6.541861721739632

Iteration Number: 545
Loss: 112.00286885452972
l2 norm of gradients: 1.2470078749901272
l2 norm of weights: 6.5415525453336345

Iteration Number: 546
Loss: 111.92705927001902
l2 norm of gradients: 1.2463802349549107
l2 norm of weights: 6.541243601321068

Iteration Number: 547
Loss: 111.85132591765569
l2 norm of gradients: 1.2457523050415262
l2 norm of weights: 6.540934889489473

Iteration Number: 548
Loss: 111.77566879300035
l2 norm of gradients: 1.2451240856045285
l2 norm of weights: 6.540626409626189

Iteration Number: 549
Loss: 111.70008789187686
l2 norm of gradients: 1.2444955770250383
l2 norm of weights: 6.540318161518352

Iteration Number: 550
Loss: 111.62458320952027
l2 norm of gradients: 1.243866779710887
l2 norm of weights: 6.540010144952903

Iteration Number: 551
Loss: 111.54915474134347
l2 norm of gradients: 1.243237694096755
l2 norm of weights: 6.539702359716579

Iteration Number: 552
Loss: 111.4738024827204
l2 norm of gradients: 1.2426083206443104
l2 norm of weights: 6.539394805595921

Iteration Number: 553
Loss: 111.39852642861946
l2 norm of gradients: 1.2419786598423428
l2 norm of weights: 6.539087482377266

Iteration Number: 554
Loss: 111.32332657427837
l2 norm of gradients: 1.2413487122068998
l2 norm of weights: 6.5387803898467505

Iteration Number: 555
Loss: 111.2482029145835
l2 norm of gradients: 1.240718478281416
l2 norm of weights: 6.538473527790311

Iteration Number: 556
Loss: 111.17315544432046
l2 norm of gradients: 1.2400879586368425
l2 norm of weights: 6.538166895993687

Iteration Number: 557
Loss: 111.09818415820662
l2 norm of gradients: 1.2394571538717736
l2 norm of weights: 6.537860494242413

Iteration Number: 558
Loss: 111.02328905083522
l2 norm of gradients: 1.2388260646125706

```

12 norm of weights: 6.5375543223218235
-----
Iteration Number: 559
Loss: 110.94847011652155
12 norm of gradients: 1.2381946915134838
12 norm of weights: 6.537248380017058
-----
Iteration Number: 560
Loss: 110.8737273497148
12 norm of gradients: 1.2375630352567708
12 norm of weights: 6.53694266711305
-----
Iteration Number: 561
Loss: 110.79906074449485
12 norm of gradients: 1.2369310965528137
12 norm of weights: 6.536637183394535
-----
Iteration Number: 562
Loss: 110.7244702949102
12 norm of gradients: 1.2362988761402312
12 norm of weights: 6.53633192864605
-----
Iteration Number: 563
Loss: 110.6499559949771
12 norm of gradients: 1.2356663747859897
12 norm of weights: 6.536026902651931
-----
Iteration Number: 564
Loss: 110.5755178383676
12 norm of gradients: 1.2350335932855125
12 norm of weights: 6.535722105196313
-----
Iteration Number: 565
Loss: 110.50115581856736
12 norm of gradients: 1.234400532462783
12 norm of weights: 6.535417536063134
-----
Iteration Number: 566
Loss: 110.42686992929173
12 norm of gradients: 1.2337671931704464
12 norm of weights: 6.535113195036135
-----
Iteration Number: 567
Loss: 110.35266016374801
12 norm of gradients: 1.2331335762899107
12 norm of weights: 6.534809081898852
-----
Iteration Number: 568

```

```
Loss: 110.27852651503271
l2 norm of gradients: 1.2324996827314385
l2 norm of weights: 6.534505196434627
-----
Iteration Number: 569
Loss: 110.20446897614819
l2 norm of gradients: 1.2318655134342429
l2 norm of weights: 6.534201538426604
-----
Iteration Number: 570
Loss: 110.13048753989322
l2 norm of gradients: 1.231231069366575
l2 norm of weights: 6.533898107657728
-----
Iteration Number: 571
Loss: 110.05658219904853
l2 norm of gradients: 1.2305963515258076
l2 norm of weights: 6.533594903910746
-----
Iteration Number: 572
Loss: 109.98275294600828
l2 norm of gradients: 1.229961360938522
l2 norm of weights: 6.533291926968209
-----
Iteration Number: 573
Loss: 109.90899977321835
l2 norm of gradients: 1.2293260986605816
l2 norm of weights: 6.532989176612471
-----
Iteration Number: 574
Loss: 109.83532267261259
l2 norm of gradients: 1.2286905657772105
l2 norm of weights: 6.5326866526256895
-----
Iteration Number: 575
Loss: 109.76172163639934
l2 norm of gradients: 1.2280547634030632
l2 norm of weights: 6.5323843547898255
-----
Iteration Number: 576
Loss: 109.6881966561542
l2 norm of gradients: 1.2274186926822923
l2 norm of weights: 6.532082282886648
-----
Iteration Number: 577
Loss: 109.6147477237067
l2 norm of gradients: 1.2267823547886134
l2 norm of weights: 6.531780436697728
```

```
-----
Iteration Number: 578
Loss: 109.54137483027166
l2 norm of gradients: 1.226145750925364
l2 norm of weights: 6.5314788160044435
-----
Iteration Number: 579
Loss: 109.46807796707017
l2 norm of gradients: 1.2255088823255602
l2 norm of weights: 6.531177420587977
-----
Iteration Number: 580
Loss: 109.39485712537869
l2 norm of gradients: 1.224871750251948
l2 norm of weights: 6.530876250229323
-----
Iteration Number: 581
Loss: 109.3217122956988
l2 norm of gradients: 1.2242343559970528
l2 norm of weights: 6.530575304709279
-----
Iteration Number: 582
Loss: 109.2486434689007
l2 norm of gradients: 1.2235967008832214
l2 norm of weights: 6.530274583808452
-----
Iteration Number: 583
Loss: 109.17565063530483
l2 norm of gradients: 1.2229587862626627
l2 norm of weights: 6.529974087307256
-----
Iteration Number: 584
Loss: 109.10273378524764
l2 norm of gradients: 1.2223206135174836
l2 norm of weights: 6.52967381498592
-----
Iteration Number: 585
Loss: 109.02989290853913
l2 norm of gradients: 1.221682184059718
l2 norm of weights: 6.529373766624478
-----
Iteration Number: 586
Loss: 108.95712799519373
l2 norm of gradients: 1.2210434993313564
l2 norm of weights: 6.52907394200278
-----
Iteration Number: 587
Loss: 108.88443903458561
```

```

12 norm of gradients: 1.220404560804365
12 norm of weights: 6.528774340900482
-----
Iteration Number: 588
Loss: 108.81182601620313
12 norm of gradients: 1.2197653699807054
12 norm of weights: 6.528474963097057
-----
Iteration Number: 589
Loss: 108.73928892907418
12 norm of gradients: 1.2191259283923472
12 norm of weights: 6.528175808371793
-----
Iteration Number: 590
Loss: 108.66682776222216
12 norm of gradients: 1.2184862376012764
12 norm of weights: 6.527876876503787
-----
Iteration Number: 591
Loss: 108.5944425043575
12 norm of gradients: 1.2178462991994983
12 norm of weights: 6.5275781672719555
-----
Iteration Number: 592
Loss: 108.52213314381017
12 norm of gradients: 1.2172061148090374
12 norm of weights: 6.52727968045503
-----
Iteration Number: 593
Loss: 108.44989966881117
12 norm of gradients: 1.216565686081932
12 norm of weights: 6.526981415831561
-----
Iteration Number: 594
Loss: 108.37774206735276
12 norm of gradients: 1.2159250147002214
12 norm of weights: 6.526683373179914
-----
Iteration Number: 595
Loss: 108.30566032715502
12 norm of gradients: 1.2152841023759322
12 norm of weights: 6.526385552278276
-----
Iteration Number: 596
Loss: 108.23365443583906
12 norm of gradients: 1.2146429508510557
12 norm of weights: 6.526087952904654
-----

```

Iteration Number: 597
Loss: 108.16172438044616
l2 norm of gradients: 1.2140015618975244
l2 norm of weights: 6.525790574836877

Iteration Number: 598
Loss: 108.08987014807103
l2 norm of gradients: 1.2133599373171786
l2 norm of weights: 6.525493417852595

Iteration Number: 599
Loss: 108.01809172534566
l2 norm of gradients: 1.2127180789417324
l2 norm of weights: 6.5251964817292825

Iteration Number: 600
Loss: 107.94638909898744
l2 norm of gradients: 1.2120759886327308
l2 norm of weights: 6.524899766244237

Iteration Number: 601
Loss: 107.87476225489542
l2 norm of gradients: 1.2114336682815037
l2 norm of weights: 6.524603271174587

Iteration Number: 602
Loss: 107.80321117925449
l2 norm of gradients: 1.2107911198091146
l2 norm of weights: 6.524306996297281

Iteration Number: 603
Loss: 107.73173585774467
l2 norm of gradients: 1.2101483451663024
l2 norm of weights: 6.524010941389101

Iteration Number: 604
Loss: 107.66033627560786
l2 norm of gradients: 1.2095053463334198
l2 norm of weights: 6.5237151062266605

Iteration Number: 605
Loss: 107.5890124181963
l2 norm of gradients: 1.208862125320363
l2 norm of weights: 6.523419490586396

Iteration Number: 606
Loss: 107.51776427019202
l2 norm of gradients: 1.2082186841665012

```

12 norm of weights: 6.523124094244583
-----
Iteration Number: 607
Loss: 107.44659181639832
12 norm of gradients: 1.2075750249405943
12 norm of weights: 6.52282891697733
-----
Iteration Number: 608
Loss: 107.37549504094635
12 norm of gradients: 1.2069311497407107
12 norm of weights: 6.522533958560577
-----
Iteration Number: 609
Loss: 107.30447392796947
12 norm of gradients: 1.2062870606941343
12 norm of weights: 6.522239218770106
-----
Iteration Number: 610
Loss: 107.233528461116
12 norm of gradients: 1.2056427599572714
12 norm of weights: 6.521944697381532
-----
Iteration Number: 611
Loss: 107.16265862397276
12 norm of gradients: 1.2049982497155463
12 norm of weights: 6.521650394170311
-----
Iteration Number: 612
Loss: 107.09186439970422
12 norm of gradients: 1.2043535321832943
12 norm of weights: 6.5213563089117415
-----
Iteration Number: 613
Loss: 107.0211457710704
12 norm of gradients: 1.2037086096036504
12 norm of weights: 6.521062441380963
-----
Iteration Number: 614
Loss: 106.95050272067544
12 norm of gradients: 1.2030634842484274
12 norm of weights: 6.520768791352959
-----
Iteration Number: 615
Loss: 106.87993523078914
12 norm of gradients: 1.2024181584179925
12 norm of weights: 6.520475358602559
-----
Iteration Number: 616

```

Loss: 106.80944328339433
l2 norm of gradients: 1.2017726344411364
l2 norm of weights: 6.5201821429044395

Iteration Number: 617
Loss: 106.7390268602648
l2 norm of gradients: 1.2011269146749362
l2 norm of weights: 6.519889144033128

Iteration Number: 618
Loss: 106.66868594255219
l2 norm of gradients: 1.200481001504613
l2 norm of weights: 6.519596361762998

Iteration Number: 619
Loss: 106.59842051141936
l2 norm of gradients: 1.1998348973433837
l2 norm of weights: 6.51930379586828

Iteration Number: 620
Loss: 106.52823054749344
l2 norm of gradients: 1.199188604632305
l2 norm of weights: 6.519011446123056

Iteration Number: 621
Loss: 106.45811603130919
l2 norm of gradients: 1.1985421258401154
l2 norm of weights: 6.518719312301266

Iteration Number: 622
Loss: 106.3880769429782
l2 norm of gradients: 1.1978954634630665
l2 norm of weights: 6.518427394176704

Iteration Number: 623
Loss: 106.3181132621659
l2 norm of gradients: 1.1972486200247519
l2 norm of weights: 6.518135691523026

Iteration Number: 624
Loss: 106.24822496833718
l2 norm of gradients: 1.1966015980759275
l2 norm of weights: 6.517844204113751

Iteration Number: 625
Loss: 106.17841204062262
l2 norm of gradients: 1.1959544001943285
l2 norm of weights: 6.517552931722258


```

-----
Iteration Number: 626
Loss: 106.10867445793647
l2 norm of gradients: 1.1953070289844774
l2 norm of weights: 6.517261874121791
-----
Iteration Number: 627
Loss: 106.03901219864068
l2 norm of gradients: 1.1946594870774883
l2 norm of weights: 6.516971031085465
-----
Iteration Number: 628
Loss: 105.96942524076813
l2 norm of gradients: 1.1940117771308632
l2 norm of weights: 6.5166804023862595
-----
Iteration Number: 629
Loss: 105.89991356234977
l2 norm of gradients: 1.193363901828285
l2 norm of weights: 6.516389987797028
-----
Iteration Number: 630
Loss: 105.83047714061502
l2 norm of gradients: 1.1927158638794009
l2 norm of weights: 6.5160997870904955
-----
Iteration Number: 631
Loss: 105.76111595282129
l2 norm of gradients: 1.1920676660196032
l2 norm of weights: 6.515809800039263
-----
Iteration Number: 632
Loss: 105.6918299756233
l2 norm of gradients: 1.1914193110098017
l2 norm of weights: 6.5155200264158095
-----
Iteration Number: 633
Loss: 105.6226191855776
l2 norm of gradients: 1.1907708016361895
l2 norm of weights: 6.51523046599249
-----
Iteration Number: 634
Loss: 105.5534835586536
l2 norm of gradients: 1.1901221407100073
l2 norm of weights: 6.514941118541545
-----
Iteration Number: 635
Loss: 105.48442307071633

```

l2 norm of gradients: 1.1894733310672951
l2 norm of weights: 6.514651983835095

Iteration Number: 636
Loss: 105.41543769693146
l2 norm of gradients: 1.188824375568644
l2 norm of weights: 6.514363061645149

Iteration Number: 637
Loss: 105.34652741254928
l2 norm of gradients: 1.1881752770989376
l2 norm of weights: 6.514074351743604

Iteration Number: 638
Loss: 105.27769219220075
l2 norm of gradients: 1.1875260385670892
l2 norm of weights: 6.513785853902245

Iteration Number: 639
Loss: 105.20893201010085
l2 norm of gradients: 1.186876662905775
l2 norm of weights: 6.5134975678927525

Iteration Number: 640
Loss: 105.14024684027085
l2 norm of gradients: 1.1862271530711566
l2 norm of weights: 6.5132094934866975

Iteration Number: 641
Loss: 105.07163665619302
l2 norm of gradients: 1.1855775120426018
l2 norm of weights: 6.512921630455553

Iteration Number: 642
Loss: 105.00310143132455
l2 norm of gradients: 1.1849277428224
l2 norm of weights: 6.5126339785706895

Iteration Number: 643
Loss: 104.93464113823421
l2 norm of gradients: 1.184277848435466
l2 norm of weights: 6.512346537603379

Iteration Number: 644
Loss: 104.86625574973569
l2 norm of gradients: 1.1836278319290447
l2 norm of weights: 6.512059307324797

Iteration Number: 645
Loss: 104.79794523764004
l2 norm of gradients: 1.1829776963724083
l2 norm of weights: 6.511772287506028

Iteration Number: 646
Loss: 104.72970957393677
l2 norm of gradients: 1.1823274448565435
l2 norm of weights: 6.511485477918063

Iteration Number: 647
Loss: 104.66154873002853
l2 norm of gradients: 1.1816770804938388
l2 norm of weights: 6.5111988783318075

Iteration Number: 648
Loss: 104.59346267670963
l2 norm of gradients: 1.181026606417763
l2 norm of weights: 6.510912488518079

Iteration Number: 649
Loss: 104.52545138474842
l2 norm of gradients: 1.1803760257825395
l2 norm of weights: 6.510626308247611

Iteration Number: 650
Loss: 104.45751482450414
l2 norm of gradients: 1.1797253417628135
l2 norm of weights: 6.510340337291058

Iteration Number: 651
Loss: 104.38965296566681
l2 norm of gradients: 1.1790745575533155
l2 norm of weights: 6.510054575418996

Iteration Number: 652
Loss: 104.32186577794955
l2 norm of gradients: 1.1784236763685174
l2 norm of weights: 6.509769022401924

Iteration Number: 653
Loss: 104.25415323017786
l2 norm of gradients: 1.1777727014422863
l2 norm of weights: 6.5094836780102705

Iteration Number: 654
Loss: 104.18651529140811
l2 norm of gradients: 1.1771216360275294

```

12 norm of weights: 6.509198542014389
-----
Iteration Number: 655
Loss: 104.11895192979222
12 norm of gradients: 1.1764704833958368
12 norm of weights: 6.508913614184571
-----
Iteration Number: 656
Loss: 104.0514631134204
12 norm of gradients: 1.1758192468371165
12 norm of weights: 6.508628894291038
-----
Iteration Number: 657
Loss: 103.9840488097736
12 norm of gradients: 1.1751679296592257
12 norm of weights: 6.508344382103952
-----
Iteration Number: 658
Loss: 103.91670898605565
12 norm of gradients: 1.174516535187599
12 norm of weights: 6.508060077393412
-----
Iteration Number: 659
Loss: 103.84944360908251
12 norm of gradients: 1.1738650667648658
12 norm of weights: 6.507775979929465
-----
Iteration Number: 660
Loss: 103.7822526453812
12 norm of gradients: 1.1732135277504707
12 norm of weights: 6.507492089482099
-----
Iteration Number: 661
Loss: 103.7151360609014
12 norm of gradients: 1.1725619215202823
12 norm of weights: 6.507208405821255
-----
Iteration Number: 662
Loss: 103.64809382115163
12 norm of gradients: 1.1719102514662008
12 norm of weights: 6.5069249287168205
-----
Iteration Number: 663
Loss: 103.58112589160895
12 norm of gradients: 1.1712585209957604
12 norm of weights: 6.506641657938639
-----
Iteration Number: 664

```

Loss: 103.51423223691862
l2 norm of gradients: 1.1706067335317274
l2 norm of weights: 6.506358593256514

Iteration Number: 665
Loss: 103.44741282174779
l2 norm of gradients: 1.169954892511692
l2 norm of weights: 6.506075734440204

Iteration Number: 666
Loss: 103.38066760984191
l2 norm of gradients: 1.169303001387659
l2 norm of weights: 6.5057930812594345

Iteration Number: 667
Loss: 103.31399656520863
l2 norm of gradients: 1.1686510636256306
l2 norm of weights: 6.505510633483893

Iteration Number: 668
Loss: 103.24739965078948
l2 norm of gradients: 1.1679990827051887
l2 norm of weights: 6.5052283908832385

Iteration Number: 669
Loss: 103.18087682976348
l2 norm of gradients: 1.1673470621190696
l2 norm of weights: 6.504946353227099

Iteration Number: 670
Loss: 103.11442806432608
l2 norm of gradients: 1.1666950053727374
l2 norm of weights: 6.504664520285077

Iteration Number: 671
Loss: 103.04805331674558
l2 norm of gradients: 1.166042915983952
l2 norm of weights: 6.504382891826755

Iteration Number: 672
Loss: 102.98175254860448
l2 norm of gradients: 1.165390797482335
l2 norm of weights: 6.504101467621692

Iteration Number: 673
Loss: 102.91552572117918
l2 norm of gradients: 1.164738653408929
l2 norm of weights: 6.503820247439433

```
-----  
Iteration Number: 674  
Loss: 102.84937279535566  
l2 norm of gradients: 1.1640864873157573  
l2 norm of weights: 6.503539231049508  
-----  
Iteration Number: 675  
Loss: 102.78329373161343  
l2 norm of gradients: 1.1634343027653762  
l2 norm of weights: 6.503258418221435  
-----  
Iteration Number: 676  
Loss: 102.71728849003748  
l2 norm of gradients: 1.1627821033304273  
l2 norm of weights: 6.502977808724727  
-----  
Iteration Number: 677  
Loss: 102.65135703020118  
l2 norm of gradients: 1.1621298925931836  
l2 norm of weights: 6.50269740232889  
-----  
Iteration Number: 678  
Loss: 102.58549931159214  
l2 norm of gradients: 1.1614776741450963  
l2 norm of weights: 6.502417198803428  
-----  
Iteration Number: 679  
Loss: 102.51971529288537  
l2 norm of gradients: 1.1608254515863339  
l2 norm of weights: 6.5021371979178495  
-----  
Iteration Number: 680  
Loss: 102.45400493256246  
l2 norm of gradients: 1.1601732285253235  
l2 norm of weights: 6.501857399441663  
-----  
Iteration Number: 681  
Loss: 102.38836818878696  
l2 norm of gradients: 1.1595210085782843  
l2 norm of weights: 6.501577803144386  
-----  
Iteration Number: 682  
Loss: 102.32280501915504  
l2 norm of gradients: 1.1588687953687635  
l2 norm of weights: 6.501298408795549  
-----  
Iteration Number: 683  
Loss: 102.25731538102647
```

l2 norm of gradients: 1.1582165925271652
l2 norm of weights: 6.501019216164693

Iteration Number: 684
Loss: 102.19189923119069
l2 norm of gradients: 1.1575644036902808
l2 norm of weights: 6.500740225021377

Iteration Number: 685
Loss: 102.12655652610569
l2 norm of gradients: 1.1569122325008139
l2 norm of weights: 6.500461435135178

Iteration Number: 686
Loss: 102.06128722185574
l2 norm of gradients: 1.1562600826069058
l2 norm of weights: 6.5001828462757

Iteration Number: 687
Loss: 101.99609127412249
l2 norm of gradients: 1.155607957661658
l2 norm of weights: 6.499904458212568

Iteration Number: 688
Loss: 101.93096863816803
l2 norm of gradients: 1.1549558613226518
l2 norm of weights: 6.4996262707154395

Iteration Number: 689
Loss: 101.86591926887363
l2 norm of gradients: 1.1543037972514676
l2 norm of weights: 6.499348283554005

Iteration Number: 690
Loss: 101.8009431206235
l2 norm of gradients: 1.1536517691132027
l2 norm of weights: 6.4990704964979855

Iteration Number: 691
Loss: 101.73604014769603
l2 norm of gradients: 1.1529997805759875
l2 norm of weights: 6.498792909317145

Iteration Number: 692
Loss: 101.67121030356587
l2 norm of gradients: 1.152347835310499
l2 norm of weights: 6.498515521781288

Iteration Number: 693
Loss: 101.60645354152607
l2 norm of gradients: 1.151695936989476
l2 norm of weights: 6.4982383336602645

Iteration Number: 694
Loss: 101.54176981456665
l2 norm of gradients: 1.15104408928723
l2 norm of weights: 6.4979613447239695

Iteration Number: 695
Loss: 101.47715907518962
l2 norm of gradients: 1.1503922958791584
l2 norm of weights: 6.497684554742354

Iteration Number: 696
Loss: 101.41262127528789
l2 norm of gradients: 1.1497405604412547
l2 norm of weights: 6.497407963485419

Iteration Number: 697
Loss: 101.34815636686483
l2 norm of gradients: 1.1490888866496194
l2 norm of weights: 6.497131570723225

Iteration Number: 698
Loss: 101.28376430094494
l2 norm of gradients: 1.1484372781799694
l2 norm of weights: 6.496855376225892

Iteration Number: 699
Loss: 101.21944502854603
l2 norm of gradients: 1.1477857387071486
l2 norm of weights: 6.496579379763606

Iteration Number: 700
Loss: 101.15519850036237
l2 norm of gradients: 1.147134271904638
l2 norm of weights: 6.496303581106616

Iteration Number: 701
Loss: 101.0910246662202
l2 norm of gradients: 1.1464828814440629
l2 norm of weights: 6.496027980025245

Iteration Number: 702
Loss: 101.02692347603258
l2 norm of gradients: 1.1458315709947058


```

12 norm of weights: 6.495752576289888
-----
Iteration Number: 703
Loss: 100.96289487926597
12 norm of gradients: 1.145180344223014
12 norm of weights: 6.495477369671015
-----
Iteration Number: 704
Loss: 100.89893882470099
12 norm of gradients: 1.1445292047921123
12 norm of weights: 6.495202359939177
-----
Iteration Number: 705
Loss: 100.83505526094503
12 norm of gradients: 1.1438781563613118
12 norm of weights: 6.494927546865007
-----
Iteration Number: 706
Loss: 100.77124413639949
12 norm of gradients: 1.1432272025856247
12 norm of weights: 6.494652930219225
-----
Iteration Number: 707
Loss: 100.70750539870434
12 norm of gradients: 1.1425763471152734
12 norm of weights: 6.494378509772636
-----
Iteration Number: 708
Loss: 100.64383899529845
12 norm of gradients: 1.141925593595208
12 norm of weights: 6.494104285296143
-----
Iteration Number: 709
Loss: 100.58024487336968
12 norm of gradients: 1.1412749456646185
12 norm of weights: 6.4938302565607415
-----
Iteration Number: 710
Loss: 100.51672297950378
12 norm of gradients: 1.140624406956452
12 norm of weights: 6.493556423337524
-----
Iteration Number: 711
Loss: 100.45327326013192
12 norm of gradients: 1.1399739810969307
12 norm of weights: 6.493282785397685
-----
Iteration Number: 712

```

```
Loss: 100.38989566098147
l2 norm of gradients: 1.1393236717050708
l2 norm of weights: 6.4930093425125275
-----
Iteration Number: 713
Loss: 100.32659012783593
l2 norm of gradients: 1.138673482392204
l2 norm of weights: 6.492736094453455
-----
Iteration Number: 714
Loss: 100.26335660580293
l2 norm of gradients: 1.1380234167614989
l2 norm of weights: 6.4924630409919875
-----
Iteration Number: 715
Loss: 100.20019503970634
l2 norm of gradients: 1.1373734784074883
l2 norm of weights: 6.492190181899757
-----
Iteration Number: 716
Loss: 100.13710537401144
l2 norm of gradients: 1.136723670915594
l2 norm of weights: 6.491917516948512
-----
Iteration Number: 717
Loss: 100.07408755282357
l2 norm of gradients: 1.1360739978616574
l2 norm of weights: 6.491645045910123
-----
Iteration Number: 718
Loss: 100.01114151987905
l2 norm of gradients: 1.1354244628114718
l2 norm of weights: 6.491372768556581
-----
Iteration Number: 719
Loss: 99.94826721855709
l2 norm of gradients: 1.1347750693203147
l2 norm of weights: 6.491100684660003
-----
Iteration Number: 720
Loss: 99.88546459177023
l2 norm of gradients: 1.134125820932487
l2 norm of weights: 6.490828793992638
-----
Iteration Number: 721
Loss: 99.8227335822771
l2 norm of gradients: 1.133476721180852
l2 norm of weights: 6.4905570963268655
```

```

-----
Iteration Number: 722
Loss: 99.76007413228464
l2 norm of gradients: 1.1328277735863783
l2 norm of weights: 6.4902855914352005
-----
Iteration Number: 723
Loss: 99.69748618384617
l2 norm of gradients: 1.132178981657686
l2 norm of weights: 6.490014279090294
-----
Iteration Number: 724
Loss: 99.63496967846685
l2 norm of gradients: 1.131530348890595
l2 norm of weights: 6.489743159064941
-----
Iteration Number: 725
Loss: 99.57252455749284
l2 norm of gradients: 1.1308818787676798
l2 norm of weights: 6.48947223113208
-----
Iteration Number: 726
Loss: 99.51015076172745
l2 norm of gradients: 1.1302335747578227
l2 norm of weights: 6.489201495064796
-----
Iteration Number: 727
Loss: 99.44784823184797
l2 norm of gradients: 1.1295854403157761
l2 norm of weights: 6.488930950636323
-----
Iteration Number: 728
Loss: 99.38561690786892
l2 norm of gradients: 1.128937478881724
l2 norm of weights: 6.4886605976200515
-----
Iteration Number: 729
Loss: 99.32345672987323
l2 norm of gradients: 1.1282896938808498
l2 norm of weights: 6.488390435789524
-----
Iteration Number: 730
Loss: 99.26136763730159
l2 norm of gradients: 1.127642088722909
l2 norm of weights: 6.488120464918444
-----
Iteration Number: 731
Loss: 99.19934956936946

```

```
12 norm of gradients: 1.126994666801802
12 norm of weights: 6.487850684780674
-----
Iteration Number: 732
Loss: 99.13740246486583
12 norm of gradients: 1.1263474314951565
12 norm of weights: 6.487581095150245
-----
Iteration Number: 733
Loss: 99.07552626253516
12 norm of gradients: 1.1257003861639083
12 norm of weights: 6.487311695801352
-----
Iteration Number: 734
Loss: 99.01372090051197
12 norm of gradients: 1.1250535341518928
12 norm of weights: 6.487042486508362
-----
Iteration Number: 735
Loss: 98.9519863165951
12 norm of gradients: 1.1244068787854353
12 norm of weights: 6.486773467045814
-----
Iteration Number: 736
Loss: 98.8903224484569
12 norm of gradients: 1.123760423372951
12 norm of weights: 6.486504637188422
-----
Iteration Number: 737
Loss: 98.82872923327511
12 norm of gradients: 1.1231141712045443
12 norm of weights: 6.4862359967110805
-----
Iteration Number: 738
Loss: 98.76720660819123
12 norm of gradients: 1.1224681255516193
12 norm of weights: 6.485967545388863
-----
Iteration Number: 739
Loss: 98.70575450955215
12 norm of gradients: 1.1218222896664893
12 norm of weights: 6.48569928299703
-----
Iteration Number: 740
Loss: 98.64437287396424
12 norm of gradients: 1.1211766667819958
12 norm of weights: 6.485431209311025
-----
```

Iteration Number: 741
Loss: 98.58306163716098
l2 norm of gradients: 1.1205312601111295
l2 norm of weights: 6.485163324106482

Iteration Number: 742
Loss: 98.52182073506106
l2 norm of gradients: 1.1198860728466593
l2 norm of weights: 6.484895627159229

Iteration Number: 743
Loss: 98.46065010307849
l2 norm of gradients: 1.1192411081607643
l2 norm of weights: 6.484628118245286

Iteration Number: 744
Loss: 98.39954967612593
l2 norm of gradients: 1.1185963692046725
l2 norm of weights: 6.484360797140873

Iteration Number: 745
Loss: 98.3385193892296
l2 norm of gradients: 1.1179518591083055
l2 norm of weights: 6.4840936636224065

Iteration Number: 746
Loss: 98.27755917682136
l2 norm of gradients: 1.1173075809799262
l2 norm of weights: 6.483826717466506

Iteration Number: 747
Loss: 98.21666897313993
l2 norm of gradients: 1.1166635379057965
l2 norm of weights: 6.483559958449999

Iteration Number: 748
Loss: 98.15584871214662
l2 norm of gradients: 1.1160197329498365
l2 norm of weights: 6.483293386349917

Iteration Number: 749
Loss: 98.09509832751186
l2 norm of gradients: 1.1153761691532922
l2 norm of weights: 6.483027000943504

Iteration Number: 750
Loss: 98.03441775263825
l2 norm of gradients: 1.1147328495344078

```

12 norm of weights: 6.4827608020082135
-----
Iteration Number: 751
Loss: 97.97380692063147
12 norm of gradients: 1.1140897770881053
12 norm of weights: 6.482494789321715
-----
Iteration Number: 752
Loss: 97.91326576424959
12 norm of gradients: 1.113446954785668
12 norm of weights: 6.482228962661898
-----
Iteration Number: 753
Loss: 97.85279421626309
12 norm of gradients: 1.1128043855744332
12 norm of weights: 6.481963321806867
-----
Iteration Number: 754
Loss: 97.79239220868867
12 norm of gradients: 1.1121620723774872
12 norm of weights: 6.481697866534952
-----
Iteration Number: 755
Loss: 97.73205967386214
12 norm of gradients: 1.1115200180933709
12 norm of weights: 6.481432596624706
-----
Iteration Number: 756
Loss: 97.67179654343063
12 norm of gradients: 1.1108782255957879
12 norm of weights: 6.481167511854907
-----
Iteration Number: 757
Loss: 97.6116027489204
12 norm of gradients: 1.110236697733321
12 norm of weights: 6.4809026120045665
-----
Iteration Number: 758
Loss: 97.55147822155897
12 norm of gradients: 1.1095954373291548
12 norm of weights: 6.480637896852922
-----
Iteration Number: 759
Loss: 97.49142289247398
12 norm of gradients: 1.1089544471808046
12 norm of weights: 6.480373366179448
-----
Iteration Number: 760

```

```

Loss: 97.43143669234631
l2 norm of gradients: 1.1083137300598511
l2 norm of weights: 6.480109019763852
-----
Iteration Number: 761
Loss: 97.37151955171673
l2 norm of gradients: 1.1076732887116834
l2 norm of weights: 6.47984485738608
-----
Iteration Number: 762
Loss: 97.31167140099231
l2 norm of gradients: 1.1070331258552466
l2 norm of weights: 6.479580878826321
-----
Iteration Number: 763
Loss: 97.25189217011433
l2 norm of gradients: 1.106393244182797
l2 norm of weights: 6.479317083865001
-----
Iteration Number: 764
Loss: 97.19218178893162
l2 norm of gradients: 1.105753646359665
l2 norm of weights: 6.479053472282794
-----
Iteration Number: 765
Loss: 97.13254018704507
l2 norm of gradients: 1.1051143350240213
l2 norm of weights: 6.478790043860618
-----
Iteration Number: 766
Loss: 97.07296729381844
l2 norm of gradients: 1.1044753127866542
l2 norm of weights: 6.47852679837964
-----
Iteration Number: 767
Loss: 97.01346303841603
l2 norm of gradients: 1.1038365822307505
l2 norm of weights: 6.478263735621276
-----
Iteration Number: 768
Loss: 96.9540273498708
l2 norm of gradients: 1.1031981459116842
l2 norm of weights: 6.4780008553671955
-----
Iteration Number: 769
Loss: 96.894660156705
l2 norm of gradients: 1.102560006356812
l2 norm of weights: 6.4777381573993225

```

```

-----
Iteration Number: 770
Loss: 96.83536138751629
l2 norm of gradients: 1.1019221660652754
l2 norm of weights: 6.477475641499835
-----
Iteration Number: 771
Loss: 96.77613097051996
l2 norm of gradients: 1.1012846275078099
l2 norm of weights: 6.47721330745117
-----
Iteration Number: 772
Loss: 96.71696883400672
l2 norm of gradients: 1.100647393126561
l2 norm of weights: 6.476951155036025
-----
Iteration Number: 773
Loss: 96.65787490566501
l2 norm of gradients: 1.1000104653349068
l2 norm of weights: 6.476689184037356
-----
Iteration Number: 774
Loss: 96.59884911330366
l2 norm of gradients: 1.0993738465172866
l2 norm of weights: 6.476427394238388
-----
Iteration Number: 775
Loss: 96.53989138442708
l2 norm of gradients: 1.0987375390290388
l2 norm of weights: 6.476165785422603
-----
Iteration Number: 776
Loss: 96.48100164625245
l2 norm of gradients: 1.098101545196243
l2 norm of weights: 6.475904357373757
-----
Iteration Number: 777
Loss: 96.42217982600322
l2 norm of gradients: 1.0974658673155706
l2 norm of weights: 6.475643109875871
-----
Iteration Number: 778
Loss: 96.36342585051734
l2 norm of gradients: 1.096830507654142
l2 norm of weights: 6.4753820427132345
-----
Iteration Number: 779
Loss: 96.30473964658283

```


l2 norm of gradients: 1.0961954684493884
l2 norm of weights: 6.475121155670414

Iteration Number: 780
Loss: 96.24612114087554
l2 norm of gradients: 1.0955607519089252
l2 norm of weights: 6.474860448532243

Iteration Number: 781
Loss: 96.18757025981856
l2 norm of gradients: 1.0949263602104278
l2 norm of weights: 6.474599921083834

Iteration Number: 782
Loss: 96.12908692946989
l2 norm of gradients: 1.094292295501515
l2 norm of weights: 6.474339573110575

Iteration Number: 783
Loss: 96.07067107611338
l2 norm of gradients: 1.093658559899642
l2 norm of weights: 6.47407940439813

Iteration Number: 784
Loss: 96.01232262556303
l2 norm of gradients: 1.0930251554919967
l2 norm of weights: 6.4738194147324455

Iteration Number: 785
Loss: 95.95404150361094
l2 norm of gradients: 1.0923920843354047
l2 norm of weights: 6.473559603899747

Iteration Number: 786
Loss: 95.89582763595828
l2 norm of gradients: 1.0917593484562393
l2 norm of weights: 6.473299971686541

Iteration Number: 787
Loss: 95.83768094787204
l2 norm of gradients: 1.0911269498503418
l2 norm of weights: 6.473040517879619

Iteration Number: 788
Loss: 95.77960136477819
l2 norm of gradients: 1.090494890482944
l2 norm of weights: 6.472781242266058

Iteration Number: 789
Loss: 95.72158881194206
l2 norm of gradients: 1.0898631722885994
l2 norm of weights: 6.472522144633219

Iteration Number: 790
Loss: 95.66364321421074
l2 norm of gradients: 1.0892317971711234
l2 norm of weights: 6.472263224768753

Iteration Number: 791
Loss: 95.60576449658045
l2 norm of gradients: 1.0886007670035345
l2 norm of weights: 6.472004482460597

Iteration Number: 792
Loss: 95.54795258392573
l2 norm of gradients: 1.0879700836280073
l2 norm of weights: 6.471745917496979

Iteration Number: 793
Loss: 95.49020740058691
l2 norm of gradients: 1.08733974885583
l2 norm of weights: 6.471487529666419

Iteration Number: 794
Loss: 95.43252887138155
l2 norm of gradients: 1.0867097644673664
l2 norm of weights: 6.471229318757727

Iteration Number: 795
Loss: 95.37491692051809
l2 norm of gradients: 1.0860801322120284
l2 norm of weights: 6.47097128456001

Iteration Number: 796
Loss: 95.31737147244951
l2 norm of gradients: 1.085450853808252
l2 norm of weights: 6.470713426862664

Iteration Number: 797
Loss: 95.25989245106564
l2 norm of gradients: 1.0848219309434792
l2 norm of weights: 6.470455745455385

Iteration Number: 798
Loss: 95.20247978061005
l2 norm of gradients: 1.0841933652741496

```

12 norm of weights: 6.470198240128163
-----
Iteration Number: 799
Loss: 95.1451333850382
12 norm of gradients: 1.0835651584256936
12 norm of weights: 6.469940910671287
-----
Iteration Number: 800
Loss: 95.08785318818839
12 norm of gradients: 1.082937311992536
12 norm of weights: 6.469683756875342
-----
Iteration Number: 801
Loss: 95.0306391136206
12 norm of gradients: 1.0823098275381027
12 norm of weights: 6.469426778531215
-----
Iteration Number: 802
Loss: 94.97349108525862
12 norm of gradients: 1.0816827065948353
12 norm of weights: 6.469169975430091
-----
Iteration Number: 803
Loss: 94.91640902639394
12 norm of gradients: 1.0810559506642108
12 norm of weights: 6.468913347363456
-----
Iteration Number: 804
Loss: 94.85939286067449
12 norm of gradients: 1.0804295612167678
12 norm of weights: 6.468656894123101
-----
Iteration Number: 805
Loss: 94.80244251139374
12 norm of gradients: 1.0798035396921384
12 norm of weights: 6.468400615501115
-----
Iteration Number: 806
Loss: 94.74555790194621
12 norm of gradients: 1.0791778874990852
12 norm of weights: 6.468144511289893
-----
Iteration Number: 807
Loss: 94.68873895550766
12 norm of gradients: 1.0785526060155461
12 norm of weights: 6.467888581282134
-----
Iteration Number: 808

```

```
Loss: 94.63198559512946
l2 norm of gradients: 1.0779276965886833
l2 norm of weights: 6.46763282527084
-----
Iteration Number: 809
Loss: 94.57529774394287
l2 norm of gradients: 1.0773031605349366
l2 norm of weights: 6.467377243049322
-----
Iteration Number: 810
Loss: 94.51867532495211
l2 norm of gradients: 1.0766789991400862
l2 norm of weights: 6.467121834411193
-----
Iteration Number: 811
Loss: 94.46211826110475
l2 norm of gradients: 1.0760552136593173
l2 norm of weights: 6.466866599150376
-----
Iteration Number: 812
Loss: 94.40562647538329
l2 norm of gradients: 1.0754318053172924
l2 norm of weights: 6.466611537061098
-----
Iteration Number: 813
Loss: 94.34919989041383
l2 norm of gradients: 1.0748087753082278
l2 norm of weights: 6.466356647937897
-----
Iteration Number: 814
Loss: 94.29283842906545
l2 norm of gradients: 1.074186124795976
l2 norm of weights: 6.466101931575617
-----
Iteration Number: 815
Loss: 94.23654201405978
l2 norm of gradients: 1.0735638549141133
l2 norm of weights: 6.465847387769411
-----
Iteration Number: 816
Loss: 94.18031056806765
l2 norm of gradients: 1.0729419667660323
l2 norm of weights: 6.465593016314743
-----
Iteration Number: 817
Loss: 94.12414401373275
l2 norm of gradients: 1.0723204614250406
l2 norm of weights: 6.465338817007385
```

```

-----
Iteration Number: 818
Loss: 94.068042273619
l2 norm of gradients: 1.0716993399344625
l2 norm of weights: 6.465084789643419
-----
Iteration Number: 819
Loss: 94.01200527019465
l2 norm of gradients: 1.0710786033077473
l2 norm of weights: 6.464830934019238
-----
Iteration Number: 820
Loss: 93.95603292616197
l2 norm of gradients: 1.0704582525285835
l2 norm of weights: 6.464577249931547
-----
Iteration Number: 821
Loss: 93.9001251637137
l2 norm of gradients: 1.0698382885510134
l2 norm of weights: 6.464323737177358
-----
Iteration Number: 822
Loss: 93.84428190559773
l2 norm of gradients: 1.0692187122995582
l2 norm of weights: 6.464070395553999
-----
Iteration Number: 823
Loss: 93.78850307405314
l2 norm of gradients: 1.0685995246693447
l2 norm of weights: 6.463817224859106
-----
Iteration Number: 824
Loss: 93.7327885915764
l2 norm of gradients: 1.0679807265262347
l2 norm of weights: 6.46356422489063
-----
Iteration Number: 825
Loss: 93.67713838047548
l2 norm of gradients: 1.0673623187069645
l2 norm of weights: 6.463311395446831
-----
Iteration Number: 826
Loss: 93.62155236321043
l2 norm of gradients: 1.0667443020192826
l2 norm of weights: 6.463058736326282
-----
Iteration Number: 827
Loss: 93.56603046204646

```

l2 norm of gradients: 1.0661266772420968
l2 norm of weights: 6.46280624732787

Iteration Number: 828
Loss: 93.51057259944601
l2 norm of gradients: 1.0655094451256222
l2 norm of weights: 6.46255392825079

Iteration Number: 829
Loss: 93.45517869763972
l2 norm of gradients: 1.064892606391535
l2 norm of weights: 6.462301778894556

Iteration Number: 830
Loss: 93.39984867908805
l2 norm of gradients: 1.06427616173313
l2 norm of weights: 6.462049799058988

Iteration Number: 831
Loss: 93.34458246602247
l2 norm of gradients: 1.0636601118154834
l2 norm of weights: 6.461797988544223

Iteration Number: 832
Loss: 93.2893799807899
l2 norm of gradients: 1.0630444572756168
l2 norm of weights: 6.4615463471507075

Iteration Number: 833
Loss: 93.23424114575047
l2 norm of gradients: 1.0624291987226673
l2 norm of weights: 6.4612948746792025

Iteration Number: 834
Loss: 93.17916588336364
l2 norm of gradients: 1.061814336738062
l2 norm of weights: 6.461043570930782

Iteration Number: 835
Loss: 93.1241541158888
l2 norm of gradients: 1.0611998718756932
l2 norm of weights: 6.4607924357068285

Iteration Number: 836
Loss: 93.06920576573027
l2 norm of gradients: 1.0605858046621008
l2 norm of weights: 6.46054146880904

Iteration Number: 837
Loss: 93.01432075528173
l2 norm of gradients: 1.0599721355966558
l2 norm of weights: 6.460290670039426

Iteration Number: 838
Loss: 92.95949900706216
l2 norm of gradients: 1.059358865151748
l2 norm of weights: 6.46004003920031

Iteration Number: 839
Loss: 92.90474044333696
l2 norm of gradients: 1.0587459937729777
l2 norm of weights: 6.459789576094322

Iteration Number: 840
Loss: 92.85004498666413
l2 norm of gradients: 1.0581335218793506
l2 norm of weights: 6.4595392805244085

Iteration Number: 841
Loss: 92.79541255964287
l2 norm of gradients: 1.0575214498634746
l2 norm of weights: 6.459289152293824

Iteration Number: 842
Loss: 92.74084308461462
l2 norm of gradients: 1.0569097780917618
l2 norm of weights: 6.459039191206136

Iteration Number: 843
Loss: 92.68633648433386
l2 norm of gradients: 1.0562985069046322
l2 norm of weights: 6.458789397065223

Iteration Number: 844
Loss: 92.63189268121795
l2 norm of gradients: 1.05568763661672
l2 norm of weights: 6.458539769675273

Iteration Number: 845
Loss: 92.57751159800041
l2 norm of gradients: 1.0550771675170858
l2 norm of weights: 6.458290308840784

Iteration Number: 846
Loss: 92.52319315736395
l2 norm of gradients: 1.0544670998694279

```

12 norm of weights: 6.4580410143665645
-----
Iteration Number: 847
Loss: 92.46893728202765
12 norm of gradients: 1.0538574339122986
12 norm of weights: 6.457791886057731
-----
Iteration Number: 848
Loss: 92.41474389478009
12 norm of gradients: 1.0532481698593212
12 norm of weights: 6.4575429237197115
-----
Iteration Number: 849
Loss: 92.36061291834592
12 norm of gradients: 1.0526393078994138
12 norm of weights: 6.4572941271582405
-----
Iteration Number: 850
Loss: 92.3065442756852
12 norm of gradients: 1.052030848197011
12 norm of weights: 6.457045496179362
-----
Iteration Number: 851
Loss: 92.25253788974294
12 norm of gradients: 1.0514227908922904
12 norm of weights: 6.456797030589428
-----
Iteration Number: 852
Loss: 92.19859368348857
12 norm of gradients: 1.0508151361014002
12 norm of weights: 6.456548730195094
-----
Iteration Number: 853
Loss: 92.14471157989104
12 norm of gradients: 1.0502078839166917
12 norm of weights: 6.456300594803329
-----
Iteration Number: 854
Loss: 92.09089150218117
12 norm of gradients: 1.0496010344069515
12 norm of weights: 6.456052624221401
-----
Iteration Number: 855
Loss: 92.0371333733528
12 norm of gradients: 1.0489945876176363
12 norm of weights: 6.4558048182568895
-----
Iteration Number: 856

```


Loss: 91.98343711677823
l2 norm of gradients: 1.0483885435711107
l2 norm of weights: 6.455557176717675

Iteration Number: 857
Loss: 91.92980265566622
l2 norm of gradients: 1.047782902266886
l2 norm of weights: 6.455309699411945

Iteration Number: 858
Loss: 91.87622991349285
l2 norm of gradients: 1.0471776636818622
l2 norm of weights: 6.455062386148189

Iteration Number: 859
Loss: 91.82271881370147
l2 norm of gradients: 1.0465728277705704
l2 norm of weights: 6.454815236735202

Iteration Number: 860
Loss: 91.76926927972077
l2 norm of gradients: 1.0459683944654181
l2 norm of weights: 6.4545682509820805

Iteration Number: 861
Loss: 91.71588123514854
l2 norm of gradients: 1.0453643636769354
l2 norm of weights: 6.454321428698223

Iteration Number: 862
Loss: 91.66255460367576
l2 norm of gradients: 1.0447607352940242
l2 norm of weights: 6.454074769693327

Iteration Number: 863
Loss: 91.60928930914585
l2 norm of gradients: 1.0441575091842068
l2 norm of weights: 6.453828273777398

Iteration Number: 864
Loss: 91.5560852754092
l2 norm of gradients: 1.0435546851938786
l2 norm of weights: 6.45358194076073

Iteration Number: 865
Loss: 91.50294242619584
l2 norm of gradients: 1.0429522631485586
l2 norm of weights: 6.453335770453927

```

-----
Iteration Number: 866
Loss: 91.44986068572783
l2 norm of gradients: 1.042350242853147
l2 norm of weights: 6.4530897626678865
-----
Iteration Number: 867
Loss: 91.39683997800881
l2 norm of gradients: 1.0417486240921783
l2 norm of weights: 6.452843917213802
-----
Iteration Number: 868
Loss: 91.34388022722706
l2 norm of gradients: 1.0411474066300779
l2 norm of weights: 6.452598233903168
-----
Iteration Number: 869
Loss: 91.29098135778193
l2 norm of gradients: 1.0405465902114208
l2 norm of weights: 6.452352712547771
-----
Iteration Number: 870
Loss: 91.23814329395736
l2 norm of gradients: 1.039946174561191
l2 norm of weights: 6.452107352959697
-----
Iteration Number: 871
Loss: 91.18536596014009
l2 norm of gradients: 1.0393461593850402
l2 norm of weights: 6.451862154951323
-----
Iteration Number: 872
Loss: 91.1326492811094
l2 norm of gradients: 1.0387465443695492
l2 norm of weights: 6.451617118335322
-----
Iteration Number: 873
Loss: 91.07999318144213
l2 norm of gradients: 1.0381473291824903
l2 norm of weights: 6.451372242924656
-----
Iteration Number: 874
Loss: 91.02739758593242
l2 norm of gradients: 1.0375485134730897
l2 norm of weights: 6.451127528532583
-----
Iteration Number: 875
Loss: 90.97486241944648

```

l2 norm of gradients: 1.0369500968722907
l2 norm of weights: 6.45088297497265

Iteration Number: 876
Loss: 90.92238760700046
l2 norm of gradients: 1.0363520789930176
l2 norm of weights: 6.450638582058695

Iteration Number: 877
Loss: 90.86997307369153
l2 norm of gradients: 1.0357544594304422
l2 norm of weights: 6.450394349604845

Iteration Number: 878
Loss: 90.8176187447168
l2 norm of gradients: 1.0351572377622462
l2 norm of weights: 6.450150277425512

Iteration Number: 879
Loss: 90.7653245454045
l2 norm of gradients: 1.0345604135488906
l2 norm of weights: 6.449906365335401

Iteration Number: 880
Loss: 90.71309040118801
l2 norm of gradients: 1.0339639863338796
l2 norm of weights: 6.4496626131495

Iteration Number: 881
Loss: 90.6609162376074
l2 norm of gradients: 1.0333679556440278
l2 norm of weights: 6.449419020683081

Iteration Number: 882
Loss: 90.60880198023595
l2 norm of gradients: 1.0327723209897282
l2 norm of weights: 6.449175587751702

Iteration Number: 883
Loss: 90.55674755502244
l2 norm of gradients: 1.0321770818652196
l2 norm of weights: 6.4489323141712065

Iteration Number: 884
Loss: 90.5047528876963
l2 norm of gradients: 1.0315822377488528
l2 norm of weights: 6.448689199757717

Iteration Number: 885
Loss: 90.45281790434606
l2 norm of gradients: 1.0309877881033602
l2 norm of weights: 6.448446244327638

Iteration Number: 886
Loss: 90.40094253102414
l2 norm of gradients: 1.0303937323761239
l2 norm of weights: 6.448203447697655

Iteration Number: 887
Loss: 90.34912669416735
l2 norm of gradients: 1.0298000699994407
l2 norm of weights: 6.447960809684734

Iteration Number: 888
Loss: 90.29737031997291
l2 norm of gradients: 1.0292068003907953
l2 norm of weights: 6.447718330106116

Iteration Number: 889
Loss: 90.245673335035
l2 norm of gradients: 1.0286139229531226
l2 norm of weights: 6.447476008779323

Iteration Number: 890
Loss: 90.19403566589843
l2 norm of gradients: 1.0280214370750806
l2 norm of weights: 6.447233845522148

Iteration Number: 891
Loss: 90.14245723940502
l2 norm of gradients: 1.0274293421313148
l2 norm of weights: 6.446991840152665

Iteration Number: 892
Loss: 90.09093798244493
l2 norm of gradients: 1.0268376374827282
l2 norm of weights: 6.446749992489218

Iteration Number: 893
Loss: 90.0394778219386
l2 norm of gradients: 1.0262463224767464
l2 norm of weights: 6.446508302350425

Iteration Number: 894
Loss: 89.98807668512052
l2 norm of gradients: 1.025655396447587

```

12 norm of weights: 6.446266769555176
-----
Iteration Number: 895
Loss: 89.93673449918235
12 norm of gradients: 1.025064858716524
12 norm of weights: 6.44602539392263
-----
Iteration Number: 896
Loss: 89.88545119162143
12 norm of gradients: 1.0244747085921566
12 norm of weights: 6.445784175272216
-----
Iteration Number: 897
Loss: 89.83422668998291
12 norm of gradients: 1.0238849453706727
12 norm of weights: 6.445543113423634
-----
Iteration Number: 898
Loss: 89.78306092181194
12 norm of gradients: 1.0232955683361167
12 norm of weights: 6.445302208196849
-----
Iteration Number: 899
Loss: 89.73195381510197
12 norm of gradients: 1.0227065767606516
12 norm of weights: 6.44506145941209
-----
Iteration Number: 900
Loss: 89.68090529756849
12 norm of gradients: 1.0221179699048268
12 norm of weights: 6.444820866889855
-----
Iteration Number: 901
Loss: 89.62991529757848
12 norm of gradients: 1.021529747017839
12 norm of weights: 6.444580430450903
-----
Iteration Number: 902
Loss: 89.57898374310159
12 norm of gradients: 1.0209419073377968
12 norm of weights: 6.444340149916257
-----
Iteration Number: 903
Loss: 89.52811056265764
12 norm of gradients: 1.0203544500919821
12 norm of weights: 6.4441000251072
-----
Iteration Number: 904

```

Loss: 89.47729568468544
l2 norm of gradients: 1.0197673744971127
l2 norm of weights: 6.443860055845276

Iteration Number: 905
Loss: 89.42653903791027
l2 norm of gradients: 1.019180679759603
l2 norm of weights: 6.443620241952288

Iteration Number: 906
Loss: 89.3758405509401
l2 norm of gradients: 1.0185943650758231
l2 norm of weights: 6.443380583250297

Iteration Number: 907
Loss: 89.32520015290235
l2 norm of gradients: 1.01800842963236
l2 norm of weights: 6.443141079561621

Iteration Number: 908
Loss: 89.2746177728592
l2 norm of gradients: 1.0174228726062728
l2 norm of weights: 6.442901730708833

Iteration Number: 909
Loss: 89.22409333984753
l2 norm of gradients: 1.016837693165354
l2 norm of weights: 6.442662536514757

Iteration Number: 910
Loss: 89.17362678341745
l2 norm of gradients: 1.0162528904683825
l2 norm of weights: 6.442423496802477

Iteration Number: 911
Loss: 89.1232180329952
l2 norm of gradients: 1.0156684636653808
l2 norm of weights: 6.442184611395324

Iteration Number: 912
Loss: 89.07286701821592
l2 norm of gradients: 1.015084411897869
l2 norm of weights: 6.441945880116878

Iteration Number: 913
Loss: 89.02257366890368
l2 norm of gradients: 1.0145007342991177
l2 norm of weights: 6.441707302790973

```

-----
Iteration Number: 914
Loss: 88.97233791508502
l2 norm of gradients: 1.0139174299944012
l2 norm of weights: 6.441468879241688
-----
Iteration Number: 915
Loss: 88.9221596866778
l2 norm of gradients: 1.0133344981012467
l2 norm of weights: 6.441230609293349
-----
Iteration Number: 916
Loss: 88.87203891410748
l2 norm of gradients: 1.0127519377296863
l2 norm of weights: 6.44099249277053
-----
Iteration Number: 917
Loss: 88.82197552758328
l2 norm of gradients: 1.0121697479825045
l2 norm of weights: 6.440754529498048
-----
Iteration Number: 918
Loss: 88.7719694578372
l2 norm of gradients: 1.011587927955485
l2 norm of weights: 6.44051671930096
-----
Iteration Number: 919
Loss: 88.72202063546419
l2 norm of gradients: 1.011006476737658
l2 norm of weights: 6.440279062004571
-----
Iteration Number: 920
Loss: 88.67212899121546
l2 norm of gradients: 1.010425393411545
l2 norm of weights: 6.440041557434421
-----
Iteration Number: 921
Loss: 88.62229445626848
l2 norm of gradients: 1.0098446770534015
l2 norm of weights: 6.439804205416294
-----
Iteration Number: 922
Loss: 88.57251696167317
l2 norm of gradients: 1.0092643267334591
l2 norm of weights: 6.43956700577621
-----
Iteration Number: 923
Loss: 88.52279643862906

```

```

12 norm of gradients: 1.008684341516168
12 norm of weights: 6.439329958340424
-----
Iteration Number: 924
Loss: 88.47313281876546
12 norm of gradients: 1.0081047204604332
12 norm of weights: 6.439093062935432
-----
Iteration Number: 925
Loss: 88.42352603349742
12 norm of gradients: 1.007525462619856
12 norm of weights: 6.438856319387957
-----
Iteration Number: 926
Loss: 88.37397601467116
12 norm of gradients: 1.0069465670429678
12 norm of weights: 6.438619727524963
-----
Iteration Number: 927
Loss: 88.324482694181
12 norm of gradients: 1.0063680327734663
12 norm of weights: 6.43838328717364
-----
Iteration Number: 928
Loss: 88.27504600407781
12 norm of gradients: 1.0057898588504477
12 norm of weights: 6.438146998161413
-----
Iteration Number: 929
Loss: 88.2256658765721
12 norm of gradients: 1.00521204430864
12 norm of weights: 6.437910860315932
-----
Iteration Number: 930
Loss: 88.17634224402197
12 norm of gradients: 1.0046345881786336
12 norm of weights: 6.437674873465078
-----
Iteration Number: 931
Loss: 88.12707503884604
12 norm of gradients: 1.0040574894871075
12 norm of weights: 6.43743903743696
-----
Iteration Number: 932
Loss: 88.0778641938028
12 norm of gradients: 1.0034807472570588
12 norm of weights: 6.437203352059909
-----

```


Iteration Number: 933
Loss: 88.02870964170017
l2 norm of gradients: 1.0029043605080274
l2 norm of weights: 6.436967817162484

Iteration Number: 934
Loss: 87.97961131552022
l2 norm of gradients: 1.0023283282563205
l2 norm of weights: 6.436732432573463

Iteration Number: 935
Loss: 87.93056914829505
l2 norm of gradients: 1.0017526495152327
l2 norm of weights: 6.436497198121852

Iteration Number: 936
Loss: 87.88158307338233
l2 norm of gradients: 1.001177323295269
l2 norm of weights: 6.436262113636872

Iteration Number: 937
Loss: 87.83265302422292
l2 norm of gradients: 1.0006023486043618
l2 norm of weights: 6.436027178947966

Iteration Number: 938
Loss: 87.78377893429976
l2 norm of gradients: 1.0000277244480886
l2 norm of weights: 6.435792393884793

Iteration Number: 939
Loss: 87.73496073749402
l2 norm of gradients: 0.9994534498298876
l2 norm of weights: 6.435557758277233

Iteration Number: 940
Loss: 87.68619836760394
l2 norm of gradients: 0.9988795237512696
l2 norm of weights: 6.435323271955378

Iteration Number: 941
Loss: 87.63749175862927
l2 norm of gradients: 0.9983059452120318
l2 norm of weights: 6.435088934749535

Iteration Number: 942
Loss: 87.58884084484588
l2 norm of gradients: 0.9977327132104655

```
l2 norm of weights: 6.434854746490226
-----
Iteration Number: 943
Loss: 87.54024556052002
l2 norm of gradients: 0.9971598267435668
l2 norm of weights: 6.434620707008183
-----
Iteration Number: 944
Loss: 87.49170584024782
l2 norm of gradients: 0.9965872848072398
l2 norm of weights: 6.434386816134349
-----
Iteration Number: 945
Loss: 87.44322161867682
l2 norm of gradients: 0.9960150863965034
l2 norm of weights: 6.434153073699878
-----
Iteration Number: 946
Loss: 87.39479283052808
l2 norm of gradients: 0.9954432305056922
l2 norm of weights: 6.433919479536131
-----
Iteration Number: 947
Loss: 87.34641941075931
l2 norm of gradients: 0.9948717161286582
l2 norm of weights: 6.433686033474676
-----
Iteration Number: 948
Loss: 87.29810129456753
l2 norm of gradients: 0.994300542258969
l2 norm of weights: 6.433452735347286
-----
Iteration Number: 949
Loss: 87.24983841721834
l2 norm of gradients: 0.9937297078901043
l2 norm of weights: 6.433219584985942
-----
Iteration Number: 950
Loss: 87.20163071404349
l2 norm of gradients: 0.9931592120156514
l2 norm of weights: 6.432986582222823
-----
Iteration Number: 951
Loss: 87.15347812059846
l2 norm of gradients: 0.9925890536294976
l2 norm of weights: 6.432753726890316
-----
Iteration Number: 952
```

```
Loss: 87.10538057270082
12 norm of gradients: 0.9920192317260214
12 norm of weights: 6.432521018821004
-----
Iteration Number: 953
Loss: 87.05733800611942
12 norm of gradients: 0.9914497453002822
12 norm of weights: 6.432288457847673
-----
Iteration Number: 954
Loss: 87.00935035687681
12 norm of gradients: 0.9908805933482058
12 norm of weights: 6.432056043803307
-----
Iteration Number: 955
Loss: 86.96141756122648
12 norm of gradients: 0.9903117748667708
12 norm of weights: 6.431823776521086
-----
Iteration Number: 956
Loss: 86.91353955540666
12 norm of gradients: 0.989743288854191
12 norm of weights: 6.431591655834389
-----
Iteration Number: 957
Loss: 86.86571627596749
12 norm of gradients: 0.9891751343100975
12 norm of weights: 6.431359681576785
-----
Iteration Number: 958
Loss: 86.817947659344
12 norm of gradients: 0.9886073102357159
12 norm of weights: 6.431127853582042
-----
Iteration Number: 959
Loss: 86.77023364249115
12 norm of gradients: 0.9880398156340453
12 norm of weights: 6.430896171684119
-----
Iteration Number: 960
Loss: 86.72257416230798
12 norm of gradients: 0.9874726495100316
12 norm of weights: 6.430664635717165
-----
Iteration Number: 961
Loss: 86.67496915569437
12 norm of gradients: 0.9869058108707421
12 norm of weights: 6.430433245515522
```

```

-----
Iteration Number: 962
Loss: 86.62741856001877
l2 norm of gradients: 0.9863392987255356
l2 norm of weights: 6.430202000913719
-----
Iteration Number: 963
Loss: 86.57992231256063
l2 norm of gradients: 0.9857731120862306
l2 norm of weights: 6.429970901746474
-----
Iteration Number: 964
Loss: 86.5324803509059
l2 norm of gradients: 0.9852072499672737
l2 norm of weights: 6.429739947848692
-----
Iteration Number: 965
Loss: 86.48509261253155
l2 norm of gradients: 0.9846417113859032
l2 norm of weights: 6.429509139055462
-----
Iteration Number: 966
Loss: 86.43775903539645
l2 norm of gradients: 0.984076495362312
l2 norm of weights: 6.429278475202061
-----
Iteration Number: 967
Loss: 86.39047955736584
l2 norm of gradients: 0.9835116009198084
l2 norm of weights: 6.429047956123946
-----
Iteration Number: 968
Loss: 86.34325411649934
l2 norm of gradients: 0.9829470270849748
l2 norm of weights: 6.4288175816567605
-----
Iteration Number: 969
Loss: 86.29608265105165
l2 norm of gradients: 0.9823827728878233
l2 norm of weights: 6.428587351636324
-----
Iteration Number: 970
Loss: 86.24896509936916
l2 norm of gradients: 0.9818188373619513
l2 norm of weights: 6.428357265898642
-----
Iteration Number: 971
Loss: 86.20190139997595

```

l2 norm of gradients: 0.9812552195446921
l2 norm of weights: 6.428127324279895

Iteration Number: 972
Loss: 86.15489149152252
l2 norm of gradients: 0.9806919184772676
l2 norm of weights: 6.427897526616442

Iteration Number: 973
Loss: 86.10793531289573
l2 norm of gradients: 0.9801289332049332
l2 norm of weights: 6.427667872744821

Iteration Number: 974
Loss: 86.06103280285828
l2 norm of gradients: 0.9795662627771262
l2 norm of weights: 6.427438362501742

Iteration Number: 975
Loss: 86.01418390065517
l2 norm of gradients: 0.9790039062476085
l2 norm of weights: 6.4272089957240945

Iteration Number: 976
Loss: 85.96738854533554
l2 norm of gradients: 0.9784418626746084
l2 norm of weights: 6.426979772248939

Iteration Number: 977
Loss: 85.92064667647222
l2 norm of gradients: 0.9778801311209608
l2 norm of weights: 6.42675069191351

Iteration Number: 978
Loss: 85.87395823329379
l2 norm of gradients: 0.9773187106542436
l2 norm of weights: 6.4265217545552105

Iteration Number: 979
Loss: 85.82732315565319
l2 norm of gradients: 0.9767576003469146
l2 norm of weights: 6.426292960011617

Iteration Number: 980
Loss: 85.7807413831445
l2 norm of gradients: 0.9761967992764433
l2 norm of weights: 6.426064308120476

Iteration Number: 981
Loss: 85.73421285585584
l2 norm of gradients: 0.9756363065254428
l2 norm of weights: 6.425835798719701

Iteration Number: 982
Loss: 85.68773751363901
l2 norm of gradients: 0.9750761211817979
l2 norm of weights: 6.425607431647372

Iteration Number: 983
Loss: 85.64131529676531
l2 norm of gradients: 0.9745162423387943
l2 norm of weights: 6.425379206741739

Iteration Number: 984
Loss: 85.59494614554843
l2 norm of gradients: 0.9739566690952405
l2 norm of weights: 6.425151123841213

Iteration Number: 985
Loss: 85.54863000041819
l2 norm of gradients: 0.9733974005555921
l2 norm of weights: 6.424923182784374

Iteration Number: 986
Loss: 85.50236680196362
l2 norm of gradients: 0.9728384358300718
l2 norm of weights: 6.424695383409963

Iteration Number: 987
Loss: 85.45615649088474
l2 norm of gradients: 0.9722797740347889
l2 norm of weights: 6.424467725556883

Iteration Number: 988
Loss: 85.40999900794024
l2 norm of gradients: 0.9717214142918537
l2 norm of weights: 6.424240209064199

Iteration Number: 989
Loss: 85.36389429427304
l2 norm of gradients: 0.9711633557294929
l2 norm of weights: 6.4240128337711395

Iteration Number: 990
Loss: 85.31784229080255
l2 norm of gradients: 0.9706055974821619

```

12 norm of weights: 6.423785599517088
-----
Iteration Number: 991
Loss: 85.27184293885027
12 norm of gradients: 0.9700481386906543
12 norm of weights: 6.423558506141591
-----
Iteration Number: 992
Loss: 85.22589617969375
12 norm of gradients: 0.9694909785022092
12 norm of weights: 6.423331553484349
-----
Iteration Number: 993
Loss: 85.18000195496994
12 norm of gradients: 0.9689341160706182
12 norm of weights: 6.423104741385221
-----
Iteration Number: 994
Loss: 85.13416020612861
12 norm of gradients: 0.9683775505563278
12 norm of weights: 6.422878069684225
-----
Iteration Number: 995
Loss: 85.08837087498641
12 norm of gradients: 0.9678212811265424
12 norm of weights: 6.422651538221528
-----
Iteration Number: 996
Loss: 85.04263390332754
12 norm of gradients: 0.9672653069553222
12 norm of weights: 6.4224251468374565
-----
Iteration Number: 997
Loss: 84.99694923330213
12 norm of gradients: 0.9667096272236815
12 norm of weights: 6.422198895372488
-----
Iteration Number: 998
Loss: 84.95131680684145
12 norm of gradients: 0.9661542411196835
12 norm of weights: 6.421972783667251
-----
Iteration Number: 999
Loss: 84.90573656619013
12 norm of gradients: 0.9655991478385345
12 norm of weights: 6.42174681156253
-----
Iteration Number: 1000

```

Loss: 84.86020845385194
l2 norm of gradients: 0.9650443465826732
l2 norm of weights: 6.421520978899253

Iteration Number: 1001
Loss: 84.81473241205777
l2 norm of gradients: 0.9644898365618617
l2 norm of weights: 6.421295285518506

Iteration Number: 1002
Loss: 84.76930838364794
l2 norm of gradients: 0.9639356169932709
l2 norm of weights: 6.421069731261518

Iteration Number: 1003
Loss: 84.72393631116971
l2 norm of gradients: 0.9633816871015664
l2 norm of weights: 6.420844315969669

Iteration Number: 1004
Loss: 84.6786161375317
l2 norm of gradients: 0.9628280461189911
l2 norm of weights: 6.420619039484485

Iteration Number: 1005
Loss: 84.63334780559556
l2 norm of gradients: 0.9622746932854455
l2 norm of weights: 6.42039390164764

Iteration Number: 1006
Loss: 84.58813125850831
l2 norm of gradients: 0.9617216278485672
l2 norm of weights: 6.420168902300951

Iteration Number: 1007
Loss: 84.54296643936281
l2 norm of gradients: 0.9611688490638072
l2 norm of weights: 6.419944041286382

Iteration Number: 1008
Loss: 84.49785329153897
l2 norm of gradients: 0.9606163561945044
l2 norm of weights: 6.419719318446043

Iteration Number: 1009
Loss: 84.45279175844078
l2 norm of gradients: 0.9600641485119591
l2 norm of weights: 6.419494733622181


```

-----
Iteration Number: 1010
Loss: 84.40778178349768
l2 norm of gradients: 0.959512225295502
l2 norm of weights: 6.419270286657193
-----
Iteration Number: 1011
Loss: 84.3628233103463
l2 norm of gradients: 0.9589605858325647
l2 norm of weights: 6.419045977393614
-----
Iteration Number: 1012
Loss: 84.31791628280979
l2 norm of gradients: 0.958409229418745
l2 norm of weights: 6.41882180567412
-----
Iteration Number: 1013
Loss: 84.27306064477578
l2 norm of gradients: 0.957858155357872
l2 norm of weights: 6.418597771341527
-----
Iteration Number: 1014
Loss: 84.22825633996284
l2 norm of gradients: 0.9573073629620685
l2 norm of weights: 6.418373874238793
-----
Iteration Number: 1015
Loss: 84.18350331272708
l2 norm of gradients: 0.9567568515518121
l2 norm of weights: 6.418150114209013
-----
Iteration Number: 1016
Loss: 84.13880150699029
l2 norm of gradients: 0.9562066204559933
l2 norm of weights: 6.417926491095419
-----
Iteration Number: 1017
Loss: 84.09415086719738
l2 norm of gradients: 0.9556566690119723
l2 norm of weights: 6.417703004741384
-----
Iteration Number: 1018
Loss: 84.0495513377477
l2 norm of gradients: 0.955106996565634
l2 norm of weights: 6.417479654990415
-----
Iteration Number: 1019
Loss: 84.00500286303384

```

l2 norm of gradients: 0.9545576024714407
l2 norm of weights: 6.417256441686156

Iteration Number: 1020
Loss: 83.9605053876662
l2 norm of gradients: 0.9540084860924823
l2 norm of weights: 6.4170333646723865

Iteration Number: 1021
Loss: 83.91605885632757
l2 norm of gradients: 0.9534596468005269
l2 norm of weights: 6.416810423793022

Iteration Number: 1022
Loss: 83.87166321390566
l2 norm of gradients: 0.9529110839760649
l2 norm of weights: 6.416587618892108

Iteration Number: 1023
Loss: 83.82731840521977
l2 norm of gradients: 0.9523627970083578
l2 norm of weights: 6.416364949813829

Iteration Number: 1024
Loss: 83.78302437535095
l2 norm of gradients: 0.9518147852954779
l2 norm of weights: 6.4161424164025

Iteration Number: 1025
Loss: 83.73878106922412
l2 norm of gradients: 0.9512670482443514
l2 norm of weights: 6.415920018502565

Iteration Number: 1026
Loss: 83.69458843219839
l2 norm of gradients: 0.9507195852707977
l2 norm of weights: 6.415697755958606

Iteration Number: 1027
Loss: 83.65044640949726
l2 norm of gradients: 0.9501723957995664
l2 norm of weights: 6.415475628615329

Iteration Number: 1028
Loss: 83.60635494658074
l2 norm of gradients: 0.9496254792643734
l2 norm of weights: 6.415253636317579

Iteration Number: 1029
Loss: 83.56231398885677
l2 norm of gradients: 0.949078835107934
l2 norm of weights: 6.415031778910322

Iteration Number: 1030
Loss: 83.51832348180366
l2 norm of gradients: 0.9485324627819961
l2 norm of weights: 6.414810056238659

Iteration Number: 1031
Loss: 83.4743833713588
l2 norm of gradients: 0.9479863617473686
l2 norm of weights: 6.414588468147818

Iteration Number: 1032
Loss: 83.4304936030289
l2 norm of gradients: 0.9474405314739512
l2 norm of weights: 6.414367014483156

Iteration Number: 1033
Loss: 83.38665412274689
l2 norm of gradients: 0.9468949714407605
l2 norm of weights: 6.4141456950901565

Iteration Number: 1034
Loss: 83.34286487653786
l2 norm of gradients: 0.9463496811359536
l2 norm of weights: 6.4139245098144295

Iteration Number: 1035
Loss: 83.29912581027588
l2 norm of gradients: 0.9458046600568529
l2 norm of weights: 6.413703458501713

Iteration Number: 1036
Loss: 83.25543687023746
l2 norm of gradients: 0.9452599077099656
l2 norm of weights: 6.413482540997872

Iteration Number: 1037
Loss: 83.21179800255739
l2 norm of gradients: 0.9447154236110039
l2 norm of weights: 6.413261757148897

Iteration Number: 1038
Loss: 83.16820915352835
l2 norm of gradients: 0.9441712072849022

```

12 norm of weights: 6.4130411068008994
-----
Iteration Number: 1039
Loss: 83.12467026952537
12 norm of gradients: 0.9436272582658332
12 norm of weights: 6.412820589800121
-----
Iteration Number: 1040
Loss: 83.08118129702136
12 norm of gradients: 0.9430835760972218
12 norm of weights: 6.412600205992923
-----
Iteration Number: 1041
Loss: 83.0377421826404
12 norm of gradients: 0.9425401603317579
12 norm of weights: 6.412379955225794
-----
Iteration Number: 1042
Loss: 82.99435287284793
12 norm of gradients: 0.9419970105314066
12 norm of weights: 6.412159837345344
-----
Iteration Number: 1043
Loss: 82.95101331444678
12 norm of gradients: 0.9414541262674172
12 norm of weights: 6.411939852198305
-----
Iteration Number: 1044
Loss: 82.9077234542213
12 norm of gradients: 0.9409115071203316
12 norm of weights: 6.411719999631535
-----
Iteration Number: 1045
Loss: 82.86448323905039
12 norm of gradients: 0.9403691526799884
12 norm of weights: 6.411500279492008
-----
Iteration Number: 1046
Loss: 82.82129261596002
12 norm of gradients: 0.9398270625455278
12 norm of weights: 6.411280691626826
-----
Iteration Number: 1047
Loss: 82.77815153186745
12 norm of gradients: 0.9392852363253938
12 norm of weights: 6.411061235883208
-----
Iteration Number: 1048

```

```

Loss: 82.73505993386912
l2 norm of gradients: 0.9387436736373355
l2 norm of weights: 6.410841912108494
-----
Iteration Number: 1049
Loss: 82.69201776919786
l2 norm of gradients: 0.9382023741084047
l2 norm of weights: 6.410622720150146
-----
Iteration Number: 1050
Loss: 82.6490249851053
l2 norm of gradients: 0.9376613373749547
l2 norm of weights: 6.410403659855745
-----
Iteration Number: 1051
Loss: 82.60608152885219
l2 norm of gradients: 0.9371205630826367
l2 norm of weights: 6.41018473107299
-----
Iteration Number: 1052
Loss: 82.56318734798894
l2 norm of gradients: 0.9365800508863923
l2 norm of weights: 6.409965933649705
-----
Iteration Number: 1053
Loss: 82.5203423899318
l2 norm of gradients: 0.9360398004504477
l2 norm of weights: 6.409747267433826
-----
Iteration Number: 1054
Loss: 82.47754660212979
l2 norm of gradients: 0.9354998114483044
l2 norm of weights: 6.409528732273411
-----
Iteration Number: 1055
Loss: 82.43479993230356
l2 norm of gradients: 0.9349600835627299
l2 norm of weights: 6.409310328016635
-----
Iteration Number: 1056
Loss: 82.39210232812694
l2 norm of gradients: 0.9344206164857437
l2 norm of weights: 6.4090920545117935
-----
Iteration Number: 1057
Loss: 82.34945373736367
l2 norm of gradients: 0.9338814099186067
l2 norm of weights: 6.408873911607296

```

```

-----
Iteration Number: 1058
Loss: 82.30685410784406
l2 norm of gradients: 0.9333424635718041
l2 norm of weights: 6.408655899151673
-----
Iteration Number: 1059
Loss: 82.2643033873758
l2 norm of gradients: 0.9328037771650304
l2 norm of weights: 6.408438016993569
-----
Iteration Number: 1060
Loss: 82.22180152399743
l2 norm of gradients: 0.9322653504271708
l2 norm of weights: 6.408220264981746
-----
Iteration Number: 1061
Loss: 82.17934846573458
l2 norm of gradients: 0.9317271830962833
l2 norm of weights: 6.408002642965083
-----
Iteration Number: 1062
Loss: 82.13694416067611
l2 norm of gradients: 0.9311892749195765
l2 norm of weights: 6.407785150792576
-----
Iteration Number: 1063
Loss: 82.09458855689607
l2 norm of gradients: 0.9306516256533884
l2 norm of weights: 6.407567788313335
-----
Iteration Number: 1064
Loss: 82.05228160268588
l2 norm of gradients: 0.9301142350631627
l2 norm of weights: 6.407350555376588
-----
Iteration Number: 1065
Loss: 82.01002324632628
l2 norm of gradients: 0.9295771029234244
l2 norm of weights: 6.407133451831676
-----
Iteration Number: 1066
Loss: 81.96781343606928
l2 norm of gradients: 0.9290402290177525
l2 norm of weights: 6.406916477528057
-----
Iteration Number: 1067
Loss: 81.92565212038667

```

l2 norm of gradients: 0.9285036131387536
l2 norm of weights: 6.406699632315302

Iteration Number: 1068
Loss: 81.8835392477495
l2 norm of gradients: 0.9279672550880322
l2 norm of weights: 6.406482916043101

Iteration Number: 1069
Loss: 81.84147476657482
l2 norm of gradients: 0.9274311546761607
l2 norm of weights: 6.406266328561255

Iteration Number: 1070
Loss: 81.79945862553042
l2 norm of gradients: 0.9268953117226472
l2 norm of weights: 6.406049869719679

Iteration Number: 1071
Loss: 81.75749077314936
l2 norm of gradients: 0.9263597260559038
l2 norm of weights: 6.4058335393684045

Iteration Number: 1072
Loss: 81.71557115820472
l2 norm of gradients: 0.9258243975132117
l2 norm of weights: 6.405617337357576

Iteration Number: 1073
Loss: 81.67369972943233
l2 norm of gradients: 0.9252893259406852
l2 norm of weights: 6.405401263537453

Iteration Number: 1074
Loss: 81.63187643554576
l2 norm of gradients: 0.9247545111932362
l2 norm of weights: 6.405185317758406

Iteration Number: 1075
Loss: 81.59010122538552
l2 norm of gradients: 0.9242199531345352
l2 norm of weights: 6.404969499870921

Iteration Number: 1076
Loss: 81.5483740479338
l2 norm of gradients: 0.9236856516369728
l2 norm of weights: 6.4047538097256

Iteration Number: 1077
Loss: 81.50669485214522
l2 norm of gradients: 0.9231516065816185
l2 norm of weights: 6.404538247173151

Iteration Number: 1078
Loss: 81.46506358693036
l2 norm of gradients: 0.9226178178581808
l2 norm of weights: 6.404322812064404

Iteration Number: 1079
Loss: 81.42348020135069
l2 norm of gradients: 0.9220842853649621
l2 norm of weights: 6.4041075042502955

Iteration Number: 1080
Loss: 81.38194464462013
l2 norm of gradients: 0.9215510090088168
l2 norm of weights: 6.403892323581878

Iteration Number: 1081
Loss: 81.3404568657985
l2 norm of gradients: 0.9210179887051054
l2 norm of weights: 6.4036772699103155

Iteration Number: 1082
Loss: 81.29901681405097
l2 norm of gradients: 0.9204852243776479
l2 norm of weights: 6.403462343086885

Iteration Number: 1083
Loss: 81.25762443862192
l2 norm of gradients: 0.9199527159586778
l2 norm of weights: 6.403247542962978

Iteration Number: 1084
Loss: 81.21627968879322
l2 norm of gradients: 0.9194204633887916
l2 norm of weights: 6.403032869390096

Iteration Number: 1085
Loss: 81.17498251398841
l2 norm of gradients: 0.9188884666169024
l2 norm of weights: 6.402818322219855

Iteration Number: 1086
Loss: 81.13373286349326
l2 norm of gradients: 0.9183567256001858


```
12 norm of weights: 6.402603901303983
-----
Iteration Number: 1087
Loss: 81.09253068681356
12 norm of gradients: 0.9178252403040315
12 norm of weights: 6.402389606494317
-----
Iteration Number: 1088
Loss: 81.05137593334389
12 norm of gradients: 0.9172940107019882
12 norm of weights: 6.4021754376428115
-----
Iteration Number: 1089
Loss: 81.01026855257894
12 norm of gradients: 0.9167630367757117
12 norm of weights: 6.401961394601531
-----
Iteration Number: 1090
Loss: 80.969208494083
12 norm of gradients: 0.9162323185149096
12 norm of weights: 6.401747477222652
-----
Iteration Number: 1091
Loss: 80.92819570751044
12 norm of gradients: 0.9157018559172855
12 norm of weights: 6.401533685358463
-----
Iteration Number: 1092
Loss: 80.88723014233682
12 norm of gradients: 0.9151716489884824
12 norm of weights: 6.401320018861367
-----
Iteration Number: 1093
Loss: 80.84631174847848
12 norm of gradients: 0.9146416977420257
12 norm of weights: 6.401106477583875
-----
Iteration Number: 1094
Loss: 80.80544047539678
12 norm of gradients: 0.9141120021992645
12 norm of weights: 6.400893061378615
-----
Iteration Number: 1095
Loss: 80.76461627292213
12 norm of gradients: 0.9135825623893121
12 norm of weights: 6.400679770098324
-----
Iteration Number: 1096
```

Loss: 80.72383909091812
l2 norm of gradients: 0.9130533783489856
l2 norm of weights: 6.400466603595851

Iteration Number: 1097
Loss: 80.6831088791208
l2 norm of gradients: 0.9125244501227451
l2 norm of weights: 6.40025356172416

Iteration Number: 1098
Loss: 80.64242558739645
l2 norm of gradients: 0.9119957777626317
l2 norm of weights: 6.4000406443363245

Iteration Number: 1099
Loss: 80.60178916562288
l2 norm of gradients: 0.9114673613282038
l2 norm of weights: 6.399827851285531

Iteration Number: 1100
Loss: 80.5611995637418
l2 norm of gradients: 0.9109392008864747
l2 norm of weights: 6.399615182425079

Iteration Number: 1101
Loss: 80.52065673170804
l2 norm of gradients: 0.9104112965118472
l2 norm of weights: 6.3994026376083815

Iteration Number: 1102
Loss: 80.48016061951606
l2 norm of gradients: 0.9098836482860483
l2 norm of weights: 6.39919021668896

Iteration Number: 1103
Loss: 80.43971117719053
l2 norm of gradients: 0.9093562562980637
l2 norm of weights: 6.398977919520451

Iteration Number: 1104
Loss: 80.39930835487607
l2 norm of gradients: 0.9088291206440704
l2 norm of weights: 6.398765745956605

Iteration Number: 1105
Loss: 80.35895210249676
l2 norm of gradients: 0.9083022414273689
l2 norm of weights: 6.398553695851282

Iteration Number: 1106
Loss: 80.31864237026033
l2 norm of gradients: 0.9077756187583159
l2 norm of weights: 6.398341769058455

Iteration Number: 1107
Loss: 80.27837910831259
l2 norm of gradients: 0.9072492527542542
l2 norm of weights: 6.398129965432212

Iteration Number: 1108
Loss: 80.23816226692412
l2 norm of gradients: 0.9067231435394427
l2 norm of weights: 6.397918284826752

Iteration Number: 1109
Loss: 80.19799179614192
l2 norm of gradients: 0.906197291244988
l2 norm of weights: 6.397706727096388

Iteration Number: 1110
Loss: 80.15786764628872
l2 norm of gradients: 0.9056716960087705
l2 norm of weights: 6.397495292095544

Iteration Number: 1111
Loss: 80.11778976765913
l2 norm of gradients: 0.9051463579753752
l2 norm of weights: 6.39728397967876

Iteration Number: 1112
Loss: 80.07775811048333
l2 norm of gradients: 0.9046212772960175
l2 norm of weights: 6.397072789700687

Iteration Number: 1113
Loss: 80.0377726250336
l2 norm of gradients: 0.9040964541284712
l2 norm of weights: 6.396861722016088

Iteration Number: 1114
Loss: 79.99783326182522
l2 norm of gradients: 0.9035718886369947
l2 norm of weights: 6.396650776479843

Iteration Number: 1115
Loss: 79.95793997106671

l2 norm of gradients: 0.9030475809922569
l2 norm of weights: 6.396439952946943

Iteration Number: 1116
Loss: 79.9180927032396
l2 norm of gradients: 0.9025235313712624
l2 norm of weights: 6.396229251272494

Iteration Number: 1117
Loss: 79.87829140876495
l2 norm of gradients: 0.9019997399572763
l2 norm of weights: 6.396018671311714

Iteration Number: 1118
Loss: 79.83853603809075
l2 norm of gradients: 0.9014762069397491
l2 norm of weights: 6.395808212919938

Iteration Number: 1119
Loss: 79.79882654161952
l2 norm of gradients: 0.9009529325142387
l2 norm of weights: 6.395597875952612

Iteration Number: 1120
Loss: 79.75916286990831
l2 norm of gradients: 0.9004299168823353
l2 norm of weights: 6.395387660265297

Iteration Number: 1121
Loss: 79.71954497349022
l2 norm of gradients: 0.899907160251584
l2 norm of weights: 6.39517756571367

Iteration Number: 1122
Loss: 79.67997280289505
l2 norm of gradients: 0.8993846628354063
l2 norm of weights: 6.394967592153519

Iteration Number: 1123
Loss: 79.64044630861109
l2 norm of gradients: 0.8988624248530223
l2 norm of weights: 6.3947577394407515

Iteration Number: 1124
Loss: 79.60096544129642
l2 norm of gradients: 0.8983404465293725
l2 norm of weights: 6.394548007431385

Iteration Number: 1125
Loss: 79.56153015148223
l2 norm of gradients: 0.8978187280950389
l2 norm of weights: 6.3943383959815545

Iteration Number: 1126
Loss: 79.52214038986196
l2 norm of gradients: 0.8972972697861646
l2 norm of weights: 6.394128904947512

Iteration Number: 1127
Loss: 79.48279610707723
l2 norm of gradients: 0.8967760718443757
l2 norm of weights: 6.39391953418562

Iteration Number: 1128
Loss: 79.44349725373145
l2 norm of gradients: 0.8962551345167
l2 norm of weights: 6.39371028355236

Iteration Number: 1129
Loss: 79.40424378056335
l2 norm of gradients: 0.895734458055487
l2 norm of weights: 6.393501152904328

Iteration Number: 1130
Loss: 79.36503563820806
l2 norm of gradients: 0.8952140427183273
l2 norm of weights: 6.39329214209824

Iteration Number: 1131
Loss: 79.3258727774627
l2 norm of gradients: 0.8946938887679714
l2 norm of weights: 6.393083250990919

Iteration Number: 1132
Loss: 79.28675514898187
l2 norm of gradients: 0.8941739964722488
l2 norm of weights: 6.392874479439313

Iteration Number: 1133
Loss: 79.24768270360347
l2 norm of gradients: 0.8936543661039857
l2 norm of weights: 6.392665827300483

Iteration Number: 1134
Loss: 79.20865539200892
l2 norm of gradients: 0.8931349979409241

```

12 norm of weights: 6.3924572944316065
-----
Iteration Number: 1135
Loss: 79.16967316499516
12 norm of gradients: 0.8926158922656386
12 norm of weights: 6.392248880689977
-----
Iteration Number: 1136
Loss: 79.1307359734497
12 norm of gradients: 0.8920970493654556
12 norm of weights: 6.392040585933009
-----
Iteration Number: 1137
Loss: 79.09184376817986
12 norm of gradients: 0.8915784695323691
12 norm of weights: 6.39183241001823
-----
Iteration Number: 1138
Loss: 79.0529964998724
12 norm of gradients: 0.8910601530629596
12 norm of weights: 6.391624352803287
-----
Iteration Number: 1139
Loss: 79.01419411953067
12 norm of gradients: 0.8905421002583108
12 norm of weights: 6.391416414145945
-----
Iteration Number: 1140
Loss: 78.97543657806202
12 norm of gradients: 0.8900243114239258
12 norm of weights: 6.391208593904086
-----
Iteration Number: 1141
Loss: 78.93672382611804
12 norm of gradients: 0.8895067868696452
12 norm of weights: 6.39100089193571
-----
Iteration Number: 1142
Loss: 78.89805581477167
12 norm of gradients: 0.8889895269095633
12 norm of weights: 6.390793308098938
-----
Iteration Number: 1143
Loss: 78.8594324948732
12 norm of gradients: 0.8884725318619449
12 norm of weights: 6.390585842252006
-----
Iteration Number: 1144

```

Loss: 78.82085381740838
l2 norm of gradients: 0.8879558020491414
l2 norm of weights: 6.390378494253272

Iteration Number: 1145
Loss: 78.78231973315948
l2 norm of gradients: 0.8874393377975082
l2 norm of weights: 6.390171263961211

Iteration Number: 1146
Loss: 78.74383019322946
l2 norm of gradients: 0.8869231394373209
l2 norm of weights: 6.389964151234418

Iteration Number: 1147
Loss: 78.70538514850706
l2 norm of gradients: 0.8864072073026914
l2 norm of weights: 6.389757155931608

Iteration Number: 1148
Loss: 78.66698454995988
l2 norm of gradients: 0.8858915417314842
l2 norm of weights: 6.389550277911617

Iteration Number: 1149
Loss: 78.62862834856827
l2 norm of gradients: 0.8853761430652342
l2 norm of weights: 6.389343517033398

Iteration Number: 1150
Loss: 78.59031649533637
l2 norm of gradients: 0.8848610116490611
l2 norm of weights: 6.389136873156025

Iteration Number: 1151
Loss: 78.55204894135565
l2 norm of gradients: 0.8843461478315873
l2 norm of weights: 6.3889303461386975

Iteration Number: 1152
Loss: 78.51382563748035
l2 norm of gradients: 0.8838315519648537
l2 norm of weights: 6.388723935840728

Iteration Number: 1153
Loss: 78.47564653490099
l2 norm of gradients: 0.8833172244042378
l2 norm of weights: 6.388517642121557

```

-----
Iteration Number: 1154
Loss: 78.43751158450262
l2 norm of gradients: 0.8828031655083673
l2 norm of weights: 6.388311464840741
-----
Iteration Number: 1155
Loss: 78.399420737422
l2 norm of gradients: 0.8822893756390406
l2 norm of weights: 6.3881054038579625
-----
Iteration Number: 1156
Loss: 78.36137394479076
l2 norm of gradients: 0.8817758551611408
l2 norm of weights: 6.387899459033022
-----
Iteration Number: 1157
Loss: 78.3233711575328
l2 norm of gradients: 0.8812626044425536
l2 norm of weights: 6.387693630225846
-----
Iteration Number: 1158
Loss: 78.2854123268101
l2 norm of gradients: 0.880749623854085
l2 norm of weights: 6.387487917296482
-----
Iteration Number: 1159
Loss: 78.2474974037156
l2 norm of gradients: 0.880236913769378
l2 norm of weights: 6.387282320105095
-----
Iteration Number: 1160
Loss: 78.20962633934698
l2 norm of gradients: 0.8797244745648303
l2 norm of weights: 6.387076838511982
-----
Iteration Number: 1161
Loss: 78.17179908482251
l2 norm of gradients: 0.8792123066195117
l2 norm of weights: 6.386871472377556
-----
Iteration Number: 1162
Loss: 78.13401559125924
l2 norm of gradients: 0.8787004103150818
l2 norm of weights: 6.386666221562358
-----
Iteration Number: 1163
Loss: 78.09627580972342

```


l2 norm of gradients: 0.8781887860357096
l2 norm of weights: 6.386461085927049

Iteration Number: 1164
Loss: 78.05857969141934
l2 norm of gradients: 0.8776774341679893
l2 norm of weights: 6.386256065332417

Iteration Number: 1165
Loss: 78.02092718749626
l2 norm of gradients: 0.8771663551008605
l2 norm of weights: 6.386051159639372

Iteration Number: 1166
Loss: 77.98331824911736
l2 norm of gradients: 0.8766555492255271
l2 norm of weights: 6.38584636870895

Iteration Number: 1167
Loss: 77.94575282737594
l2 norm of gradients: 0.8761450169353759
l2 norm of weights: 6.385641692402311

Iteration Number: 1168
Loss: 77.90823087352149
l2 norm of gradients: 0.8756347586258955
l2 norm of weights: 6.3854371305807405

Iteration Number: 1169
Loss: 77.87075233876682
l2 norm of gradients: 0.875124774694597
l2 norm of weights: 6.385232683105649

Iteration Number: 1170
Loss: 77.83331717415746
l2 norm of gradients: 0.8746150655409335
l2 norm of weights: 6.385028349838571

Iteration Number: 1171
Loss: 77.7959253310086
l2 norm of gradients: 0.8741056315662206
l2 norm of weights: 6.384824130641171

Iteration Number: 1172
Loss: 77.75857676044778
l2 norm of gradients: 0.8735964731735567
l2 norm of weights: 6.384620025375235

Iteration Number: 1173
Loss: 77.72127141377425
l2 norm of gradients: 0.8730875907677446
l2 norm of weights: 6.384416033902679

Iteration Number: 1174
Loss: 77.68400924220074
l2 norm of gradients: 0.8725789847552125
l2 norm of weights: 6.3842121560855425

Iteration Number: 1175
Loss: 77.6467901968944
l2 norm of gradients: 0.8720706555439361
l2 norm of weights: 6.384008391785996

Iteration Number: 1176
Loss: 77.60961422908687
l2 norm of gradients: 0.871562603543361
l2 norm of weights: 6.383804740866332

Iteration Number: 1177
Loss: 77.57248129010567
l2 norm of gradients: 0.8710548291643242
l2 norm of weights: 6.3836012031889755

Iteration Number: 1178
Loss: 77.53539133112703
l2 norm of gradients: 0.8705473328189791
l2 norm of weights: 6.383397778616476

Iteration Number: 1179
Loss: 77.49834430348436
l2 norm of gradients: 0.8700401149207173
l2 norm of weights: 6.383194467011512

Iteration Number: 1180
Loss: 77.46134015837569
l2 norm of gradients: 0.869533175884093
l2 norm of weights: 6.38299126823689

Iteration Number: 1181
Loss: 77.42437884707805
l2 norm of gradients: 0.8690265161247482
l2 norm of weights: 6.382788182155546

Iteration Number: 1182
Loss: 77.38746032095312
l2 norm of gradients: 0.8685201360593366

```

12 norm of weights: 6.382585208630545
-----
Iteration Number: 1183
Loss: 77.35058453121727
12 norm of gradients: 0.86801403610545
12 norm of weights: 6.382382347525078
-----
Iteration Number: 1184
Loss: 77.3137514292413
12 norm of gradients: 0.8675082166815428
12 norm of weights: 6.382179598702469
-----
Iteration Number: 1185
Loss: 77.27696096618743
12 norm of gradients: 0.86700267820686
12 norm of weights: 6.38197696202617
-----
Iteration Number: 1186
Loss: 77.24021309351271
12 norm of gradients: 0.866497421101363
12 norm of weights: 6.381774437359763
-----
Iteration Number: 1187
Loss: 77.2035077624671
12 norm of gradients: 0.8659924457856576
12 norm of weights: 6.38157202456696
-----
Iteration Number: 1188
Loss: 77.16684492438496
12 norm of gradients: 0.8654877526809224
12 norm of weights: 6.381369723511606
-----
Iteration Number: 1189
Loss: 77.13022453058912
12 norm of gradients: 0.8649833422088362
12 norm of weights: 6.381167534057672
-----
Iteration Number: 1190
Loss: 77.093646532442
12 norm of gradients: 0.8644792147915085
12 norm of weights: 6.380965456069262
-----
Iteration Number: 1191
Loss: 77.05711088135241
12 norm of gradients: 0.8639753708514087
12 norm of weights: 6.380763489410613
-----
Iteration Number: 1192

```

```
Loss: 77.02061752853916
12 norm of gradients: 0.8634718108112965
12 norm of weights: 6.380561633946092
-----
Iteration Number: 1193
Loss: 76.98416642551899
12 norm of gradients: 0.862968535094153
12 norm of weights: 6.380359889540197
-----
Iteration Number: 1194
Loss: 76.94775752350333
12 norm of gradients: 0.8624655441231115
12 norm of weights: 6.38015825605756
-----
Iteration Number: 1195
Loss: 76.91139077403894
12 norm of gradients: 0.8619628383213911
12 norm of weights: 6.379956733362942
-----
Iteration Number: 1196
Loss: 76.87506612839748
12 norm of gradients: 0.8614604181122287
12 norm of weights: 6.379755321321239
-----
Iteration Number: 1197
Loss: 76.83878353797192
12 norm of gradients: 0.8609582839188128
12 norm of weights: 6.379554019797482
-----
Iteration Number: 1198
Loss: 76.80254295420772
12 norm of gradients: 0.8604564361642172
12 norm of weights: 6.379352828656828
-----
Iteration Number: 1199
Loss: 76.76634432851664
12 norm of gradients: 0.8599548752713362
12 norm of weights: 6.379151747764575
-----
Iteration Number: 1200
Loss: 76.73018761231714
12 norm of gradients: 0.8594536016628203
12 norm of weights: 6.378950776986148
-----
Iteration Number: 1201
Loss: 76.69407275703914
12 norm of gradients: 0.8589526157610121
12 norm of weights: 6.378749916187111
```

```

-----
Iteration Number: 1202
Loss: 76.65799971404446
l2 norm of gradients: 0.8584519179878833
l2 norm of weights: 6.378549165233158
-----
Iteration Number: 1203
Loss: 76.62196843485286
l2 norm of gradients: 0.8579515087649726
l2 norm of weights: 6.3783485239901205
-----
Iteration Number: 1204
Loss: 76.585978870915
l2 norm of gradients: 0.8574513885133226
l2 norm of weights: 6.378147992323962
-----
Iteration Number: 1205
Loss: 76.55003097369645
l2 norm of gradients: 0.8569515576534212
l2 norm of weights: 6.377947570100782
-----
Iteration Number: 1206
Loss: 76.51412469459434
l2 norm of gradients: 0.8564520166051394
l2 norm of weights: 6.3777472571868135
-----
Iteration Number: 1207
Loss: 76.47825998507244
l2 norm of gradients: 0.8559527657876727
l2 norm of weights: 6.377547053448428
-----
Iteration Number: 1208
Loss: 76.44243679677943
l2 norm of gradients: 0.8554538056194815
l2 norm of weights: 6.377346958752126
-----
Iteration Number: 1209
Loss: 76.4066550810602
l2 norm of gradients: 0.8549551365182347
l2 norm of weights: 6.377146972964552
-----
Iteration Number: 1210
Loss: 76.37091478942058
l2 norm of gradients: 0.8544567589007509
l2 norm of weights: 6.37694709595248
-----
Iteration Number: 1211
Loss: 76.33521587343851

```

l2 norm of gradients: 0.8539586731829423
l2 norm of weights: 6.376747327582823

Iteration Number: 1212
Loss: 76.299558284652
l2 norm of gradients: 0.8534608797797597
l2 norm of weights: 6.37654766772263

Iteration Number: 1213
Loss: 76.26394197450807
l2 norm of gradients: 0.8529633791051369
l2 norm of weights: 6.376348116239084

Iteration Number: 1214
Loss: 76.22836689455316
l2 norm of gradients: 0.8524661715719359
l2 norm of weights: 6.376148672999508

Iteration Number: 1215
Loss: 76.19283299640847
l2 norm of gradients: 0.8519692575918958
l2 norm of weights: 6.375949337871362

Iteration Number: 1216
Loss: 76.15734023156955
l2 norm of gradients: 0.8514726375755773
l2 norm of weights: 6.37575011072224

Iteration Number: 1217
Loss: 76.12188855159509
l2 norm of gradients: 0.8509763119323135
l2 norm of weights: 6.3755509914198765

Iteration Number: 1218
Loss: 76.08647790814616
l2 norm of gradients: 0.8504802810701578
l2 norm of weights: 6.375351979832143

Iteration Number: 1219
Loss: 76.05110825273898
l2 norm of gradients: 0.8499845453958337
l2 norm of weights: 6.375153075827046

Iteration Number: 1220
Loss: 76.01577953697138
l2 norm of gradients: 0.8494891053146851
l2 norm of weights: 6.374954279272735

Iteration Number: 1221
Loss: 75.98049171246385
l2 norm of gradients: 0.8489939612306296
l2 norm of weights: 6.374755590037494

Iteration Number: 1222
Loss: 75.9452447309418
l2 norm of gradients: 0.8484991135461091
l2 norm of weights: 6.374557007989747

Iteration Number: 1223
Loss: 75.91003854382214
l2 norm of gradients: 0.8480045626620432
l2 norm of weights: 6.374358532998054

Iteration Number: 1224
Loss: 75.87487310292782
l2 norm of gradients: 0.8475103089777841
l2 norm of weights: 6.374160164931117

Iteration Number: 1225
Loss: 75.83974835984633
l2 norm of gradients: 0.8470163528910709
l2 norm of weights: 6.3739619036577775

Iteration Number: 1226
Loss: 75.804664266241
l2 norm of gradients: 0.8465226947979855
l2 norm of weights: 6.373763749047012

Iteration Number: 1227
Loss: 75.76962077379504
l2 norm of gradients: 0.8460293350929082
l2 norm of weights: 6.37356570096794

Iteration Number: 1228
Loss: 75.73461783418537
l2 norm of gradients: 0.8455362741684777
l2 norm of weights: 6.37336775928982

Iteration Number: 1229
Loss: 75.69965539905888
l2 norm of gradients: 0.8450435124155453
l2 norm of weights: 6.373169923882048

Iteration Number: 1230
Loss: 75.66473342026566
l2 norm of gradients: 0.8445510502231383

```

12 norm of weights: 6.372972194614164
-----
Iteration Number: 1231
Loss: 75.62985184946054
12 norm of gradients: 0.8440588879784168
12 norm of weights: 6.372774571355842
-----
Iteration Number: 1232
Loss: 75.59501063830642
12 norm of gradients: 0.843567026066637
12 norm of weights: 6.372577053976903
-----
Iteration Number: 1233
Loss: 75.56020973870743
12 norm of gradients: 0.8430754648711104
12 norm of weights: 6.372379642347305
-----
Iteration Number: 1234
Loss: 75.52544910227965
12 norm of gradients: 0.8425842047731686
12 norm of weights: 6.3721823363371435
-----
Iteration Number: 1235
Loss: 75.49072868079878
12 norm of gradients: 0.8420932461521256
12 norm of weights: 6.3719851358166615
-----
Iteration Number: 1236
Loss: 75.45604842620199
12 norm of gradients: 0.8416025893852425
12 norm of weights: 6.371788040656238
-----
Iteration Number: 1237
Loss: 75.42140829014875
12 norm of gradients: 0.8411122348476918
12 norm of weights: 6.371591050726395
-----
Iteration Number: 1238
Loss: 75.3868082245348
12 norm of gradients: 0.8406221829125254
12 norm of weights: 6.371394165897794
-----
Iteration Number: 1239
Loss: 75.35224818118468
12 norm of gradients: 0.8401324339506399
12 norm of weights: 6.371197386041239
-----
Iteration Number: 1240

```


Loss: 75.31772811187378
l2 norm of gradients: 0.8396429883307444
l2 norm of weights: 6.371000711027676

Iteration Number: 1241
Loss: 75.28324796846756
l2 norm of gradients: 0.8391538464193301
l2 norm of weights: 6.370804140728192

Iteration Number: 1242
Loss: 75.24880770298113
l2 norm of gradients: 0.8386650085806397
l2 norm of weights: 6.370607675014016

Iteration Number: 1243
Loss: 75.21440726709827
l2 norm of gradients: 0.8381764751766375
l2 norm of weights: 6.370411313756517

Iteration Number: 1244
Loss: 75.18004661285738
l2 norm of gradients: 0.8376882465669803
l2 norm of weights: 6.370215056827209

Iteration Number: 1245
Loss: 75.14572569217789
l2 norm of gradients: 0.8372003231089901
l2 norm of weights: 6.370018904097745

Iteration Number: 1246
Loss: 75.11144445692972
l2 norm of gradients: 0.8367127051576276
l2 norm of weights: 6.369822855439926

Iteration Number: 1247
Loss: 75.07720285909765
l2 norm of gradients: 0.8362253930654653
l2 norm of weights: 6.369626910725687

Iteration Number: 1248
Loss: 75.04300085065098
l2 norm of gradients: 0.8357383871826628
l2 norm of weights: 6.369431069827112

Iteration Number: 1249
Loss: 75.00883838353556
l2 norm of gradients: 0.8352516878569419
l2 norm of weights: 6.369235332616425

```

-----
Iteration Number: 1250
Loss: 74.97471540975843
l2 norm of gradients: 0.8347652954335636
l2 norm of weights: 6.3690396989659925
-----
Iteration Number: 1251
Loss: 74.94063188133161
l2 norm of gradients: 0.8342792102553053
l2 norm of weights: 6.368844168748327
-----
Iteration Number: 1252
Loss: 74.90658775028548
l2 norm of gradients: 0.8337934326624395
l2 norm of weights: 6.368648741836077
-----
Iteration Number: 1253
Loss: 74.8725829687463
l2 norm of gradients: 0.8333079629927123
l2 norm of weights: 6.3684534181020425
-----
Iteration Number: 1254
Loss: 74.83861748870889
l2 norm of gradients: 0.832822801581323
l2 norm of weights: 6.368258197419162
-----
Iteration Number: 1255
Loss: 74.80469126226171
l2 norm of gradients: 0.8323379487609065
l2 norm of weights: 6.368063079660517
-----
Iteration Number: 1256
Loss: 74.77080424144113
l2 norm of gradients: 0.831853404861513
l2 norm of weights: 6.367868064699333
-----
Iteration Number: 1257
Loss: 74.73695637849852
l2 norm of gradients: 0.8313691702105922
l2 norm of weights: 6.367673152408979
-----
Iteration Number: 1258
Loss: 74.70314762551187
l2 norm of gradients: 0.8308852451329755
l2 norm of weights: 6.367478342662968
-----
Iteration Number: 1259
Loss: 74.6693779346408

```

l2 norm of gradients: 0.8304016299508603
l2 norm of weights: 6.367283635334956

Iteration Number: 1260
Loss: 74.63564725805952
l2 norm of gradients: 0.8299183249837957
l2 norm of weights: 6.367089030298744

Iteration Number: 1261
Loss: 74.60195554789219
l2 norm of gradients: 0.8294353305486682
l2 norm of weights: 6.366894527428273

Iteration Number: 1262
Loss: 74.56830275642557
l2 norm of gradients: 0.8289526469596876
l2 norm of weights: 6.366700126597633

Iteration Number: 1263
Loss: 74.53468883590031
l2 norm of gradients: 0.8284702745283757
l2 norm of weights: 6.366505827681052

Iteration Number: 1264
Loss: 74.50111373856389
l2 norm of gradients: 0.8279882135635531
l2 norm of weights: 6.366311630552907

Iteration Number: 1265
Loss: 74.46757741669522
l2 norm of gradients: 0.8275064643713301
l2 norm of weights: 6.366117535087718

Iteration Number: 1266
Loss: 74.43407982251267
l2 norm of gradients: 0.8270250272550954
l2 norm of weights: 6.3659235411601465

Iteration Number: 1267
Loss: 74.40062090841126
l2 norm of gradients: 0.8265439025155074
l2 norm of weights: 6.365729648645

Iteration Number: 1268
Loss: 74.36720062665023
l2 norm of gradients: 0.8260630904504851
l2 norm of weights: 6.365535857417229

Iteration Number: 1269
Loss: 74.33381892973291
l2 norm of gradients: 0.8255825913552014
l2 norm of weights: 6.36534216735193

Iteration Number: 1270
Loss: 74.30047576982193
l2 norm of gradients: 0.8251024055220759
l2 norm of weights: 6.365148578324343

Iteration Number: 1271
Loss: 74.26717109949641
l2 norm of gradients: 0.8246225332407682
l2 norm of weights: 6.36495509020985

Iteration Number: 1272
Loss: 74.23390487109931
l2 norm of gradients: 0.8241429747981732
l2 norm of weights: 6.364761702883982

Iteration Number: 1273
Loss: 74.20067703704832
l2 norm of gradients: 0.8236637304784167
l2 norm of weights: 6.36456841622241

Iteration Number: 1274
Loss: 74.16748754980004
l2 norm of gradients: 0.8231848005628508
l2 norm of weights: 6.364375230100951

Iteration Number: 1275
Loss: 74.13433636191984
l2 norm of gradients: 0.8227061853300518
l2 norm of weights: 6.364182144395566

Iteration Number: 1276
Loss: 74.10122342584677
l2 norm of gradients: 0.822227885055817
l2 norm of weights: 6.3639891589823625

Iteration Number: 1277
Loss: 74.06814869410636
l2 norm of gradients: 0.8217499000131644
l2 norm of weights: 6.363796273737589

Iteration Number: 1278
Loss: 74.035112119268
l2 norm of gradients: 0.8212722304723301

```

12 norm of weights: 6.363603488537641
-----
Iteration Number: 1279
Loss: 74.00211365396979
12 norm of gradients: 0.82079487670077
12 norm of weights: 6.3634108032590575
-----
Iteration Number: 1280
Loss: 73.96915325075189
12 norm of gradients: 0.8203178389631598
12 norm of weights: 6.363218217778522
-----
Iteration Number: 1281
Loss: 73.93623086223573
12 norm of gradients: 0.8198411175213954
12 norm of weights: 6.363025731972863
-----
Iteration Number: 1282
Loss: 73.9033464410809
12 norm of gradients: 0.8193647126345968
12 norm of weights: 6.362833345719053
-----
Iteration Number: 1283
Loss: 73.87049993995113
12 norm of gradients: 0.8188886245591087
12 norm of weights: 6.3626410588942095
-----
Iteration Number: 1284
Loss: 73.83769131155351
12 norm of gradients: 0.8184128535485062
12 norm of weights: 6.362448871375594
-----
Iteration Number: 1285
Loss: 73.8049205086113
12 norm of gradients: 0.8179373998535968
12 norm of weights: 6.362256783040613
-----
Iteration Number: 1286
Loss: 73.77218748387533
12 norm of gradients: 0.8174622637224267
12 norm of weights: 6.362064793766815
-----
Iteration Number: 1287
Loss: 73.73949219005054
12 norm of gradients: 0.8169874454002856
12 norm of weights: 6.361872903431899
-----
Iteration Number: 1288

```

```
Loss: 73.70683458007149
12 norm of gradients: 0.8165129451297128
12 norm of weights: 6.3616811119137004
-----
Iteration Number: 1289
Loss: 73.67421460662905
12 norm of gradients: 0.8160387631505039
12 norm of weights: 6.361489419090207
-----
Iteration Number: 1290
Loss: 73.64163222271826
12 norm of gradients: 0.8155648996997189
12 norm of weights: 6.361297824839545
-----
Iteration Number: 1291
Loss: 73.60908738099592
12 norm of gradients: 0.8150913550116892
12 norm of weights: 6.361106329039988
-----
Iteration Number: 1292
Loss: 73.57658003458567
12 norm of gradients: 0.8146181293180266
12 norm of weights: 6.360914931569953
-----
Iteration Number: 1293
Loss: 73.5441101362733
12 norm of gradients: 0.8141452228476329
12 norm of weights: 6.3607236323080025
-----
Iteration Number: 1294
Loss: 73.5116776390866
12 norm of gradients: 0.8136726358267097
12 norm of weights: 6.3605324311328415
-----
Iteration Number: 1295
Loss: 73.47928249598274
12 norm of gradients: 0.8132003684787675
12 norm of weights: 6.36034132792332
-----
Iteration Number: 1296
Loss: 73.44692465996172
12 norm of gradients: 0.8127284210246396
12 norm of weights: 6.360150322558433
-----
Iteration Number: 1297
Loss: 73.41460408410133
12 norm of gradients: 0.8122567936824905
12 norm of weights: 6.359959414917321
```

Iteration Number: 1298
Loss: 73.38232072139962
l2 norm of gradients: 0.8117854866678305
l2 norm of weights: 6.359768604879266

Iteration Number: 1299
Loss: 73.35007452495168
l2 norm of gradients: 0.8113145001935261
l2 norm of weights: 6.359577892323692

Iteration Number: 1300
Loss: 73.31786544795307
l2 norm of gradients: 0.810843834469816
l2 norm of weights: 6.359387277130174

Iteration Number: 1301
Loss: 73.28569344347939
l2 norm of gradients: 0.8103734897043225
l2 norm of weights: 6.359196759178426

Iteration Number: 1302
Loss: 73.25355846470956
l2 norm of gradients: 0.8099034661020676
l2 norm of weights: 6.359006338348307

Iteration Number: 1303
Loss: 73.22146046485402
l2 norm of gradients: 0.8094337638654863
l2 norm of weights: 6.358816014519819

Iteration Number: 1304
Loss: 73.189399397219
l2 norm of gradients: 0.8089643831944429
l2 norm of weights: 6.3586257875731125

Iteration Number: 1305
Loss: 73.15737521500036
l2 norm of gradients: 0.8084953242862463
l2 norm of weights: 6.358435657388474

Iteration Number: 1306
Loss: 73.12538787149522
l2 norm of gradients: 0.8080265873356673
l2 norm of weights: 6.358245623846342

Iteration Number: 1307
Loss: 73.09343732002898

l2 norm of gradients: 0.8075581725349541
l2 norm of weights: 6.358055686827291

Iteration Number: 1308
Loss: 73.06152351394881
l2 norm of gradients: 0.8070900800738504
l2 norm of weights: 6.357865846212047

Iteration Number: 1309
Loss: 73.02964640664598
l2 norm of gradients: 0.8066223101396126
l2 norm of weights: 6.357676101881473

Iteration Number: 1310
Loss: 72.99780595153659
l2 norm of gradients: 0.8061548629170285
l2 norm of weights: 6.357486453716577

Iteration Number: 1311
Loss: 72.96600210205709
l2 norm of gradients: 0.8056877385884356
l2 norm of weights: 6.357296901598514

Iteration Number: 1312
Loss: 72.93423481161862
l2 norm of gradients: 0.8052209373337398
l2 norm of weights: 6.3571074454085785

Iteration Number: 1313
Loss: 72.90250403386656
l2 norm of gradients: 0.804754459330435
l2 norm of weights: 6.356918085028209

Iteration Number: 1314
Loss: 72.87080972219717
l2 norm of gradients: 0.8042883047536241
l2 norm of weights: 6.356728820338988

Iteration Number: 1315
Loss: 72.8391518303144
l2 norm of gradients: 0.8038224737760368
l2 norm of weights: 6.356539651222641

Iteration Number: 1316
Loss: 72.80753031161191
l2 norm of gradients: 0.8033569665680524
l2 norm of weights: 6.356350577561035

Iteration Number: 1317
Loss: 72.77594511996398
l2 norm of gradients: 0.8028917832977197
l2 norm of weights: 6.356161599236184

Iteration Number: 1318
Loss: 72.74439620881992
l2 norm of gradients: 0.8024269241307783
l2 norm of weights: 6.355972716130241

Iteration Number: 1319
Loss: 72.71288353197599
l2 norm of gradients: 0.801962389230681
l2 norm of weights: 6.3557839281255

Iteration Number: 1320
Loss: 72.68140704309872
l2 norm of gradients: 0.801498178758615
l2 norm of weights: 6.355595235104404

Iteration Number: 1321
Loss: 72.64996669600752
l2 norm of gradients: 0.8010342928735245
l2 norm of weights: 6.355406636949533

Iteration Number: 1322
Loss: 72.61856244441599
l2 norm of gradients: 0.800570731732134
l2 norm of weights: 6.355218133543612

Iteration Number: 1323
Loss: 72.58719424215705
l2 norm of gradients: 0.8001074954889702
l2 norm of weights: 6.3550297247695084

Iteration Number: 1324
Loss: 72.55586204315098
l2 norm of gradients: 0.7996445842963857
l2 norm of weights: 6.354841410510231

Iteration Number: 1325
Loss: 72.52456580121671
l2 norm of gradients: 0.7991819983045831
l2 norm of weights: 6.354653190648931

Iteration Number: 1326
Loss: 72.49330547034283
l2 norm of gradients: 0.7987197376616377

```

12 norm of weights: 6.354465065068902
-----
Iteration Number: 1327
Loss: 72.46208100440427
12 norm of gradients: 0.7982578025135229
12 norm of weights: 6.354277033653578
-----
Iteration Number: 1328
Loss: 72.43089235740406
12 norm of gradients: 0.7977961930041331
12 norm of weights: 6.354089096286538
-----
Iteration Number: 1329
Loss: 72.39973948335266
12 norm of gradients: 0.7973349092753099
12 norm of weights: 6.3539012528515
-----
Iteration Number: 1330
Loss: 72.36862233630742
12 norm of gradients: 0.7968739514668656
12 norm of weights: 6.353713503232325
-----
Iteration Number: 1331
Loss: 72.33754087034728
12 norm of gradients: 0.7964133197166087
12 norm of weights: 6.353525847313015
-----
Iteration Number: 1332
Loss: 72.30649503959603
12 norm of gradients: 0.795953014160369
12 norm of weights: 6.3533382849777125
-----
Iteration Number: 1333
Loss: 72.27548479821245
12 norm of gradients: 0.7954930349320238
12 norm of weights: 6.353150816110705
-----
Iteration Number: 1334
Loss: 72.24451010045034
12 norm of gradients: 0.7950333821635223
12 norm of weights: 6.352963440596415
-----
Iteration Number: 1335
Loss: 72.21357090039365
12 norm of gradients: 0.7945740559849126
12 norm of weights: 6.352776158319411
-----
Iteration Number: 1336

```

```

Loss: 72.18266715230851
12 norm of gradients: 0.794115056524367
12 norm of weights: 6.352588969164403
-----
Iteration Number: 1337
Loss: 72.15179881060789
12 norm of gradients: 0.7936563839082079
12 norm of weights: 6.3524018730162375
-----
Iteration Number: 1338
Loss: 72.12096582949572
12 norm of gradients: 0.7931980382609354
12 norm of weights: 6.352214869759905
-----
Iteration Number: 1339
Loss: 72.09016816342236
12 norm of gradients: 0.7927400197052525
12 norm of weights: 6.352027959280535
-----
Iteration Number: 1340
Loss: 72.05940576672798
12 norm of gradients: 0.7922823283620917
12 norm of weights: 6.351841141463399
-----
Iteration Number: 1341
Loss: 72.02867859380233
12 norm of gradients: 0.7918249643506424
12 norm of weights: 6.351654416193907
-----
Iteration Number: 1342
Loss: 71.99798659917991
12 norm of gradients: 0.7913679277883775
12 norm of weights: 6.35146778335761
-----
Iteration Number: 1343
Loss: 71.9673297373579
12 norm of gradients: 0.7909112187910801
12 norm of weights: 6.3512812428402015
-----
Iteration Number: 1344
Loss: 71.93670796281292
12 norm of gradients: 0.7904548374728705
12 norm of weights: 6.35109479452751
-----
Iteration Number: 1345
Loss: 71.90612123010897
12 norm of gradients: 0.7899987839462332
12 norm of weights: 6.350908438305509

```

Iteration Number: 1346
Loss: 71.87556949392123
l2 norm of gradients: 0.7895430583220446
l2 norm of weights: 6.350722174060309

Iteration Number: 1347
Loss: 71.84505270883363
l2 norm of gradients: 0.7890876607095996
l2 norm of weights: 6.350536001678157

Iteration Number: 1348
Loss: 71.8145708295175
l2 norm of gradients: 0.7886325912166389
l2 norm of weights: 6.3503499210454475

Iteration Number: 1349
Loss: 71.78412381075084
l2 norm of gradients: 0.7881778499493768
l2 norm of weights: 6.350163932048707

Iteration Number: 1350
Loss: 71.75371160722673
l2 norm of gradients: 0.7877234370125279
l2 norm of weights: 6.349978034574605

Iteration Number: 1351
Loss: 71.72333417365982
l2 norm of gradients: 0.787269352509335
l2 norm of weights: 6.349792228509948

Iteration Number: 1352
Loss: 71.6929914650658
l2 norm of gradients: 0.7868155965415959
l2 norm of weights: 6.349606513741683

Iteration Number: 1353
Loss: 71.66268343611827
l2 norm of gradients: 0.7863621692096916
l2 norm of weights: 6.349420890156894

Iteration Number: 1354
Loss: 71.63241004176291
l2 norm of gradients: 0.7859090706126127
l2 norm of weights: 6.349235357642807

Iteration Number: 1355
Loss: 71.60217123694444

l2 norm of gradients: 0.7854563008479878
l2 norm of weights: 6.349049916086782

Iteration Number: 1356
Loss: 71.57196697662219
l2 norm of gradients: 0.7850038600121095
l2 norm of weights: 6.348864565376322

Iteration Number: 1357
Loss: 71.5417972158095
l2 norm of gradients: 0.784551748199963
l2 norm of weights: 6.348679305399063

Iteration Number: 1358
Loss: 71.51166190953712
l2 norm of gradients: 0.7840999655052527
l2 norm of weights: 6.348494136042785

Iteration Number: 1359
Loss: 71.48156101282252
l2 norm of gradients: 0.7836485120204298
l2 norm of weights: 6.348309057195401

Iteration Number: 1360
Loss: 71.45149448094806
l2 norm of gradients: 0.783197387836719
l2 norm of weights: 6.348124068744965

Iteration Number: 1361
Loss: 71.42146226885859
l2 norm of gradients: 0.7827465930441453
l2 norm of weights: 6.347939170579666

Iteration Number: 1362
Loss: 71.39146433178603
l2 norm of gradients: 0.7822961277315622
l2 norm of weights: 6.3477543625878345

Iteration Number: 1363
Loss: 71.36150062508399
l2 norm of gradients: 0.781845991986678
l2 norm of weights: 6.347569644657934

Iteration Number: 1364
Loss: 71.33157110381076
l2 norm of gradients: 0.7813961858960823
l2 norm of weights: 6.347385016678568

Iteration Number: 1365
Loss: 71.30167572347291
l2 norm of gradients: 0.7809467095452729
l2 norm of weights: 6.347200478538476

Iteration Number: 1366
Loss: 71.27181443927601
l2 norm of gradients: 0.7804975630186829
l2 norm of weights: 6.347016030126538

Iteration Number: 1367
Loss: 71.24198720658949
l2 norm of gradients: 0.780048746399707
l2 norm of weights: 6.3468316713317625

Iteration Number: 1368
Loss: 71.2121939808301
l2 norm of gradients: 0.7796002597707272
l2 norm of weights: 6.346647402043304

Iteration Number: 1369
Loss: 71.18243471750755
l2 norm of gradients: 0.7791521032131405
l2 norm of weights: 6.346463222150448

Iteration Number: 1370
Loss: 71.15270937204646
l2 norm of gradients: 0.778704276807384
l2 norm of weights: 6.346279131542619

Iteration Number: 1371
Loss: 71.12301790003659
l2 norm of gradients: 0.7782567806329609
l2 norm of weights: 6.346095130109377

Iteration Number: 1372
Loss: 71.09336025690814
l2 norm of gradients: 0.7778096147684667
l2 norm of weights: 6.345911217740416

Iteration Number: 1373
Loss: 71.063736398391
l2 norm of gradients: 0.7773627792916148
l2 norm of weights: 6.345727394325571

Iteration Number: 1374
Loss: 71.03414628005706
l2 norm of gradients: 0.7769162742792625

```

12 norm of weights: 6.345543659754808
-----
Iteration Number: 1375
Loss: 71.00458985758611
12 norm of gradients: 0.7764700998074348
12 norm of weights: 6.345360013918232
-----
Iteration Number: 1376
Loss: 70.97506708668189
12 norm of gradients: 0.7760242559513517
12 norm of weights: 6.345176456706081
-----
Iteration Number: 1377
Loss: 70.94557792303368
12 norm of gradients: 0.7755787427854515
12 norm of weights: 6.344992988008731
-----
Iteration Number: 1378
Loss: 70.91612232256279
12 norm of gradients: 0.7751335603834165
12 norm of weights: 6.344809607716692
-----
Iteration Number: 1379
Loss: 70.8867002409596
12 norm of gradients: 0.7746887088181973
12 norm of weights: 6.344626315720609
-----
Iteration Number: 1380
Loss: 70.85731163422072
12 norm of gradients: 0.7742441881620368
12 norm of weights: 6.3444431119112625
-----
Iteration Number: 1381
Loss: 70.8279564581529
12 norm of gradients: 0.7737999984864958
12 norm of weights: 6.344259996179568
-----
Iteration Number: 1382
Loss: 70.79863466862678
12 norm of gradients: 0.7733561398624748
12 norm of weights: 6.344076968416576
-----
Iteration Number: 1383
Loss: 70.76934622179176
12 norm of gradients: 0.7729126123602398
12 norm of weights: 6.34389402851347
-----
Iteration Number: 1384

```

Loss: 70.74009107352218
l2 norm of gradients: 0.7724694160494446
l2 norm of weights: 6.343711176361571

Iteration Number: 1385
Loss: 70.71086917987313
l2 norm of gradients: 0.7720265509991543
l2 norm of weights: 6.343528411852332

Iteration Number: 1386
Loss: 70.68168049701046
l2 norm of gradients: 0.7715840172778691
l2 norm of weights: 6.3433457348773405

Iteration Number: 1387
Loss: 70.65252498105906
l2 norm of gradients: 0.7711418149535455
l2 norm of weights: 6.343163145328319

Iteration Number: 1388
Loss: 70.62340258805283
l2 norm of gradients: 0.7706999440936207
l2 norm of weights: 6.342980643097124

Iteration Number: 1389
Loss: 70.59431327440065
l2 norm of gradients: 0.7702584047650338
l2 norm of weights: 6.342798228075743

Iteration Number: 1390
Loss: 70.56525699615051
l2 norm of gradients: 0.769817197034248
l2 norm of weights: 6.342615900156303

Iteration Number: 1391
Loss: 70.53623370970806
l2 norm of gradients: 0.7693763209672726
l2 norm of weights: 6.342433659231058

Iteration Number: 1392
Loss: 70.50724337133639
l2 norm of gradients: 0.768935776629684
l2 norm of weights: 6.342251505192401

Iteration Number: 1393
Loss: 70.47828593738961
l2 norm of gradients: 0.7684955640866479
l2 norm of weights: 6.342069437932855


```
-----  
Iteration Number: 1394  
Loss: 70.449361364266  
l2 norm of gradients: 0.7680556834029392  
l2 norm of weights: 6.341887457345076  
-----  
Iteration Number: 1395  
Loss: 70.4204696084718  
l2 norm of gradients: 0.7676161346429636  
l2 norm of weights: 6.341705563321856  
-----  
Iteration Number: 1396  
Loss: 70.39161062634211  
l2 norm of gradients: 0.7671769178707772  
l2 norm of weights: 6.341523755756118  
-----  
Iteration Number: 1397  
Loss: 70.36278437446148  
l2 norm of gradients: 0.7667380331501068  
l2 norm of weights: 6.341342034540915  
-----  
Iteration Number: 1398  
Loss: 70.3339908093667  
l2 norm of gradients: 0.7662994805443705  
l2 norm of weights: 6.341160399569438  
-----  
Iteration Number: 1399  
Loss: 70.30522988765752  
l2 norm of gradients: 0.765861260116696  
l2 norm of weights: 6.340978850735008  
-----  
Iteration Number: 1400  
Loss: 70.27650156596147  
l2 norm of gradients: 0.7654233719299406  
l2 norm of weights: 6.340797387931077  
-----  
Iteration Number: 1401  
Loss: 70.24780580096491  
l2 norm of gradients: 0.7649858160467096  
l2 norm of weights: 6.340616011051233  
-----  
Iteration Number: 1402  
Loss: 70.21914254924533  
l2 norm of gradients: 0.7645485925293756  
l2 norm of weights: 6.340434719989191  
-----  
Iteration Number: 1403  
Loss: 70.19051176764818
```

l2 norm of gradients: 0.7641117014400964
l2 norm of weights: 6.340253514638802

Iteration Number: 1404
Loss: 70.16191341287214
l2 norm of gradients: 0.7636751428408324
l2 norm of weights: 6.34007239489405

Iteration Number: 1405
Loss: 70.13334744181681
l2 norm of gradients: 0.7632389167933652
l2 norm of weights: 6.3398913606490455

Iteration Number: 1406
Loss: 70.10481381132054
l2 norm of gradients: 0.7628030233593145
l2 norm of weights: 6.339710411798036

Iteration Number: 1407
Loss: 70.07631247822628
l2 norm of gradients: 0.7623674626001549
l2 norm of weights: 6.339529548235396

Iteration Number: 1408
Loss: 70.04784339945432
l2 norm of gradients: 0.7619322345772331
l2 norm of weights: 6.339348769855636

Iteration Number: 1409
Loss: 70.01940653204569
l2 norm of gradients: 0.7614973393517839
l2 norm of weights: 6.3391680765533955

Iteration Number: 1410
Loss: 69.99100183294428
l2 norm of gradients: 0.7610627769849463
l2 norm of weights: 6.338987468223443

Iteration Number: 1411
Loss: 69.96262925925197
l2 norm of gradients: 0.7606285475377791
l2 norm of weights: 6.338806944760682

Iteration Number: 1412
Loss: 69.93428876793129
l2 norm of gradients: 0.7601946510712765
l2 norm of weights: 6.338626506060146

Iteration Number: 1413
Loss: 69.9059803162328
l2 norm of gradients: 0.759761087646383
l2 norm of weights: 6.338446152016998

Iteration Number: 1414
Loss: 69.87770386125231
l2 norm of gradients: 0.7593278573240076
l2 norm of weights: 6.338265882526532

Iteration Number: 1415
Loss: 69.8494593601822
l2 norm of gradients: 0.7588949601650395
l2 norm of weights: 6.338085697484173

Iteration Number: 1416
Loss: 69.82124677019421
l2 norm of gradients: 0.7584623962303605
l2 norm of weights: 6.337905596785477

Iteration Number: 1417
Loss: 69.7930660487002
l2 norm of gradients: 0.7580301655808598
l2 norm of weights: 6.33772558032613

Iteration Number: 1418
Loss: 69.76491715293434
l2 norm of gradients: 0.7575982682774465
l2 norm of weights: 6.337545648001948

Iteration Number: 1419
Loss: 69.73680004019948
l2 norm of gradients: 0.7571667043810631
l2 norm of weights: 6.337365799708877

Iteration Number: 1420
Loss: 69.70871466793406
l2 norm of gradients: 0.7567354739526987
l2 norm of weights: 6.337186035342994

Iteration Number: 1421
Loss: 69.68066099359612
l2 norm of gradients: 0.7563045770533995
l2 norm of weights: 6.337006354800507

Iteration Number: 1422
Loss: 69.65263897458959
l2 norm of gradients: 0.7558740137442822

```

12 norm of weights: 6.336826757977748
-----
Iteration Number: 1423
Loss: 69.62464856838479
12 norm of gradients: 0.7554437840865454
12 norm of weights: 6.3366472447711875
-----
Iteration Number: 1424
Loss: 69.59668973257804
12 norm of gradients: 0.7550138881414806
12 norm of weights: 6.3364678150774205
-----
Iteration Number: 1425
Loss: 69.56876242475161
12 norm of gradients: 0.7545843259704829
12 norm of weights: 6.336288468793171
-----
Iteration Number: 1426
Loss: 69.54086660245642
12 norm of gradients: 0.7541550976350617
12 norm of weights: 6.336109205815295
-----
Iteration Number: 1427
Loss: 69.51300222333938
12 norm of gradients: 0.7537262031968509
12 norm of weights: 6.335930026040775
-----
Iteration Number: 1428
Loss: 69.48516924515953
12 norm of gradients: 0.7532976427176189
12 norm of weights: 6.335750929366726
-----
Iteration Number: 1429
Loss: 69.45736762558222
12 norm of gradients: 0.7528694162592778
12 norm of weights: 6.335571915690391
-----
Iteration Number: 1430
Loss: 69.42959732236298
12 norm of gradients: 0.7524415238838925
12 norm of weights: 6.335392984909141
-----
Iteration Number: 1431
Loss: 69.40185829331428
12 norm of gradients: 0.7520139656536899
12 norm of weights: 6.335214136920476
-----
Iteration Number: 1432

```

Loss: 69.37415049633694
l2 norm of gradients: 0.7515867416310666
l2 norm of weights: 6.335035371622027

Iteration Number: 1433
Loss: 69.34647388924239
l2 norm of gradients: 0.7511598518785977
l2 norm of weights: 6.334856688911553

Iteration Number: 1434
Loss: 69.31882842988534
l2 norm of gradients: 0.750733296459044
l2 norm of weights: 6.334678088686939

Iteration Number: 1435
Loss: 69.29121407634334
l2 norm of gradients: 0.7503070754353593
l2 norm of weights: 6.3344995708462015

Iteration Number: 1436
Loss: 69.26363078654215
l2 norm of gradients: 0.7498811888706983
l2 norm of weights: 6.334321135287485

Iteration Number: 1437
Loss: 69.23607851845533
l2 norm of gradients: 0.7494556368284225
l2 norm of weights: 6.334142781909063

Iteration Number: 1438
Loss: 69.20855723025443
l2 norm of gradients: 0.7490304193721063
l2 norm of weights: 6.333964510609335

Iteration Number: 1439
Loss: 69.18106687993097
l2 norm of gradients: 0.7486055365655433
l2 norm of weights: 6.333786321286833

Iteration Number: 1440
Loss: 69.15360742571974
l2 norm of gradients: 0.7481809884727526
l2 norm of weights: 6.333608213840212

Iteration Number: 1441
Loss: 69.12617882566461
l2 norm of gradients: 0.747756775157983
l2 norm of weights: 6.333430188168259

```

-----
Iteration Number: 1442
Loss: 69.09878103813247
l2 norm of gradients: 0.7473328966857185
l2 norm of weights: 6.333252244169888
-----
Iteration Number: 1443
Loss: 69.07141402125734
l2 norm of gradients: 0.7469093531206827
l2 norm of weights: 6.33307438174414
-----
Iteration Number: 1444
Loss: 69.04407773332109
l2 norm of gradients: 0.7464861445278436
l2 norm of weights: 6.3328966007901855
-----
Iteration Number: 1445
Loss: 69.01677213271367
l2 norm of gradients: 0.7460632709724163
l2 norm of weights: 6.332718901207321
-----
Iteration Number: 1446
Loss: 68.98949717773837
l2 norm of gradients: 0.7456407325198681
l2 norm of weights: 6.332541282894972
-----
Iteration Number: 1447
Loss: 68.96225282678115
l2 norm of gradients: 0.7452185292359208
l2 norm of weights: 6.332363745752693
-----
Iteration Number: 1448
Loss: 68.93503903835405
l2 norm of gradients: 0.7447966611865536
l2 norm of weights: 6.332186289680163
-----
Iteration Number: 1449
Loss: 68.90785577085973
l2 norm of gradients: 0.7443751284380062
l2 norm of weights: 6.33200891457719
-----
Iteration Number: 1450
Loss: 68.8807029828106
l2 norm of gradients: 0.7439539310567808
l2 norm of weights: 6.331831620343709
-----
Iteration Number: 1451
Loss: 68.85358063281228

```

l2 norm of gradients: 0.7435330691096431
l2 norm of weights: 6.331654406879784

Iteration Number: 1452
Loss: 68.82648867932944
l2 norm of gradients: 0.7431125426636266
l2 norm of weights: 6.331477274085605

Iteration Number: 1453
Loss: 68.79942708109549
l2 norm of gradients: 0.74269235178603
l2 norm of weights: 6.331300221861489

Iteration Number: 1454
Loss: 68.77239579665347
l2 norm of gradients: 0.7422724965444216
l2 norm of weights: 6.33112325010788

Iteration Number: 1455
Loss: 68.74539478481476
l2 norm of gradients: 0.7418529770066379
l2 norm of weights: 6.330946358725352

Iteration Number: 1456
Loss: 68.71842400418885
l2 norm of gradients: 0.7414337932407847
l2 norm of weights: 6.3307695476146

Iteration Number: 1457
Loss: 68.69148341359958
l2 norm of gradients: 0.7410149453152367
l2 norm of weights: 6.330592816676455

Iteration Number: 1458
Loss: 68.66457297188157
l2 norm of gradients: 0.7405964332986378
l2 norm of weights: 6.330416165811864

Iteration Number: 1459
Loss: 68.63769263775056
l2 norm of gradients: 0.7401782572598995
l2 norm of weights: 6.3302395949219115

Iteration Number: 1460
Loss: 68.61084237017548
l2 norm of gradients: 0.7397604172682012
l2 norm of weights: 6.330063103907802

Iteration Number: 1461
Loss: 68.58402212806902
l2 norm of gradients: 0.7393429133929877
l2 norm of weights: 6.3298866926708675

Iteration Number: 1462
Loss: 68.55723187031334
l2 norm of gradients: 0.7389257457039685
l2 norm of weights: 6.329710361112571

Iteration Number: 1463
Loss: 68.5304715559663
l2 norm of gradients: 0.7385089142711161
l2 norm of weights: 6.329534109134496

Iteration Number: 1464
Loss: 68.50374114397364
l2 norm of gradients: 0.7380924191646624
l2 norm of weights: 6.329357936638359

Iteration Number: 1465
Loss: 68.47704059346285
l2 norm of gradients: 0.7376762604550982
l2 norm of weights: 6.329181843525997

Iteration Number: 1466
Loss: 68.45036986346601
l2 norm of gradients: 0.7372604382131694
l2 norm of weights: 6.329005829699377

Iteration Number: 1467
Loss: 68.42372891310558
l2 norm of gradients: 0.7368449525098741
l2 norm of weights: 6.3288298950605935

Iteration Number: 1468
Loss: 68.39711770155367
l2 norm of gradients: 0.7364298034164592
l2 norm of weights: 6.328654039511864

Iteration Number: 1469
Loss: 68.37053618800842
l2 norm of gradients: 0.7360149910044175
l2 norm of weights: 6.328478262955535

Iteration Number: 1470
Loss: 68.34398433177199
l2 norm of gradients: 0.7356005153454828


```

12 norm of weights: 6.328302565294079
-----
Iteration Number: 1471
Loss: 68.31746209199677
12 norm of gradients: 0.7351863765116264
12 norm of weights: 6.328126946430092
-----
Iteration Number: 1472
Loss: 68.29096942810952
12 norm of gradients: 0.734772574575053
12 norm of weights: 6.327951406266302
-----
Iteration Number: 1473
Loss: 68.26450629933655
12 norm of gradients: 0.7343591096081951
12 norm of weights: 6.327775944705556
-----
Iteration Number: 1474
Loss: 68.23807266512472
12 norm of gradients: 0.733945981683709
12 norm of weights: 6.327600561650834
-----
Iteration Number: 1475
Loss: 68.21166848488114
12 norm of gradients: 0.7335331908744688
12 norm of weights: 6.327425257005237
-----
Iteration Number: 1476
Loss: 68.18529371810398
12 norm of gradients: 0.7331207372535614
12 norm of weights: 6.327250030671995
-----
Iteration Number: 1477
Loss: 68.15894832420908
12 norm of gradients: 0.732708620894281
12 norm of weights: 6.327074882554464
-----
Iteration Number: 1478
Loss: 68.1326322626945
12 norm of gradients: 0.7322968418701222
12 norm of weights: 6.326899812556123
-----
Iteration Number: 1479
Loss: 68.10634549318377
12 norm of gradients: 0.731885400254775
12 norm of weights: 6.326724820580581
-----
Iteration Number: 1480

```

```

Loss: 68.08008797520797
l2 norm of gradients: 0.7314742961221172
l2 norm of weights: 6.326549906531571
-----
Iteration Number: 1481
Loss: 68.05385966840439
l2 norm of gradients: 0.7310635295462087
l2 norm of weights: 6.326375070312951
-----
Iteration Number: 1482
Loss: 68.0276605324959
l2 norm of gradients: 0.7306531006012836
l2 norm of weights: 6.326200311828706
-----
Iteration Number: 1483
Loss: 68.00149052719537
l2 norm of gradients: 0.7302430093617441
l2 norm of weights: 6.326025630982948
-----
Iteration Number: 1484
Loss: 67.97534961217154
l2 norm of gradients: 0.729833255902152
l2 norm of weights: 6.325851027679912
-----
Iteration Number: 1485
Loss: 67.9492377471952
l2 norm of gradients: 0.7294238402972216
l2 norm of weights: 6.3256765018239625
-----
Iteration Number: 1486
Loss: 67.92315489208977
l2 norm of gradients: 0.7290147626218116
l2 norm of weights: 6.325502053319585
-----
Iteration Number: 1487
Loss: 67.89710100668542
l2 norm of gradients: 0.7286060229509181
l2 norm of weights: 6.325327682071397
-----
Iteration Number: 1488
Loss: 67.87107605087289
l2 norm of gradients: 0.7281976213596645
l2 norm of weights: 6.325153387984134
-----
Iteration Number: 1489
Loss: 67.84507998456354
l2 norm of gradients: 0.7277895579232943
l2 norm of weights: 6.324979170962665

```

```

-----
Iteration Number: 1490
Loss: 67.8191127677674
l2 norm of gradients: 0.7273818327171616
l2 norm of weights: 6.324805030911979
-----
Iteration Number: 1491
Loss: 67.7931743603573
l2 norm of gradients: 0.7269744458167233
l2 norm of weights: 6.3246309677371935
-----
Iteration Number: 1492
Loss: 67.76726472240888
l2 norm of gradients: 0.7265673972975288
l2 norm of weights: 6.32445698134355
-----
Iteration Number: 1493
Loss: 67.74138381392036
l2 norm of gradients: 0.7261606872352113
l2 norm of weights: 6.324283071636417
-----
Iteration Number: 1494
Loss: 67.715531595135
l2 norm of gradients: 0.7257543157054782
l2 norm of weights: 6.324109238521288
-----
Iteration Number: 1495
Loss: 67.68970802599894
l2 norm of gradients: 0.725348282784101
l2 norm of weights: 6.323935481903781
-----
Iteration Number: 1496
Loss: 67.66391306677119
l2 norm of gradients: 0.7249425885469061
l2 norm of weights: 6.323761801689643
-----
Iteration Number: 1497
Loss: 67.63814667757771
l2 norm of gradients: 0.7245372330697645
l2 norm of weights: 6.3235881977847415
-----
Iteration Number: 1498
Loss: 67.61240881865905
l2 norm of gradients: 0.7241322164285812
l2 norm of weights: 6.323414670095074
-----
Iteration Number: 1499
Loss: 67.58669945040593

```

```

12 norm of gradients: 0.723727538699285
12 norm of weights: 6.323241218526761
-----
Iteration Number: 1500
Loss: 67.56101853293768
12 norm of gradients: 0.7233231999578185
12 norm of weights: 6.323067842986049
-----
Iteration Number: 1501
Loss: 67.53536602671375
12 norm of gradients: 0.7229192002801262
12 norm of weights: 6.322894543379311
-----
Iteration Number: 1502
Loss: 67.50974189204466
12 norm of gradients: 0.7225155397421449
12 norm of weights: 6.322721319613044
-----
Iteration Number: 1503
Loss: 67.48414608934252
12 norm of gradients: 0.7221122184197916
12 norm of weights: 6.32254817159387
-----
Iteration Number: 1504
Loss: 67.45857857904788
12 norm of gradients: 0.7217092363889521
12 norm of weights: 6.32237509922854
-----
Iteration Number: 1505
Loss: 67.43303932170014
12 norm of gradients: 0.7213065937254715
12 norm of weights: 6.322202102423926
-----
Iteration Number: 1506
Loss: 67.40752827769933
12 norm of gradients: 0.7209042905051402
12 norm of weights: 6.322029181087029
-----
Iteration Number: 1507
Loss: 67.38204540764394
12 norm of gradients: 0.7205023268036836
12 norm of weights: 6.321856335124971
-----
Iteration Number: 1508
Loss: 67.35659067218212
12 norm of gradients: 0.7201007026967507
12 norm of weights: 6.321683564445004
-----

```

Iteration Number: 1509
Loss: 67.33116403180706
l2 norm of gradients: 0.7196994182599011
l2 norm of weights: 6.321510868954504

Iteration Number: 1510
Loss: 67.30576544718187
l2 norm of gradients: 0.7192984735685937
l2 norm of weights: 6.321338248560971

Iteration Number: 1511
Loss: 67.28039487913114
l2 norm of gradients: 0.7188978686981742
l2 norm of weights: 6.321165703172031

Iteration Number: 1512
Loss: 67.25505228821434
l2 norm of gradients: 0.7184976037238624
l2 norm of weights: 6.320993232695438

Iteration Number: 1513
Loss: 67.22973763526447
l2 norm of gradients: 0.718097678720741
l2 norm of weights: 6.320820837039065

Iteration Number: 1514
Loss: 67.20445088102272
l2 norm of gradients: 0.7176980937637418
l2 norm of weights: 6.320648516110918

Iteration Number: 1515
Loss: 67.17919198637328
l2 norm of gradients: 0.7172988489276336
l2 norm of weights: 6.320476269819125

Iteration Number: 1516
Loss: 67.15396091211517
l2 norm of gradients: 0.7168999442870096
l2 norm of weights: 6.320304098071936

Iteration Number: 1517
Loss: 67.12875761921649
l2 norm of gradients: 0.7165013799162742
l2 norm of weights: 6.320132000777731

Iteration Number: 1518
Loss: 67.10358206847738
l2 norm of gradients: 0.7161031558896299

```

12 norm of weights: 6.319959977845013
-----
Iteration Number: 1519
Loss: 67.07843422099302
12 norm of gradients: 0.7157052722810651
12 norm of weights: 6.319788029182414
-----
Iteration Number: 1520
Loss: 67.0533140377571
12 norm of gradients: 0.7153077291643397
12 norm of weights: 6.319616154698686
-----
Iteration Number: 1521
Loss: 67.02822147969071
12 norm of gradients: 0.7149105266129735
12 norm of weights: 6.319444354302709
-----
Iteration Number: 1522
Loss: 67.00315650792611
12 norm of gradients: 0.7145136647002308
12 norm of weights: 6.319272627903488
-----
Iteration Number: 1523
Loss: 66.9781190836283
12 norm of gradients: 0.7141171434991094
12 norm of weights: 6.319100975410156
-----
Iteration Number: 1524
Loss: 66.95310916781679
12 norm of gradients: 0.7137209630823245
12 norm of weights: 6.318929396731966
-----
Iteration Number: 1525
Loss: 66.92812672172643
12 norm of gradients: 0.7133251235222976
12 norm of weights: 6.318757891778301
-----
Iteration Number: 1526
Loss: 66.90317170656742
12 norm of gradients: 0.7129296248911415
12 norm of weights: 6.318586460458666
-----
Iteration Number: 1527
Loss: 66.87824408358793
12 norm of gradients: 0.712534467260647
12 norm of weights: 6.318415102682694
-----
Iteration Number: 1528

```

```
Loss: 66.85334381403557
l2 norm of gradients: 0.7121396507022686
l2 norm of weights: 6.318243818360142
-----
Iteration Number: 1529
Loss: 66.82847085922275
l2 norm of gradients: 0.711745175287112
l2 norm of weights: 6.318072607400893
-----
Iteration Number: 1530
Loss: 66.8036251805319
l2 norm of gradients: 0.7113510410859194
l2 norm of weights: 6.317901469714954
-----
Iteration Number: 1531
Loss: 66.77880673929067
l2 norm of gradients: 0.7109572481690548
l2 norm of weights: 6.31773040521246
-----
Iteration Number: 1532
Loss: 66.75401549690928
l2 norm of gradients: 0.7105637966064917
l2 norm of weights: 6.317559413803667
-----
Iteration Number: 1533
Loss: 66.72925141490623
l2 norm of gradients: 0.7101706864677982
l2 norm of weights: 6.31738849539896
-----
Iteration Number: 1534
Loss: 66.70451445470908
l2 norm of gradients: 0.7097779178221231
l2 norm of weights: 6.317217649908851
-----
Iteration Number: 1535
Loss: 66.67980457790173
l2 norm of gradients: 0.7093854907381816
l2 norm of weights: 6.31704687724397
-----
Iteration Number: 1536
Loss: 66.65512174592098
l2 norm of gradients: 0.7089934052842416
l2 norm of weights: 6.316876177315081
-----
Iteration Number: 1537
Loss: 66.63046592048109
l2 norm of gradients: 0.7086016615281093
l2 norm of weights: 6.316705550033069
```

Iteration Number: 1538
Loss: 66.60583706314625
l2 norm of gradients: 0.7082102595371154
l2 norm of weights: 6.316534995308943

Iteration Number: 1539
Loss: 66.58123513558023
l2 norm of gradients: 0.7078191993781002
l2 norm of weights: 6.31636451305384

Iteration Number: 1540
Loss: 66.55666009948527
l2 norm of gradients: 0.7074284811174005
l2 norm of weights: 6.316194103179023

Iteration Number: 1541
Loss: 66.53211191653212
l2 norm of gradients: 0.7070381048208342
l2 norm of weights: 6.316023765595877

Iteration Number: 1542
Loss: 66.50759054859803
l2 norm of gradients: 0.7066480705536866
l2 norm of weights: 6.315853500215915

Iteration Number: 1543
Loss: 66.48309595738085
l2 norm of gradients: 0.7062583783806967
l2 norm of weights: 6.315683306950776

Iteration Number: 1544
Loss: 66.45862810478954
l2 norm of gradients: 0.7058690283660417
l2 norm of weights: 6.315513185712222

Iteration Number: 1545
Loss: 66.43418695258974
l2 norm of gradients: 0.7054800205733244
l2 norm of weights: 6.31534313641214

Iteration Number: 1546
Loss: 66.40977246282972
l2 norm of gradients: 0.705091355065557
l2 norm of weights: 6.315173158962548

Iteration Number: 1547
Loss: 66.38538459727557

l2 norm of gradients: 0.7047030319051486
l2 norm of weights: 6.315003253275583

Iteration Number: 1548
Loss: 66.36102331802273
l2 norm of gradients: 0.7043150511538901
l2 norm of weights: 6.31483341926351

Iteration Number: 1549
Loss: 66.33668858708032
l2 norm of gradients: 0.7039274128729399
l2 norm of weights: 6.31466365683872

Iteration Number: 1550
Loss: 66.3123803663752
l2 norm of gradients: 0.7035401171228104
l2 norm of weights: 6.314493965913728

Iteration Number: 1551
Loss: 66.28809861811024
l2 norm of gradients: 0.7031531639633527
l2 norm of weights: 6.3143243464011745

Iteration Number: 1552
Loss: 66.26384330439572
l2 norm of gradients: 0.7027665534537433
l2 norm of weights: 6.314154798213829

Iteration Number: 1553
Loss: 66.23961438727352
l2 norm of gradients: 0.7023802856524696
l2 norm of weights: 6.31398532126458

Iteration Number: 1554
Loss: 66.21541182898193
l2 norm of gradients: 0.7019943606173161
l2 norm of weights: 6.313815915466448

Iteration Number: 1555
Loss: 66.19123559180285
l2 norm of gradients: 0.7016087784053494
l2 norm of weights: 6.3136465807325735

Iteration Number: 1556
Loss: 66.16708563787162
l2 norm of gradients: 0.7012235390729052
l2 norm of weights: 6.313477316976227

Iteration Number: 1557
Loss: 66.14296192954664
l2 norm of gradients: 0.7008386426755737
l2 norm of weights: 6.313308124110801

Iteration Number: 1558
Loss: 66.11886442914773
l2 norm of gradients: 0.7004540892681854
l2 norm of weights: 6.313139002049816

Iteration Number: 1559
Loss: 66.0947930990363
l2 norm of gradients: 0.7000698789047977
l2 norm of weights: 6.312969950706917

Iteration Number: 1560
Loss: 66.07074790164403
l2 norm of gradients: 0.6996860116386802
l2 norm of weights: 6.312800969995872

Iteration Number: 1561
Loss: 66.04672879928242
l2 norm of gradients: 0.6993024875223017
l2 norm of weights: 6.31263205983058

Iteration Number: 1562
Loss: 66.02273575444522
l2 norm of gradients: 0.6989193066073156
l2 norm of weights: 6.312463220125061

Iteration Number: 1563
Loss: 65.99876872969715
l2 norm of gradients: 0.6985364689445465
l2 norm of weights: 6.312294450793461

Iteration Number: 1564
Loss: 65.97482768757858
l2 norm of gradients: 0.698153974583976
l2 norm of weights: 6.312125751750053

Iteration Number: 1565
Loss: 65.95091259059946
l2 norm of gradients: 0.6977718235747301
l2 norm of weights: 6.311957122909235

Iteration Number: 1566
Loss: 65.92702340136852
l2 norm of gradients: 0.6973900159650641

```

12 norm of weights: 6.311788564185531
-----
Iteration Number: 1567
Loss: 65.90316008258823
12 norm of gradients: 0.6970085518023498
12 norm of weights: 6.311620075493589
-----
Iteration Number: 1568
Loss: 65.87932259682488
12 norm of gradients: 0.6966274311330622
12 norm of weights: 6.311451656748183
-----
Iteration Number: 1569
Loss: 65.85551090690906
12 norm of gradients: 0.6962466540027656
12 norm of weights: 6.311283307864213
-----
Iteration Number: 1570
Loss: 65.83172497552476
12 norm of gradients: 0.6958662204561002
12 norm of weights: 6.3111150287567055
-----
Iteration Number: 1571
Loss: 65.80796476541259
12 norm of gradients: 0.6954861305367691
12 norm of weights: 6.31094681934081
-----
Iteration Number: 1572
Loss: 65.78423023949867
12 norm of gradients: 0.6951063842875247
12 norm of weights: 6.310778679531803
-----
Iteration Number: 1573
Loss: 65.76052136053008
12 norm of gradients: 0.6947269817501556
12 norm of weights: 6.310610609245088
-----
Iteration Number: 1574
Loss: 65.73683809151244
12 norm of gradients: 0.6943479229654738
12 norm of weights: 6.31044260839619
-----
Iteration Number: 1575
Loss: 65.71318039526258
12 norm of gradients: 0.6939692079733015
12 norm of weights: 6.3102746769007645
-----
Iteration Number: 1576

```

Loss: 65.68954823476757
l2 norm of gradients: 0.6935908368124576
l2 norm of weights: 6.310106814674588

Iteration Number: 1577
Loss: 65.66594157303284
l2 norm of gradients: 0.6932128095207457
l2 norm of weights: 6.309939021633565

Iteration Number: 1578
Loss: 65.64236037311463
l2 norm of gradients: 0.6928351261349406
l2 norm of weights: 6.309771297693725

Iteration Number: 1579
Loss: 65.61880459803305
l2 norm of gradients: 0.6924577866907761
l2 norm of weights: 6.309603642771224

Iteration Number: 1580
Loss: 65.59527421095137
l2 norm of gradients: 0.6920807912229322
l2 norm of weights: 6.309436056782341

Iteration Number: 1581
Loss: 65.57176917494843
l2 norm of gradients: 0.6917041397650221
l2 norm of weights: 6.309268539643482

Iteration Number: 1582
Loss: 65.54828945321654
l2 norm of gradients: 0.6913278323495807
l2 norm of weights: 6.309101091271181

Iteration Number: 1583
Loss: 65.52483500895417
l2 norm of gradients: 0.6909518690080516
l2 norm of weights: 6.308933711582092

Iteration Number: 1584
Loss: 65.50140580548273
l2 norm of gradients: 0.6905762497707745
l2 norm of weights: 6.308766400492998

Iteration Number: 1585
Loss: 65.47800180596931
l2 norm of gradients: 0.6902009746669738
l2 norm of weights: 6.308599157920809

Iteration Number: 1586
Loss: 65.45462297385065
l2 norm of gradients: 0.6898260437247465
l2 norm of weights: 6.3084319837825555

Iteration Number: 1587
Loss: 65.43126927236872
l2 norm of gradients: 0.6894514569710495
l2 norm of weights: 6.308264877995399

Iteration Number: 1588
Loss: 65.40794066498549
l2 norm of gradients: 0.6890772144316882
l2 norm of weights: 6.3080978404766235

Iteration Number: 1589
Loss: 65.38463711512283
l2 norm of gradients: 0.6887033161313053
l2 norm of weights: 6.307930871143638

Iteration Number: 1590
Loss: 65.36135858624716
l2 norm of gradients: 0.6883297620933673
l2 norm of weights: 6.307763969913977

Iteration Number: 1591
Loss: 65.33810504190076
l2 norm of gradients: 0.6879565523401554
l2 norm of weights: 6.307597136705304

Iteration Number: 1592
Loss: 65.31487644550118
l2 norm of gradients: 0.6875836868927522
l2 norm of weights: 6.307430371435403

Iteration Number: 1593
Loss: 65.29167276065807
l2 norm of gradients: 0.6872111657710308
l2 norm of weights: 6.307263674022187

Iteration Number: 1594
Loss: 65.26849395104128
l2 norm of gradients: 0.6868389889936443
l2 norm of weights: 6.307097044383692

Iteration Number: 1595
Loss: 65.24533998029972

l2 norm of gradients: 0.6864671565780136
l2 norm of weights: 6.306930482438082

Iteration Number: 1596
Loss: 65.22221081205952
l2 norm of gradients: 0.6860956685403169
l2 norm of weights: 6.3067639881036435

Iteration Number: 1597
Loss: 65.19910641005077
l2 norm of gradients: 0.685724524895479
l2 norm of weights: 6.30659756129879

Iteration Number: 1598
Loss: 65.17602673802584
l2 norm of gradients: 0.6853537256571599
l2 norm of weights: 6.30643120194206

Iteration Number: 1599
Loss: 65.1529717597875
l2 norm of gradients: 0.6849832708377447
l2 norm of weights: 6.3062649099521195

Iteration Number: 1600
Loss: 65.12994143921627
l2 norm of gradients: 0.6846131604483331
l2 norm of weights: 6.306098685247755

Iteration Number: 1601
Loss: 65.1069357400815
l2 norm of gradients: 0.6842433944987281
l2 norm of weights: 6.305932527747882

Iteration Number: 1602
Loss: 65.0839546263354
l2 norm of gradients: 0.6838739729974265
l2 norm of weights: 6.305766437371544

Iteration Number: 1603
Loss: 65.06099806198323
l2 norm of gradients: 0.6835048959516082
l2 norm of weights: 6.305600414037901

Iteration Number: 1604
Loss: 65.03806601083333
l2 norm of gradients: 0.6831361633671272
l2 norm of weights: 6.305434457666247

Iteration Number: 1605
Loss: 65.0151584370739
l2 norm of gradients: 0.6827677752484997
l2 norm of weights: 6.305268568175997

Iteration Number: 1606
Loss: 64.99227530467292
l2 norm of gradients: 0.6823997315988959
l2 norm of weights: 6.305102745486693

Iteration Number: 1607
Loss: 64.96941657772925
l2 norm of gradients: 0.6820320324201298
l2 norm of weights: 6.3049369895180005

Iteration Number: 1608
Loss: 64.94658222038822
l2 norm of gradients: 0.6816646777126493
l2 norm of weights: 6.304771300189711

Iteration Number: 1609
Loss: 64.92377219676976
l2 norm of gradients: 0.6812976674755271
l2 norm of weights: 6.304605677421741

Iteration Number: 1610
Loss: 64.90098647107163
l2 norm of gradients: 0.6809310017064512
l2 norm of weights: 6.304440121134135

Iteration Number: 1611
Loss: 64.87822500759715
l2 norm of gradients: 0.6805646804017156
l2 norm of weights: 6.304274631247059

Iteration Number: 1612
Loss: 64.85548777049976
l2 norm of gradients: 0.6801987035562111
l2 norm of weights: 6.304109207680803

Iteration Number: 1613
Loss: 64.83277472422256
l2 norm of gradients: 0.6798330711634176
l2 norm of weights: 6.303943850355786

Iteration Number: 1614
Loss: 64.81008583304914
l2 norm of gradients: 0.679467783215393

```

12 norm of weights: 6.303778559192551
-----
Iteration Number: 1615
Loss: 64.78742106143227
12 norm of gradients: 0.6791028397027661
12 norm of weights: 6.303613334111765
-----
Iteration Number: 1616
Loss: 64.76478037368956
12 norm of gradients: 0.6787382406147284
12 norm of weights: 6.303448175034221
-----
Iteration Number: 1617
Loss: 64.74216373433994
12 norm of gradients: 0.6783739859390236
12 norm of weights: 6.303283081880837
-----
Iteration Number: 1618
Loss: 64.7195711078331
12 norm of gradients: 0.6780100756619417
12 norm of weights: 6.303118054572654
-----
Iteration Number: 1619
Loss: 64.69700245882079
12 norm of gradients: 0.6776465097683093
12 norm of weights: 6.30295309303084
-----
Iteration Number: 1620
Loss: 64.67445775176944
12 norm of gradients: 0.6772832882414821
12 norm of weights: 6.302788197176689
-----
Iteration Number: 1621
Loss: 64.65193695138036
12 norm of gradients: 0.6769204110633368
12 norm of weights: 6.302623366931616
-----
Iteration Number: 1622
Loss: 64.62944002225184
12 norm of gradients: 0.6765578782142632
12 norm of weights: 6.3024586022171665
-----
Iteration Number: 1623
Loss: 64.60696692905056
12 norm of gradients: 0.6761956896731574
12 norm of weights: 6.302293902955004
-----
Iteration Number: 1624

```



```

Loss: 64.58451763651318
l2 norm of gradients: 0.6758338454174128
l2 norm of weights: 6.3021292690669215
-----
Iteration Number: 1625
Loss: 64.56209210945353
l2 norm of gradients: 0.6754723454229143
l2 norm of weights: 6.301964700474836
-----
Iteration Number: 1626
Loss: 64.53969031263264
l2 norm of gradients: 0.6751111896640302
l2 norm of weights: 6.301800197100788
-----
Iteration Number: 1627
Loss: 64.51731221088816
l2 norm of gradients: 0.6747503781136049
l2 norm of weights: 6.301635758866946
-----
Iteration Number: 1628
Loss: 64.49495776911681
l2 norm of gradients: 0.6743899107429527
l2 norm of weights: 6.301471385695598
-----
Iteration Number: 1629
Loss: 64.47262695228054
l2 norm of gradients: 0.67402978752185
l2 norm of weights: 6.30130707750916
-----
Iteration Number: 1630
Loss: 64.45031972523698
l2 norm of gradients: 0.6736700084185299
l2 norm of weights: 6.301142834230173
-----
Iteration Number: 1631
Loss: 64.42803605309939
l2 norm of gradients: 0.6733105733996739
l2 norm of weights: 6.3009786557813
-----
Iteration Number: 1632
Loss: 64.40577590079033
l2 norm of gradients: 0.6729514824304074
l2 norm of weights: 6.300814542085331
-----
Iteration Number: 1633
Loss: 64.38353923344116
l2 norm of gradients: 0.672592735474292
l2 norm of weights: 6.300650493065178

```

```

-----
Iteration Number: 1634
Loss: 64.36132601617219
l2 norm of gradients: 0.6722343324933198
l2 norm of weights: 6.300486508643881
-----
Iteration Number: 1635
Loss: 64.3391362141048
l2 norm of gradients: 0.6718762734479073
l2 norm of weights: 6.3003225887446
-----
Iteration Number: 1636
Loss: 64.31696979242928
l2 norm of gradients: 0.6715185582968904
l2 norm of weights: 6.3001587332906235
-----
Iteration Number: 1637
Loss: 64.2948267163526
l2 norm of gradients: 0.6711611869975168
l2 norm of weights: 6.299994942205361
-----
Iteration Number: 1638
Loss: 64.27270695120548
l2 norm of gradients: 0.6708041595054421
l2 norm of weights: 6.299831215412348
-----
Iteration Number: 1639
Loss: 64.25061046229204
l2 norm of gradients: 0.6704474757747239
l2 norm of weights: 6.299667552835244
-----
Iteration Number: 1640
Loss: 64.22853721485124
l2 norm of gradients: 0.6700911357578154
l2 norm of weights: 6.299503954397832
-----
Iteration Number: 1641
Loss: 64.20648717437426
l2 norm of gradients: 0.6697351394055623
l2 norm of weights: 6.299340420024022
-----
Iteration Number: 1642
Loss: 64.18446030624207
l2 norm of gradients: 0.6693794866671956
l2 norm of weights: 6.299176949637841
-----
Iteration Number: 1643
Loss: 64.1624565759211

```

l2 norm of gradients: 0.6690241774903278
l2 norm of weights: 6.299013543163449

Iteration Number: 1644
Loss: 64.14047594890954
l2 norm of gradients: 0.6686692118209483
l2 norm of weights: 6.2988502005251235

Iteration Number: 1645
Loss: 64.11851839075321
l2 norm of gradients: 0.6683145896034182
l2 norm of weights: 6.298686921647269

Iteration Number: 1646
Loss: 64.09658386703994
l2 norm of gradients: 0.6679603107804659
l2 norm of weights: 6.298523706454412

Iteration Number: 1647
Loss: 64.07467234333323
l2 norm of gradients: 0.6676063752931825
l2 norm of weights: 6.298360554871204

Iteration Number: 1648
Loss: 64.05278378541237
l2 norm of gradients: 0.6672527830810183
l2 norm of weights: 6.29819746682242

Iteration Number: 1649
Loss: 64.03091815888745
l2 norm of gradients: 0.6668995340817783
l2 norm of weights: 6.298034442232959

Iteration Number: 1650
Loss: 64.00907542951309
l2 norm of gradients: 0.6665466282316179
l2 norm of weights: 6.297871481027844

Iteration Number: 1651
Loss: 63.98725556303459
l2 norm of gradients: 0.6661940654650395
l2 norm of weights: 6.297708583132219

Iteration Number: 1652
Loss: 63.96545852527502
l2 norm of gradients: 0.6658418457148882
l2 norm of weights: 6.297545748471355

Iteration Number: 1653
Loss: 63.943684282151985
l2 norm of gradients: 0.6654899689123497
l2 norm of weights: 6.297382976970643

Iteration Number: 1654
Loss: 63.92193279956818
l2 norm of gradients: 0.6651384349869446
l2 norm of weights: 6.297220268555601

Iteration Number: 1655
Loss: 63.900204043351835
l2 norm of gradients: 0.6647872438665272
l2 norm of weights: 6.297057623151868

Iteration Number: 1656
Loss: 63.878497979595046
l2 norm of gradients: 0.6644363954772815
l2 norm of weights: 6.2968950406852064

Iteration Number: 1657
Loss: 63.856814574265
l2 norm of gradients: 0.6640858897437174
l2 norm of weights: 6.296732521081503

Iteration Number: 1658
Loss: 63.835153793372356
l2 norm of gradients: 0.6637357265886693
l2 norm of weights: 6.296570064266767

Iteration Number: 1659
Loss: 63.81351560311817
l2 norm of gradients: 0.6633859059332922
l2 norm of weights: 6.29640767016713

Iteration Number: 1660
Loss: 63.79189996959875
l2 norm of gradients: 0.6630364276970597
l2 norm of weights: 6.2962453387088475

Iteration Number: 1661
Loss: 63.77030685893895
l2 norm of gradients: 0.6626872917977608
l2 norm of weights: 6.296083069818298

Iteration Number: 1662
Loss: 63.74873623740774
l2 norm of gradients: 0.6623384981514984

```

12 norm of weights: 6.295920863421984
-----
Iteration Number: 1663
Loss: 63.727188071312476
12 norm of gradients: 0.6619900466726866
12 norm of weights: 6.2957587194465265
-----
Iteration Number: 1664
Loss: 63.70566232683559
12 norm of gradients: 0.6616419372740483
12 norm of weights: 6.295596637818676
-----
Iteration Number: 1665
Loss: 63.68415897040132
12 norm of gradients: 0.6612941698666136
12 norm of weights: 6.295434618465298
-----
Iteration Number: 1666
Loss: 63.662677968350245
12 norm of gradients: 0.6609467443597179
12 norm of weights: 6.295272661313387
-----
Iteration Number: 1667
Loss: 63.64121928715295
12 norm of gradients: 0.6605996606610003
12 norm of weights: 6.295110766290056
-----
Iteration Number: 1668
Loss: 63.619782893226365
12 norm of gradients: 0.6602529186764015
12 norm of weights: 6.294948933322543
-----
Iteration Number: 1669
Loss: 63.5983687530715
12 norm of gradients: 0.659906518310163
12 norm of weights: 6.294787162338207
-----
Iteration Number: 1670
Loss: 63.5769768332437
12 norm of gradients: 0.6595604594648251
12 norm of weights: 6.29462545326453
-----
Iteration Number: 1671
Loss: 63.555607100381444
12 norm of gradients: 0.6592147420412258
12 norm of weights: 6.294463806029115
-----
Iteration Number: 1672

```

Loss: 63.5342595210079
l2 norm of gradients: 0.6588693659385005
l2 norm of weights: 6.294302220559688

Iteration Number: 1673
Loss: 63.51293406184585
l2 norm of gradients: 0.6585243310540797
l2 norm of weights: 6.2941406967840985

Iteration Number: 1674
Loss: 63.49163068960786
l2 norm of gradients: 0.6581796372836894
l2 norm of weights: 6.293979234630315

Iteration Number: 1675
Loss: 63.47034937104616
l2 norm of gradients: 0.6578352845213493
l2 norm of weights: 6.29381783402643

Iteration Number: 1676
Loss: 63.44909007295769
l2 norm of gradients: 0.6574912726593735
l2 norm of weights: 6.293656494900658

Iteration Number: 1677
Loss: 63.42785276214701
l2 norm of gradients: 0.6571476015883679
l2 norm of weights: 6.2934952171813325

Iteration Number: 1678
Loss: 63.406637405495076
l2 norm of gradients: 0.6568042711972325
l2 norm of weights: 6.293334000796913

Iteration Number: 1679
Loss: 63.38544396989463
l2 norm of gradients: 0.656461281373159
l2 norm of weights: 6.293172845675977

Iteration Number: 1680
Loss: 63.36427242234911
l2 norm of gradients: 0.656118632001632
l2 norm of weights: 6.293011751747224

Iteration Number: 1681
Loss: 63.34312272981483
l2 norm of gradients: 0.6557763229664278
l2 norm of weights: 6.292850718939477

```
-----  
Iteration Number: 1682  
Loss: 63.321994859329806  
l2 norm of gradients: 0.6554343541496158  
l2 norm of weights: 6.2926897471816785  
-----  
Iteration Number: 1683  
Loss: 63.30088877803683  
l2 norm of gradients: 0.6550927254315582  
l2 norm of weights: 6.292528836402892  
-----  
Iteration Number: 1684  
Loss: 63.27980445295129  
l2 norm of gradients: 0.6547514366909094  
l2 norm of weights: 6.2923679865323034  
-----  
Iteration Number: 1685  
Loss: 63.25874185134569  
l2 norm of gradients: 0.6544104878046184  
l2 norm of weights: 6.292207197499219  
-----  
Iteration Number: 1686  
Loss: 63.237700940384016  
l2 norm of gradients: 0.6540698786479275  
l2 norm of weights: 6.292046469233064  
-----  
Iteration Number: 1687  
Loss: 63.2166816872527  
l2 norm of gradients: 0.6537296090943744  
l2 norm of weights: 6.2918858016633905  
-----  
Iteration Number: 1688  
Loss: 63.19568405931035  
l2 norm of gradients: 0.6533896790157919  
l2 norm of weights: 6.291725194719863  
-----  
Iteration Number: 1689  
Loss: 63.174708023881436  
l2 norm of gradients: 0.6530500882823095  
l2 norm of weights: 6.291564648332273  
-----  
Iteration Number: 1690  
Loss: 63.15375354837836  
l2 norm of gradients: 0.6527108367623543  
l2 norm of weights: 6.291404162430532  
-----  
Iteration Number: 1691  
Loss: 63.13282060010359
```

l2 norm of gradients: 0.6523719243226521
l2 norm of weights: 6.2912437369446685

Iteration Number: 1692
Loss: 63.1119091466105
l2 norm of gradients: 0.6520333508282284
l2 norm of weights: 6.291083371804834

Iteration Number: 1693
Loss: 63.09101915539185
l2 norm of gradients: 0.6516951161424103
l2 norm of weights: 6.290923066941299

Iteration Number: 1694
Loss: 63.07015059389481
l2 norm of gradients: 0.6513572201268273
l2 norm of weights: 6.290762822284455

Iteration Number: 1695
Loss: 63.049303429810564
l2 norm of gradients: 0.6510196626414135
l2 norm of weights: 6.290602637764814

Iteration Number: 1696
Loss: 63.028477630711215
l2 norm of gradients: 0.6506824435444093
l2 norm of weights: 6.290442513313007

Iteration Number: 1697
Loss: 63.007673164331166
l2 norm of gradients: 0.6503455626923619
l2 norm of weights: 6.290282448859785

Iteration Number: 1698
Loss: 62.986889998275124
l2 norm of gradients: 0.6500090199401289
l2 norm of weights: 6.290122444336021

Iteration Number: 1699
Loss: 62.96612810035366
l2 norm of gradients: 0.6496728151408798
l2 norm of weights: 6.2899624996727015

Iteration Number: 1700
Loss: 62.945387438419246
l2 norm of gradients: 0.6493369481460974
l2 norm of weights: 6.28980261480094

Iteration Number: 1701
Loss: 62.924667980209634
l2 norm of gradients: 0.6490014188055805
l2 norm of weights: 6.289642789651965

Iteration Number: 1702
Loss: 62.90396969366194
l2 norm of gradients: 0.6486662269674461
l2 norm of weights: 6.289483024157125

Iteration Number: 1703
Loss: 62.88329254667841
l2 norm of gradients: 0.6483313724781324
l2 norm of weights: 6.28932331824789

Iteration Number: 1704
Loss: 62.862636507288684
l2 norm of gradients: 0.6479968551824006
l2 norm of weights: 6.289163671855845

Iteration Number: 1705
Loss: 62.84200154336184
l2 norm of gradients: 0.647662674923337
l2 norm of weights: 6.2890040849126985

Iteration Number: 1706
Loss: 62.821387623122526
l2 norm of gradients: 0.647328831542357
l2 norm of weights: 6.2888445573502745

Iteration Number: 1707
Loss: 62.80079471450641
l2 norm of gradients: 0.6469953248792075
l2 norm of weights: 6.288685089100516

Iteration Number: 1708
Loss: 62.780222785764316
l2 norm of gradients: 0.6466621547719689
l2 norm of weights: 6.288525680095488

Iteration Number: 1709
Loss: 62.75967180502904
l2 norm of gradients: 0.646329321057059
l2 norm of weights: 6.28836633026737

Iteration Number: 1710
Loss: 62.73914174057476
l2 norm of gradients: 0.6459968235692362

```

12 norm of weights: 6.288207039548462
-----
Iteration Number: 1711
Loss: 62.71863256057514
12 norm of gradients: 0.6456646621416016
12 norm of weights: 6.288047807871182
-----
Iteration Number: 1712
Loss: 62.698144233449014
12 norm of gradients: 0.6453328366056033
12 norm of weights: 6.287888635168068
-----
Iteration Number: 1713
Loss: 62.677676727442744
12 norm of gradients: 0.6450013467910397
12 norm of weights: 6.28772952137177
-----
Iteration Number: 1714
Loss: 62.65723001101844
12 norm of gradients: 0.6446701925260621
12 norm of weights: 6.2875704664150645
-----
Iteration Number: 1715
Loss: 62.636804052610515
12 norm of gradients: 0.6443393736371793
12 norm of weights: 6.287411470230841
-----
Iteration Number: 1716
Loss: 62.616398820738496
12 norm of gradients: 0.6440088899492604
12 norm of weights: 6.287252532752104
-----
Iteration Number: 1717
Loss: 62.5960142838794
12 norm of gradients: 0.643678741285539
12 norm of weights: 6.287093653911983
-----
Iteration Number: 1718
Loss: 62.57565041053179
12 norm of gradients: 0.6433489274676166
12 norm of weights: 6.28693483364372
-----
Iteration Number: 1719
Loss: 62.55530716946004
12 norm of gradients: 0.6430194483154668
12 norm of weights: 6.286776071880673
-----
Iteration Number: 1720

```

Loss: 62.534984529229895
l2 norm of gradients: 0.6426903036474393
l2 norm of weights: 6.286617368556322

Iteration Number: 1721
Loss: 62.51468245854918
l2 norm of gradients: 0.6423614932802637
l2 norm of weights: 6.2864587236042615

Iteration Number: 1722
Loss: 62.49440092615412
l2 norm of gradients: 0.6420330170290537
l2 norm of weights: 6.286300136958203

Iteration Number: 1723
Loss: 62.47413990089916
l2 norm of gradients: 0.6417048747073115
l2 norm of weights: 6.286141608551974

Iteration Number: 1724
Loss: 62.45389935156028
l2 norm of gradients: 0.6413770661269319
l2 norm of weights: 6.28598313831952

Iteration Number: 1725
Loss: 62.43367924696693
l2 norm of gradients: 0.6410495910982066
l2 norm of weights: 6.285824726194905

Iteration Number: 1726
Loss: 62.41347955614757
l2 norm of gradients: 0.6407224494298293
l2 norm of weights: 6.2856663721123045

Iteration Number: 1727
Loss: 62.39330024796141
l2 norm of gradients: 0.6403956409288988
l2 norm of weights: 6.285508076006015

Iteration Number: 1728
Loss: 62.37314129152395
l2 norm of gradients: 0.6400691654009254
l2 norm of weights: 6.285349837810448

Iteration Number: 1729
Loss: 62.35300265581112
l2 norm of gradients: 0.6397430226498342
l2 norm of weights: 6.285191657460129

```

-----
Iteration Number: 1730
Loss: 62.332884309872455
l2 norm of gradients: 0.6394172124779708
l2 norm of weights: 6.285033534889702
-----
Iteration Number: 1731
Loss: 62.3127862229262
l2 norm of gradients: 0.6390917346861051
l2 norm of weights: 6.284875470033925
-----
Iteration Number: 1732
Loss: 62.292708364154585
l2 norm of gradients: 0.6387665890734376
l2 norm of weights: 6.2847174628276745
-----
Iteration Number: 1733
Loss: 62.27265070273722
l2 norm of gradients: 0.6384417754376028
l2 norm of weights: 6.284559513205939
-----
Iteration Number: 1734
Loss: 62.25261320800818
l2 norm of gradients: 0.6381172935746755
l2 norm of weights: 6.284401621103823
-----
Iteration Number: 1735
Loss: 62.23259584916477
l2 norm of gradients: 0.6377931432791756
l2 norm of weights: 6.284243786456551
-----
Iteration Number: 1736
Loss: 62.21259859569317
l2 norm of gradients: 0.6374693243440722
l2 norm of weights: 6.284086009199455
-----
Iteration Number: 1737
Loss: 62.19262141692931
l2 norm of gradients: 0.637145836560791
l2 norm of weights: 6.283928289267988
-----
Iteration Number: 1738
Loss: 62.172664282336534
l2 norm of gradients: 0.6368226797192181
l2 norm of weights: 6.283770626597715
-----
Iteration Number: 1739
Loss: 62.15272716140784

```

l2 norm of gradients: 0.6364998536077053
l2 norm of weights: 6.283613021124317

Iteration Number: 1740
Loss: 62.13281002362274
l2 norm of gradients: 0.6361773580130761
l2 norm of weights: 6.28345547278359

Iteration Number: 1741
Loss: 62.11291283869987
l2 norm of gradients: 0.6358551927206313
l2 norm of weights: 6.283297981511444

Iteration Number: 1742
Loss: 62.09303557611239
l2 norm of gradients: 0.6355333575141542
l2 norm of weights: 6.2831405472439

Iteration Number: 1743
Loss: 62.07317820560934
l2 norm of gradients: 0.6352118521759165
l2 norm of weights: 6.2829831699171

Iteration Number: 1744
Loss: 62.05334069685492
l2 norm of gradients: 0.6348906764866838
l2 norm of weights: 6.282825849467293

Iteration Number: 1745
Loss: 62.03352301965713
l2 norm of gradients: 0.634569830225722
l2 norm of weights: 6.282668585830846

Iteration Number: 1746
Loss: 62.01372514377594
l2 norm of gradients: 0.6342493131708019
l2 norm of weights: 6.282511378944241

Iteration Number: 1747
Loss: 61.99394703911147
l2 norm of gradients: 0.6339291250982063
l2 norm of weights: 6.282354228744068

Iteration Number: 1748
Loss: 61.974188675503
l2 norm of gradients: 0.6336092657827352
l2 norm of weights: 6.282197135167036

Iteration Number: 1749
Loss: 61.95445002289221
l2 norm of gradients: 0.6332897349977124
l2 norm of weights: 6.282040098149964

Iteration Number: 1750
Loss: 61.93473105126038
l2 norm of gradients: 0.6329705325149907
l2 norm of weights: 6.281883117629786

Iteration Number: 1751
Loss: 61.91503173063883
l2 norm of gradients: 0.6326516581049588
l2 norm of weights: 6.281726193543547

Iteration Number: 1752
Loss: 61.8953520310939
l2 norm of gradients: 0.6323331115365475
l2 norm of weights: 6.281569325828408

Iteration Number: 1753
Loss: 61.87569192270038
l2 norm of gradients: 0.6320148925772351
l2 norm of weights: 6.28141251442164

Iteration Number: 1754
Loss: 61.856051375740314
l2 norm of gradients: 0.6316970009930545
l2 norm of weights: 6.281255759260628

Iteration Number: 1755
Loss: 61.836430360268665
l2 norm of gradients: 0.631379436548599
l2 norm of weights: 6.281099060282868

Iteration Number: 1756
Loss: 61.81682884661765
l2 norm of gradients: 0.6310621990070289
l2 norm of weights: 6.280942417425969

Iteration Number: 1757
Loss: 61.797246805024685
l2 norm of gradients: 0.6307452881300785
l2 norm of weights: 6.280785830627653

Iteration Number: 1758
Loss: 61.77768420589285
l2 norm of gradients: 0.6304287036780609

```

12 norm of weights: 6.280629299825751
-----
Iteration Number: 1759
Loss: 61.75814101954483
12 norm of gradients: 0.6301124454098765
12 norm of weights: 6.280472824958211
-----
Iteration Number: 1760
Loss: 61.73861721646698
12 norm of gradients: 0.6297965130830185
12 norm of weights: 6.280316405963087
-----
Iteration Number: 1761
Loss: 61.719112767032044
12 norm of gradients: 0.6294809064535793
12 norm of weights: 6.280160042778549
-----
Iteration Number: 1762
Loss: 61.69962764186496
12 norm of gradients: 0.629165625276258
12 norm of weights: 6.280003735342875
-----
Iteration Number: 1763
Loss: 61.68016181147847
12 norm of gradients: 0.6288506693043661
12 norm of weights: 6.279847483594456
-----
Iteration Number: 1764
Loss: 61.66071524647588
12 norm of gradients: 0.6285360382898356
12 norm of weights: 6.2796912874717945
-----
Iteration Number: 1765
Loss: 61.6412879174728
12 norm of gradients: 0.628221731983224
12 norm of weights: 6.2795351469135
-----
Iteration Number: 1766
Loss: 61.621879795289146
12 norm of gradients: 0.627907750133723
12 norm of weights: 6.2793790618582985
-----
Iteration Number: 1767
Loss: 61.60249085051607
12 norm of gradients: 0.6275940924891639
12 norm of weights: 6.279223032245023
-----
Iteration Number: 1768

```

```

Loss: 61.58312105402644
l2 norm of gradients: 0.6272807587960255
l2 norm of weights: 6.279067058012616
-----
Iteration Number: 1769
Loss: 61.563770376597695
l2 norm of gradients: 0.6269677487994402
l2 norm of weights: 6.278911139100134
-----
Iteration Number: 1770
Loss: 61.54443878915805
l2 norm of gradients: 0.6266550622432018
l2 norm of weights: 6.27875527544674
-----
Iteration Number: 1771
Loss: 61.52512626264303
l2 norm of gradients: 0.6263426988697721
l2 norm of weights: 6.278599466991709
-----
Iteration Number: 1772
Loss: 61.50583276796351
l2 norm of gradients: 0.6260306584202878
l2 norm of weights: 6.278443713674424
-----
Iteration Number: 1773
Loss: 61.48655827613581
l2 norm of gradients: 0.625718940634568
l2 norm of weights: 6.278288015434378
-----
Iteration Number: 1774
Loss: 61.46730275822213
l2 norm of gradients: 0.625407545251121
l2 norm of weights: 6.278132372211177
-----
Iteration Number: 1775
Loss: 61.44806618532907
l2 norm of gradients: 0.625096472007152
l2 norm of weights: 6.2779767839445295
-----
Iteration Number: 1776
Loss: 61.42884852860147
l2 norm of gradients: 0.6247857206385694
l2 norm of weights: 6.277821250574259
-----
Iteration Number: 1777
Loss: 61.40964975924517
l2 norm of gradients: 0.6244752908799928
l2 norm of weights: 6.277665772040295

```



```

-----
Iteration Number: 1778
Loss: 61.3904698485127
l2 norm of gradients: 0.6241651824647602
l2 norm of weights: 6.277510348282676
-----
Iteration Number: 1779
Loss: 61.37130876764503
l2 norm of gradients: 0.6238553951249345
l2 norm of weights: 6.277354979241551
-----
Iteration Number: 1780
Loss: 61.352166488003554
l2 norm of gradients: 0.6235459285913123
l2 norm of weights: 6.277199664857174
-----
Iteration Number: 1781
Loss: 61.33304298090965
l2 norm of gradients: 0.62323678259343
l2 norm of weights: 6.277044405069911
-----
Iteration Number: 1782
Loss: 61.31393821778433
l2 norm of gradients: 0.6229279568595709
l2 norm of weights: 6.276889199820234
-----
Iteration Number: 1783
Loss: 61.29485217019427
l2 norm of gradients: 0.6226194511167743
l2 norm of weights: 6.276734049048722
-----
Iteration Number: 1784
Loss: 61.27578480950495
l2 norm of gradients: 0.6223112650908416
l2 norm of weights: 6.276578952696064
-----
Iteration Number: 1785
Loss: 61.256736107370564
l2 norm of gradients: 0.6220033985063436
l2 norm of weights: 6.276423910703056
-----
Iteration Number: 1786
Loss: 61.23770603533973
l2 norm of gradients: 0.6216958510866291
l2 norm of weights: 6.2762689230106
-----
Iteration Number: 1787
Loss: 61.218694565069534

```

l2 norm of gradients: 0.6213886225538312
l2 norm of weights: 6.276113989559707

Iteration Number: 1788
Loss: 61.19970166824653
l2 norm of gradients: 0.6210817126288761
l2 norm of weights: 6.2759591102914944

Iteration Number: 1789
Loss: 61.18072731663273
l2 norm of gradients: 0.6207751210314898
l2 norm of weights: 6.275804285147185

Iteration Number: 1790
Loss: 61.161771481967094
l2 norm of gradients: 0.620468847480205
l2 norm of weights: 6.275649514068111

Iteration Number: 1791
Loss: 61.1428341361004
l2 norm of gradients: 0.6201628916923706
l2 norm of weights: 6.27549479699571

Iteration Number: 1792
Loss: 61.12391525090306
l2 norm of gradients: 0.6198572533841579
l2 norm of weights: 6.275340133871525

Iteration Number: 1793
Loss: 61.105014798253755
l2 norm of gradients: 0.6195519322705687
l2 norm of weights: 6.275185524637208

Iteration Number: 1794
Loss: 61.086132750156956
l2 norm of gradients: 0.6192469280654427
l2 norm of weights: 6.275030969234512

Iteration Number: 1795
Loss: 61.06726907858011
l2 norm of gradients: 0.6189422404814654
l2 norm of weights: 6.274876467605303

Iteration Number: 1796
Loss: 61.04842375557596
l2 norm of gradients: 0.6186378692301762
l2 norm of weights: 6.274722019691546

Iteration Number: 1797
Loss: 61.029596753294534
l2 norm of gradients: 0.6183338140219747
l2 norm of weights: 6.274567625435314

Iteration Number: 1798
Loss: 61.0107880438814
l2 norm of gradients: 0.6180300745661309
l2 norm of weights: 6.274413284778787

Iteration Number: 1799
Loss: 60.99199759942688
l2 norm of gradients: 0.6177266505707899
l2 norm of weights: 6.274258997664249

Iteration Number: 1800
Loss: 60.97322539222604
l2 norm of gradients: 0.6174235417429826
l2 norm of weights: 6.274104764034087

Iteration Number: 1801
Loss: 60.95447139462386
l2 norm of gradients: 0.6171207477886311
l2 norm of weights: 6.273950583830795

Iteration Number: 1802
Loss: 60.935735578804724
l2 norm of gradients: 0.616818268412558
l2 norm of weights: 6.2737964569969735

Iteration Number: 1803
Loss: 60.9170179172155
l2 norm of gradients: 0.6165161033184936
l2 norm of weights: 6.273642383475323

Iteration Number: 1804
Loss: 60.89831838227276
l2 norm of gradients: 0.6162142522090838
l2 norm of weights: 6.27348836320865

Iteration Number: 1805
Loss: 60.87963694644031
l2 norm of gradients: 0.6159127147858982
l2 norm of weights: 6.273334396139869

Iteration Number: 1806
Loss: 60.86097358217974
l2 norm of gradients: 0.615611490749437

```

12 norm of weights: 6.27318048221199
-----
Iteration Number: 1807
Loss: 60.84232826208623
12 norm of gradients: 0.6153105797991406
12 norm of weights: 6.273026621368136
-----
Iteration Number: 1808
Loss: 60.82370095872278
12 norm of gradients: 0.6150099816333955
12 norm of weights: 6.272872813551528
-----
Iteration Number: 1809
Loss: 60.805091644725564
12 norm of gradients: 0.6147096959495434
12 norm of weights: 6.272719058705491
-----
Iteration Number: 1810
Loss: 60.78650029283172
12 norm of gradients: 0.6144097224438887
12 norm of weights: 6.2725653567734545
-----
Iteration Number: 1811
Loss: 60.76792687573057
12 norm of gradients: 0.6141100608117068
12 norm of weights: 6.272411707698951
-----
Iteration Number: 1812
Loss: 60.74937136622825
12 norm of gradients: 0.6138107107472512
12 norm of weights: 6.2722581114256135
-----
Iteration Number: 1813
Loss: 60.730833737102024
12 norm of gradients: 0.613511671943762
12 norm of weights: 6.272104567897182
-----
Iteration Number: 1814
Loss: 60.71231396124743
12 norm of gradients: 0.6132129440934735
12 norm of weights: 6.271951077057496
-----
Iteration Number: 1815
Loss: 60.69381201157603
12 norm of gradients: 0.6129145268876226
12 norm of weights: 6.2717976388504955
-----
Iteration Number: 1816

```

Loss: 60.67532786100406
l2 norm of gradients: 0.6126164200164557
l2 norm of weights: 6.271644253220228

Iteration Number: 1817
Loss: 60.656861482586734
l2 norm of gradients: 0.612318623169238
l2 norm of weights: 6.271490920110838

Iteration Number: 1818
Loss: 60.638412849383926
l2 norm of gradients: 0.6120211360342604
l2 norm of weights: 6.271337639466575

Iteration Number: 1819
Loss: 60.61998193442998
l2 norm of gradients: 0.6117239582988477
l2 norm of weights: 6.271184411231787

Iteration Number: 1820
Loss: 60.60156871092995
l2 norm of gradients: 0.6114270896493665
l2 norm of weights: 6.271031235350928

Iteration Number: 1821
Loss: 60.58317315201482
l2 norm of gradients: 0.6111305297712335
l2 norm of weights: 6.270878111768548

Iteration Number: 1822
Loss: 60.564795230965665
l2 norm of gradients: 0.6108342783489229
l2 norm of weights: 6.2707250404293005

Iteration Number: 1823
Loss: 60.546434920996944
l2 norm of gradients: 0.6105383350659743
l2 norm of weights: 6.270572021277941

Iteration Number: 1824
Loss: 60.52809219547316
l2 norm of gradients: 0.6102426996050015
l2 norm of weights: 6.270419054259323

Iteration Number: 1825
Loss: 60.509767027741574
l2 norm of gradients: 0.6099473716476992
l2 norm of weights: 6.270266139318403

Iteration Number: 1826
Loss: 60.491459391217035
l2 norm of gradients: 0.6096523508748521
l2 norm of weights: 6.270113276400237

Iteration Number: 1827
Loss: 60.47316925930846
l2 norm of gradients: 0.609357636966342
l2 norm of weights: 6.26996046544998

Iteration Number: 1828
Loss: 60.454896605571044
l2 norm of gradients: 0.6090632296011559
l2 norm of weights: 6.269807706412888

Iteration Number: 1829
Loss: 60.43664140355328
l2 norm of gradients: 0.6087691284573942
l2 norm of weights: 6.269654999234317

Iteration Number: 1830
Loss: 60.418403626852076
l2 norm of gradients: 0.6084753332122784
l2 norm of weights: 6.269502343859721

Iteration Number: 1831
Loss: 60.40018324905685
l2 norm of gradients: 0.6081818435421592
l2 norm of weights: 6.269349740234656

Iteration Number: 1832
Loss: 60.3819802439027
l2 norm of gradients: 0.6078886591225243
l2 norm of weights: 6.269197188304774

Iteration Number: 1833
Loss: 60.36379458506898
l2 norm of gradients: 0.607595779628006
l2 norm of weights: 6.269044688015828

Iteration Number: 1834
Loss: 60.345626246356034
l2 norm of gradients: 0.6073032047323899
l2 norm of weights: 6.26889223931367

Iteration Number: 1835
Loss: 60.327475201538505

l2 norm of gradients: 0.6070109341086216
l2 norm of weights: 6.268739842144251

Iteration Number: 1836
Loss: 60.309341424551064
l2 norm of gradients: 0.6067189674288161
l2 norm of weights: 6.268587496453617

Iteration Number: 1837
Loss: 60.29122488921047
l2 norm of gradients: 0.6064273043642645
l2 norm of weights: 6.268435202187916

Iteration Number: 1838
Loss: 60.273125569565835
l2 norm of gradients: 0.6061359445854423
l2 norm of weights: 6.268282959293393

Iteration Number: 1839
Loss: 60.25504343952347
l2 norm of gradients: 0.6058448877620172
l2 norm of weights: 6.268130767716391

Iteration Number: 1840
Loss: 60.23697847317852
l2 norm of gradients: 0.6055541335628568
l2 norm of weights: 6.26797862740335

Iteration Number: 1841
Loss: 60.21893064462038
l2 norm of gradients: 0.6052636816560374
l2 norm of weights: 6.267826538300808

Iteration Number: 1842
Loss: 60.200899927930955
l2 norm of gradients: 0.6049735317088505
l2 norm of weights: 6.267674500355399

Iteration Number: 1843
Loss: 60.18288629735027
l2 norm of gradients: 0.6046836833878113
l2 norm of weights: 6.267522513513858

Iteration Number: 1844
Loss: 60.16488972709025
l2 norm of gradients: 0.6043941363586666
l2 norm of weights: 6.267370577723011

Iteration Number: 1845
Loss: 60.14691019135943
l2 norm of gradients: 0.6041048902864025
l2 norm of weights: 6.267218692929786

Iteration Number: 1846
Loss: 60.1289476645615
l2 norm of gradients: 0.6038159448352524
l2 norm of weights: 6.267066859081206

Iteration Number: 1847
Loss: 60.11100212094703
l2 norm of gradients: 0.6035272996687042
l2 norm of weights: 6.266915076124388

Iteration Number: 1848
Loss: 60.09307353495884
l2 norm of gradients: 0.603238954449509
l2 norm of weights: 6.266763344006549

Iteration Number: 1849
Loss: 60.07516188105073
l2 norm of gradients: 0.6029509088396878
l2 norm of weights: 6.266611662674998

Iteration Number: 1850
Loss: 60.05726713375949
l2 norm of gradients: 0.6026631625005404
l2 norm of weights: 6.266460032077144

Iteration Number: 1851
Loss: 60.039389267535086
l2 norm of gradients: 0.6023757150926523
l2 norm of weights: 6.266308452160489

Iteration Number: 1852
Loss: 60.02152825700705
l2 norm of gradients: 0.6020885662759027
l2 norm of weights: 6.2661569228726295

Iteration Number: 1853
Loss: 60.00368407676165
l2 norm of gradients: 0.6018017157094723
l2 norm of weights: 6.2660054441612605

Iteration Number: 1854
Loss: 59.98585670151084
l2 norm of gradients: 0.601515163051851


```
12 norm of weights: 6.265854015974169
-----
Iteration Number: 1855
Loss: 59.96804610597661
12 norm of gradients: 0.6012289079608455
12 norm of weights: 6.2657026382592385
-----
Iteration Number: 1856
Loss: 59.95025226489798
12 norm of gradients: 0.6009429500935872
12 norm of weights: 6.265551310964447
-----
Iteration Number: 1857
Loss: 59.93247515308394
12 norm of gradients: 0.6006572891065395
12 norm of weights: 6.265400034037866
-----
Iteration Number: 1858
Loss: 59.91471474535819
12 norm of gradients: 0.6003719246555057
12 norm of weights: 6.265248807427663
-----
Iteration Number: 1859
Loss: 59.896971016619304
12 norm of gradients: 0.6000868563956366
12 norm of weights: 6.265097631082098
-----
Iteration Number: 1860
Loss: 59.879243941810664
12 norm of gradients: 0.599802083981438
12 norm of weights: 6.264946504949524
-----
Iteration Number: 1861
Loss: 59.86153349591482
12 norm of gradients: 0.599517607066779
12 norm of weights: 6.264795428978393
-----
Iteration Number: 1862
Loss: 59.84383965396129
12 norm of gradients: 0.5992334253048983
12 norm of weights: 6.264644403117242
-----
Iteration Number: 1863
Loss: 59.82616239097852
12 norm of gradients: 0.5989495383484126
12 norm of weights: 6.26449342731471
-----
Iteration Number: 1864
```

```
Loss: 59.80850168217735
12 norm of gradients: 0.5986659458493248
12 norm of weights: 6.264342501519522
-----
Iteration Number: 1865
Loss: 59.79085750262871
12 norm of gradients: 0.5983826474590292
12 norm of weights: 6.2641916256805015
-----
Iteration Number: 1866
Loss: 59.77322982761456
12 norm of gradients: 0.5980996428283224
12 norm of weights: 6.264040799746561
-----
Iteration Number: 1867
Loss: 59.75561863228966
12 norm of gradients: 0.5978169316074077
12 norm of weights: 6.263890023666707
-----
Iteration Number: 1868
Loss: 59.73802389201613
12 norm of gradients: 0.5975345134459044
12 norm of weights: 6.263739297390039
-----
Iteration Number: 1869
Loss: 59.72044558209355
12 norm of gradients: 0.5972523879928546
12 norm of weights: 6.263588620865748
-----
Iteration Number: 1870
Loss: 59.70288367795386
12 norm of gradients: 0.5969705548967307
12 norm of weights: 6.263437994043116
-----
Iteration Number: 1871
Loss: 59.685338154924885
12 norm of gradients: 0.596689013805443
12 norm of weights: 6.263287416871518
-----
Iteration Number: 1872
Loss: 59.66780898858194
12 norm of gradients: 0.5964077643663471
12 norm of weights: 6.26313688930042
-----
Iteration Number: 1873
Loss: 59.65029615434432
12 norm of gradients: 0.5961268062262512
12 norm of weights: 6.262986411279382
```

```

-----
Iteration Number: 1874
Loss: 59.632799627867364
l2 norm of gradients: 0.5958461390314231
l2 norm of weights: 6.262835982758051
-----
Iteration Number: 1875
Loss: 59.61531938466211
l2 norm of gradients: 0.5955657624275982
l2 norm of weights: 6.262685603686167
-----
Iteration Number: 1876
Loss: 59.597855400425885
l2 norm of gradients: 0.5952856760599865
l2 norm of weights: 6.262535274013562
-----
Iteration Number: 1877
Loss: 59.58040765087673
l2 norm of gradients: 0.5950058795732799
l2 norm of weights: 6.262384993690156
-----
Iteration Number: 1878
Loss: 59.56297611166058
l2 norm of gradients: 0.5947263726116593
l2 norm of weights: 6.262234762665963
-----
Iteration Number: 1879
Loss: 59.54556075862738
l2 norm of gradients: 0.5944471548188023
l2 norm of weights: 6.262084580891083
-----
Iteration Number: 1880
Loss: 59.528161567547045
l2 norm of gradients: 0.5941682258378902
l2 norm of weights: 6.261934448315712
-----
Iteration Number: 1881
Loss: 59.510778514346555
l2 norm of gradients: 0.5938895853116151
l2 norm of weights: 6.261784364890127
-----
Iteration Number: 1882
Loss: 59.493411574843954
l2 norm of gradients: 0.5936112328821871
l2 norm of weights: 6.2616343305647035
-----
Iteration Number: 1883
Loss: 59.47606072508378

```

l2 norm of gradients: 0.5933331681913421
l2 norm of weights: 6.261484345289903

Iteration Number: 1884
Loss: 59.45872594103063
l2 norm of gradients: 0.5930553908803475
l2 norm of weights: 6.261334409016273

Iteration Number: 1885
Loss: 59.44140719871833
l2 norm of gradients: 0.5927779005900111
l2 norm of weights: 6.261184521694458

Iteration Number: 1886
Loss: 59.42410447419397
l2 norm of gradients: 0.5925006969606871
l2 norm of weights: 6.2610346832751835

Iteration Number: 1887
Loss: 59.40681774370442
l2 norm of gradients: 0.5922237796322831
l2 norm of weights: 6.260884893709268

Iteration Number: 1888
Loss: 59.389546983283104
l2 norm of gradients: 0.591947148244268
l2 norm of weights: 6.2607351529476185

Iteration Number: 1889
Loss: 59.37229216921827
l2 norm of gradients: 0.591670802435678
l2 norm of weights: 6.260585460941227

Iteration Number: 1890
Loss: 59.355053277779916
l2 norm of gradients: 0.5913947418451246
l2 norm of weights: 6.260435817641178

Iteration Number: 1891
Loss: 59.337830285174164
l2 norm of gradients: 0.5911189661108006
l2 norm of weights: 6.260286222998641

Iteration Number: 1892
Loss: 59.320623167890936
l2 norm of gradients: 0.590843474870488
l2 norm of weights: 6.260136676964875

Iteration Number: 1893
Loss: 59.30343190218734
l2 norm of gradients: 0.5905682677615642
l2 norm of weights: 6.259987179491226

Iteration Number: 1894
Loss: 59.2862564645882
l2 norm of gradients: 0.5902933444210096
l2 norm of weights: 6.259837730529127

Iteration Number: 1895
Loss: 59.26909683153142
l2 norm of gradients: 0.5900187044854134
l2 norm of weights: 6.259688330030099

Iteration Number: 1896
Loss: 59.251952979532824
l2 norm of gradients: 0.5897443475909817
l2 norm of weights: 6.2595389779457475

Iteration Number: 1897
Loss: 59.234824885142366
l2 norm of gradients: 0.5894702733735434
l2 norm of weights: 6.25938967422777

Iteration Number: 1898
Loss: 59.21771252501083
l2 norm of gradients: 0.5891964814685577
l2 norm of weights: 6.259240418827945

Iteration Number: 1899
Loss: 59.200615875725866
l2 norm of gradients: 0.5889229715111205
l2 norm of weights: 6.2590912116981405

Iteration Number: 1900
Loss: 59.18353491402747
l2 norm of gradients: 0.5886497431359708
l2 norm of weights: 6.258942052790311

Iteration Number: 1901
Loss: 59.16646961665288
l2 norm of gradients: 0.5883767959774986
l2 norm of weights: 6.258792942056496

Iteration Number: 1902
Loss: 59.149419960341454
l2 norm of gradients: 0.5881041296697498

```

12 norm of weights: 6.25864387944882
-----
Iteration Number: 1903
Loss: 59.13238592196116
12 norm of gradients: 0.5878317438464351
12 norm of weights: 6.258494864919496
-----
Iteration Number: 1904
Loss: 59.11536747833477
12 norm of gradients: 0.5875596381409345
12 norm of weights: 6.258345898420819
-----
Iteration Number: 1905
Loss: 59.098364606371845
12 norm of gradients: 0.5872878121863054
12 norm of weights: 6.258196979905172
-----
Iteration Number: 1906
Loss: 59.08137728302745
12 norm of gradients: 0.5870162656152886
12 norm of weights: 6.2580481093250215
-----
Iteration Number: 1907
Loss: 59.064405485362464
12 norm of gradients: 0.5867449980603147
12 norm of weights: 6.257899286632921
-----
Iteration Number: 1908
Loss: 59.04744919028935
12 norm of gradients: 0.5864740091535112
12 norm of weights: 6.257750511781506
-----
Iteration Number: 1909
Loss: 59.03050837499495
12 norm of gradients: 0.5862032985267087
12 norm of weights: 6.257601784723498
-----
Iteration Number: 1910
Loss: 59.01358301651987
12 norm of gradients: 0.5859328658114472
12 norm of weights: 6.257453105411704
-----
Iteration Number: 1911
Loss: 58.99667309202319
12 norm of gradients: 0.5856627106389828
12 norm of weights: 6.257304473799012
-----
Iteration Number: 1912

```

```

Loss: 58.979778578807796
l2 norm of gradients: 0.5853928326402941
l2 norm of weights: 6.257155889838398
-----
Iteration Number: 1913
Loss: 58.962899454043836
l2 norm of gradients: 0.5851232314460888
l2 norm of weights: 6.257007353482917
-----
Iteration Number: 1914
Loss: 58.94603569501678
l2 norm of gradients: 0.5848539066868095
l2 norm of weights: 6.256858864685713
-----
Iteration Number: 1915
Loss: 58.9291872791165
l2 norm of gradients: 0.5845848579926408
l2 norm of weights: 6.2567104234000075
-----
Iteration Number: 1916
Loss: 58.91235418371303
l2 norm of gradients: 0.584316084993515
l2 norm of weights: 6.2565620295791105
-----
Iteration Number: 1917
Loss: 58.89553638617321
l2 norm of gradients: 0.5840475873191184
l2 norm of weights: 6.2564136831764126
-----
Iteration Number: 1918
Loss: 58.878733863952405
l2 norm of gradients: 0.5837793645988986
l2 norm of weights: 6.256265384145387
-----
Iteration Number: 1919
Loss: 58.86194659462987
l2 norm of gradients: 0.583511416462069
l2 norm of weights: 6.25611713243959
-----
Iteration Number: 1920
Loss: 58.84517455566259
l2 norm of gradients: 0.5832437425376165
l2 norm of weights: 6.25596892801266
-----
Iteration Number: 1921
Loss: 58.82841772473166
l2 norm of gradients: 0.5829763424543067
l2 norm of weights: 6.255820770818319

```

```

-----
Iteration Number: 1922
Loss: 58.81167607938357
l2 norm of gradients: 0.5827092158406912
l2 norm of weights: 6.2556726608103705
-----
Iteration Number: 1923
Loss: 58.79494959733685
l2 norm of gradients: 0.582442362325112
l2 norm of weights: 6.255524597942697
-----
Iteration Number: 1924
Loss: 58.778238256264444
l2 norm of gradients: 0.5821757815357094
l2 norm of weights: 6.255376582169268
-----
Iteration Number: 1925
Loss: 58.76154203396704
l2 norm of gradients: 0.5819094731004268
l2 norm of weights: 6.255228613444131
-----
Iteration Number: 1926
Loss: 58.744860908210015
l2 norm of gradients: 0.5816434366470172
l2 norm of weights: 6.255080691721415
-----
Iteration Number: 1927
Loss: 58.72819485689634
l2 norm of gradients: 0.5813776718030493
l2 norm of weights: 6.25493281695533
-----
Iteration Number: 1928
Loss: 58.71154385782135
l2 norm of gradients: 0.5811121781959137
l2 norm of weights: 6.25478498910017
-----
Iteration Number: 1929
Loss: 58.6949078889409
l2 norm of gradients: 0.5808469554528276
l2 norm of weights: 6.254637208110305
-----
Iteration Number: 1930
Loss: 58.678286928221446
l2 norm of gradients: 0.5805820032008425
l2 norm of weights: 6.254489473940189
-----
Iteration Number: 1931
Loss: 58.66168095366859

```


l2 norm of gradients: 0.5803173210668489
l2 norm of weights: 6.254341786544354

Iteration Number: 1932
Loss: 58.64508994333962
l2 norm of gradients: 0.5800529086775822
l2 norm of weights: 6.254194145877417

Iteration Number: 1933
Loss: 58.62851387532944
l2 norm of gradients: 0.5797887656596293
l2 norm of weights: 6.254046551894069

Iteration Number: 1934
Loss: 58.611952727724486
l2 norm of gradients: 0.5795248916394335
l2 norm of weights: 6.253899004549083

Iteration Number: 1935
Loss: 58.595406478747314
l2 norm of gradients: 0.5792612862433011
l2 norm of weights: 6.2537515037973135

Iteration Number: 1936
Loss: 58.578875106640005
l2 norm of gradients: 0.578997949097406
l2 norm of weights: 6.253604049593692

Iteration Number: 1937
Loss: 58.562358589571396
l2 norm of gradients: 0.5787348798277969
l2 norm of weights: 6.253456641893233

Iteration Number: 1938
Loss: 58.54585690590843
l2 norm of gradients: 0.5784720780604019
l2 norm of weights: 6.253309280651022

Iteration Number: 1939
Loss: 58.529370033947
l2 norm of gradients: 0.5782095434210343
l2 norm of weights: 6.253161965822234

Iteration Number: 1940
Loss: 58.51289795207668
l2 norm of gradients: 0.5779472755353987
l2 norm of weights: 6.253014697362116

Iteration Number: 1941
Loss: 58.49644063871988
l2 norm of gradients: 0.5776852740290961
l2 norm of weights: 6.252867475225994

Iteration Number: 1942
Loss: 58.47999807234274
l2 norm of gradients: 0.5774235385276296
l2 norm of weights: 6.252720299369273

Iteration Number: 1943
Loss: 58.46357023144744
l2 norm of gradients: 0.5771620686564101
l2 norm of weights: 6.252573169747439

Iteration Number: 1944
Loss: 58.44715709458389
l2 norm of gradients: 0.5769008640407616
l2 norm of weights: 6.252426086316052

Iteration Number: 1945
Loss: 58.43075864029483
l2 norm of gradients: 0.5766399243059267
l2 norm of weights: 6.252279049030752

Iteration Number: 1946
Loss: 58.41437484728306
l2 norm of gradients: 0.5763792490770727
l2 norm of weights: 6.252132057847254

Iteration Number: 1947
Loss: 58.39800569416729
l2 norm of gradients: 0.5761188379792954
l2 norm of weights: 6.251985112721354

Iteration Number: 1948
Loss: 58.3816511596827
l2 norm of gradients: 0.5758586906376262
l2 norm of weights: 6.2518382136089246

Iteration Number: 1949
Loss: 58.36531122249598
l2 norm of gradients: 0.5755988066770364
l2 norm of weights: 6.251691360465912

Iteration Number: 1950
Loss: 58.34898586148455
l2 norm of gradients: 0.5753391857224436

```

12 norm of weights: 6.251544553248344
-----
Iteration Number: 1951
Loss: 58.33267505548468
12 norm of gradients: 0.5750798273987153
12 norm of weights: 6.251397791912322
-----
Iteration Number: 1952
Loss: 58.316378783278914
12 norm of gradients: 0.5748207313306758
12 norm of weights: 6.2512510764140234
-----
Iteration Number: 1953
Loss: 58.30009702386985
12 norm of gradients: 0.5745618971431106
12 norm of weights: 6.251104406709706
-----
Iteration Number: 1954
Loss: 58.28382975612837
12 norm of gradients: 0.5743033244607715
12 norm of weights: 6.250957782755699
-----
Iteration Number: 1955
Loss: 58.267576959109434
12 norm of gradients: 0.5740450129083827
12 norm of weights: 6.25081120450841
-----
Iteration Number: 1956
Loss: 58.25133861180869
12 norm of gradients: 0.5737869621106448
12 norm of weights: 6.250664671924322
-----
Iteration Number: 1957
Loss: 58.235114693323865
12 norm of gradients: 0.5735291716922408
12 norm of weights: 6.2505181849599944
-----
Iteration Number: 1958
Loss: 58.21890518272537
12 norm of gradients: 0.5732716412778402
12 norm of weights: 6.250371743572059
-----
Iteration Number: 1959
Loss: 58.202710059212194
12 norm of gradients: 0.5730143704921055
12 norm of weights: 6.250225347717228
-----
Iteration Number: 1960

```

```

Loss: 58.18652930194031
12 norm of gradients: 0.5727573589596956
12 norm of weights: 6.250078997352284
-----
Iteration Number: 1961
Loss: 58.17036289018751
12 norm of gradients: 0.5725006063052722
12 norm of weights: 6.249932692434086
-----
Iteration Number: 1962
Loss: 58.1542108031896
12 norm of gradients: 0.5722441121535039
12 norm of weights: 6.249786432919569
-----
Iteration Number: 1963
Loss: 58.138073020259014
12 norm of gradients: 0.5719878761290715
12 norm of weights: 6.249640218765739
-----
Iteration Number: 1964
Loss: 58.12194952077306
12 norm of gradients: 0.5717318978566727
12 norm of weights: 6.249494049929682
-----
Iteration Number: 1965
Loss: 58.105840284088565
12 norm of gradients: 0.571476176961027
12 norm of weights: 6.249347926368552
-----
Iteration Number: 1966
Loss: 58.089745289660115
12 norm of gradients: 0.5712207130668809
12 norm of weights: 6.249201848039582
-----
Iteration Number: 1967
Loss: 58.07366451696042
12 norm of gradients: 0.5709655057990124
12 norm of weights: 6.249055814900076
-----
Iteration Number: 1968
Loss: 58.0575979454666
12 norm of gradients: 0.5707105547822358
12 norm of weights: 6.24890982690741
-----
Iteration Number: 1969
Loss: 58.0415455547787
12 norm of gradients: 0.5704558596414062
12 norm of weights: 6.248763884019039

```

```

-----
Iteration Number: 1970
Loss: 58.02550732448494
l2 norm of gradients: 0.570201420001425
l2 norm of weights: 6.248617986192487
-----
Iteration Number: 1971
Loss: 58.009483234175256
l2 norm of gradients: 0.5699472354872441
l2 norm of weights: 6.24847213338535
-----
Iteration Number: 1972
Loss: 57.99347326352613
l2 norm of gradients: 0.5696933057238702
l2 norm of weights: 6.248326325555302
-----
Iteration Number: 1973
Loss: 57.977477392313666
l2 norm of gradients: 0.5694396303363702
l2 norm of weights: 6.248180562660086
-----
Iteration Number: 1974
Loss: 57.961495600204564
l2 norm of gradients: 0.5691862089498753
l2 norm of weights: 6.248034844657515
-----
Iteration Number: 1975
Loss: 57.94552786701301
l2 norm of gradients: 0.5689330411895859
l2 norm of weights: 6.247889171505481
-----
Iteration Number: 1976
Loss: 57.92957417252052
l2 norm of gradients: 0.5686801266807758
l2 norm of weights: 6.247743543161945
-----
Iteration Number: 1977
Loss: 57.913634496686385
l2 norm of gradients: 0.5684274650487973
l2 norm of weights: 6.247597959584937
-----
Iteration Number: 1978
Loss: 57.897708819373605
l2 norm of gradients: 0.5681750559190847
l2 norm of weights: 6.247452420732564
-----
Iteration Number: 1979
Loss: 57.881797120492124

```

l2 norm of gradients: 0.5679228989171597
l2 norm of weights: 6.2473069265630015

Iteration Number: 1980
Loss: 57.86589938006849
l2 norm of gradients: 0.5676709936686358
l2 norm of weights: 6.2471614770344965

Iteration Number: 1981
Loss: 57.85001557812228
l2 norm of gradients: 0.5674193397992221
l2 norm of weights: 6.24701607210537

Iteration Number: 1982
Loss: 57.83414569470615
l2 norm of gradients: 0.5671679369347279
l2 norm of weights: 6.246870711734012

Iteration Number: 1983
Loss: 57.81828970986894
l2 norm of gradients: 0.5669167847010675
l2 norm of weights: 6.246725395878881

Iteration Number: 1984
Loss: 57.80244760382756
l2 norm of gradients: 0.5666658827242639
l2 norm of weights: 6.246580124498513

Iteration Number: 1985
Loss: 57.78661935670897
l2 norm of gradients: 0.5664152306304536
l2 norm of weights: 6.246434897551508

Iteration Number: 1986
Loss: 57.770804948741706
l2 norm of gradients: 0.56616482804589
l2 norm of weights: 6.246289714996538

Iteration Number: 1987
Loss: 57.75500436013714
l2 norm of gradients: 0.5659146745969494
l2 norm of weights: 6.24614457679235

Iteration Number: 1988
Loss: 57.739217571269215
l2 norm of gradients: 0.565664769910133
l2 norm of weights: 6.245999482897756

Iteration Number: 1989
Loss: 57.723444562435084
l2 norm of gradients: 0.5654151136120728
l2 norm of weights: 6.245854433271637

Iteration Number: 1990
Loss: 57.70768531396676
l2 norm of gradients: 0.5651657053295346
l2 norm of weights: 6.245709427872949

Iteration Number: 1991
Loss: 57.691939806382145
l2 norm of gradients: 0.564916544689423
l2 norm of weights: 6.2455644666607135

Iteration Number: 1992
Loss: 57.676208019950124
l2 norm of gradients: 0.564667631318785
l2 norm of weights: 6.245419549594023

Iteration Number: 1993
Loss: 57.66048993527376
l2 norm of gradients: 0.5644189648448139
l2 norm of weights: 6.245274676632039

Iteration Number: 1994
Loss: 57.64478553287563
l2 norm of gradients: 0.564170544894854
l2 norm of weights: 6.24512984773399

Iteration Number: 1995
Loss: 57.62909479323478
l2 norm of gradients: 0.5639223710964042
l2 norm of weights: 6.244985062859178

Iteration Number: 1996
Loss: 57.61341769703356
l2 norm of gradients: 0.5636744430771214
l2 norm of weights: 6.2448403219669695

Iteration Number: 1997
Loss: 57.59775422491549
l2 norm of gradients: 0.563426760464826
l2 norm of weights: 6.2446956250168

Iteration Number: 1998
Loss: 57.582104357467124
l2 norm of gradients: 0.5631793228875037

```

12 norm of weights: 6.244550971968176
-----
Iteration Number: 1999
Loss: 57.566468075452775
12 norm of gradients: 0.5629321299733119
12 norm of weights: 6.244406362780672
-----
Iteration Number: 2000
Loss: 57.55084535959812
12 norm of gradients: 0.5626851813505807
12 norm of weights: 6.244261797413926
-----
Iteration Number: 2001
Loss: 57.53523619073403
12 norm of gradients: 0.5624384766478194
12 norm of weights: 6.244117275827648
-----
Iteration Number: 2002
Loss: 57.519640549628775
12 norm of gradients: 0.5621920154937188
12 norm of weights: 6.243972797981615
-----
Iteration Number: 2003
Loss: 57.50405841718877
12 norm of gradients: 0.5619457975171551
12 norm of weights: 6.2438283638356715
-----
Iteration Number: 2004
Loss: 57.48848977427689
12 norm of gradients: 0.5616998223471945
12 norm of weights: 6.243683973349729
-----
Iteration Number: 2005
Loss: 57.472934601868495
12 norm of gradients: 0.5614540896130958
12 norm of weights: 6.243539626483766
-----
Iteration Number: 2006
Loss: 57.457392880905736
12 norm of gradients: 0.561208598944315
12 norm of weights: 6.243395323197828
-----
Iteration Number: 2007
Loss: 57.44186459239839
12 norm of gradients: 0.5609633499705081
12 norm of weights: 6.243251063452028
-----
Iteration Number: 2008

```


Loss: 57.426349717440765
l2 norm of gradients: 0.5607183423215356
l2 norm of weights: 6.243106847206547

Iteration Number: 2009
Loss: 57.41084823706299
l2 norm of gradients: 0.5604735756274661
l2 norm of weights: 6.242962674421627

Iteration Number: 2010
Loss: 57.39536013238373
l2 norm of gradients: 0.5602290495185787
l2 norm of weights: 6.242818545057584

Iteration Number: 2011
Loss: 57.37988538459313
l2 norm of gradients: 0.5599847636253683
l2 norm of weights: 6.242674459074794

Iteration Number: 2012
Loss: 57.36442397489084
l2 norm of gradients: 0.5597407175785478
l2 norm of weights: 6.242530416433703

Iteration Number: 2013
Loss: 57.34897588446955
l2 norm of gradients: 0.5594969110090521
l2 norm of weights: 6.242386417094819

Iteration Number: 2014
Loss: 57.33354109464183
l2 norm of gradients: 0.5592533435480416
l2 norm of weights: 6.242242461018719

Iteration Number: 2015
Loss: 57.31811958672582
l2 norm of gradients: 0.5590100148269056
l2 norm of weights: 6.242098548166047

Iteration Number: 2016
Loss: 57.30271134201897
l2 norm of gradients: 0.558766924477266
l2 norm of weights: 6.241954678497506

Iteration Number: 2017
Loss: 57.28731634193128
l2 norm of gradients: 0.5585240721309799
l2 norm of weights: 6.24181085197387

```

-----
Iteration Number: 2018
Loss: 57.271934567840866
l2 norm of gradients: 0.5582814574201441
l2 norm of weights: 6.241667068555976
-----
Iteration Number: 2019
Loss: 57.256566001266165
l2 norm of gradients: 0.5580390799770978
l2 norm of weights: 6.241523328204724
-----
Iteration Number: 2020
Loss: 57.24121062361308
l2 norm of gradients: 0.5577969394344257
l2 norm of weights: 6.241379630881083
-----
Iteration Number: 2021
Loss: 57.225868416501896
l2 norm of gradients: 0.5575550354249617
l2 norm of weights: 6.241235976546082
-----
Iteration Number: 2022
Loss: 57.21053936140977
l2 norm of gradients: 0.5573133675817922
l2 norm of weights: 6.241092365160819
-----
Iteration Number: 2023
Loss: 57.1952234399821
l2 norm of gradients: 0.5570719355382592
l2 norm of weights: 6.240948796686452
-----
Iteration Number: 2024
Loss: 57.17992063385997
l2 norm of gradients: 0.5568307389279635
l2 norm of weights: 6.240805271084205
-----
Iteration Number: 2025
Loss: 57.164630924670725
l2 norm of gradients: 0.5565897773847681
l2 norm of weights: 6.240661788315365
-----
Iteration Number: 2026
Loss: 57.149354294190175
l2 norm of gradients: 0.556349050542801
l2 norm of weights: 6.2405183483412845
-----
Iteration Number: 2027
Loss: 57.1340907240841

```

l2 norm of gradients: 0.5561085580364589
l2 norm of weights: 6.240374951123378

Iteration Number: 2028
Loss: 57.118840196156164
l2 norm of gradients: 0.5558682995004095
l2 norm of weights: 6.240231596623123

Iteration Number: 2029
Loss: 57.10360269226149
l2 norm of gradients: 0.5556282745695954
l2 norm of weights: 6.240088284802062

Iteration Number: 2030
Loss: 57.088378194210776
l2 norm of gradients: 0.5553884828792368
l2 norm of weights: 6.2399450156217995

Iteration Number: 2031
Loss: 57.07316668386514
l2 norm of gradients: 0.5551489240648346
l2 norm of weights: 6.239801789044002

Iteration Number: 2032
Loss: 57.057968143218524
l2 norm of gradients: 0.5549095977621735
l2 norm of weights: 6.2396586050304

Iteration Number: 2033
Loss: 57.04278255415856
l2 norm of gradients: 0.5546705036073247
l2 norm of weights: 6.239515463542788

Iteration Number: 2034
Loss: 57.02760989877777
l2 norm of gradients: 0.5544316412366491
l2 norm of weights: 6.239372364543019

Iteration Number: 2035
Loss: 57.01245015895068
l2 norm of gradients: 0.5541930102868007
l2 norm of weights: 6.239229307993013

Iteration Number: 2036
Loss: 56.99730331688711
l2 norm of gradients: 0.5539546103947283
l2 norm of weights: 6.239086293854748

Iteration Number: 2037
Loss: 56.982169354601716
l2 norm of gradients: 0.5537164411976794
l2 norm of weights: 6.238943322090265

Iteration Number: 2038
Loss: 56.96704825424249
l2 norm of gradients: 0.5534785023332033
l2 norm of weights: 6.23880039266167

Iteration Number: 2039
Loss: 56.95193999797539
l2 norm of gradients: 0.5532407934391527
l2 norm of weights: 6.238657505531125

Iteration Number: 2040
Loss: 56.93684456805455
l2 norm of gradients: 0.553003314153688
l2 norm of weights: 6.23851466066086

Iteration Number: 2041
Loss: 56.92176194668007
l2 norm of gradients: 0.5527660641152784
l2 norm of weights: 6.238371858013159

Iteration Number: 2042
Loss: 56.90669211613053
l2 norm of gradients: 0.552529042962707
l2 norm of weights: 6.2382290975503745

Iteration Number: 2043
Loss: 56.89163505869925
l2 norm of gradients: 0.552292250335071
l2 norm of weights: 6.238086379234915

Iteration Number: 2044
Loss: 56.876590756775
l2 norm of gradients: 0.5520556858717863
l2 norm of weights: 6.23794370302925

Iteration Number: 2045
Loss: 56.86155919267938
l2 norm of gradients: 0.5518193492125892
l2 norm of weights: 6.237801068895914

Iteration Number: 2046
Loss: 56.84654034887451
l2 norm of gradients: 0.5515832399975399

```

12 norm of weights: 6.237658476797499
-----
Iteration Number: 2047
Loss: 56.83153420781403
12 norm of gradients: 0.5513473578670239
12 norm of weights: 6.2375159266966556
-----
Iteration Number: 2048
Loss: 56.8165407518985
12 norm of gradients: 0.5511117024617563
12 norm of weights: 6.237373418556098
-----
Iteration Number: 2049
Loss: 56.80155996377411
12 norm of gradients: 0.5508762734227826
12 norm of weights: 6.2372309523385985
-----
Iteration Number: 2050
Loss: 56.78659182587909
12 norm of gradients: 0.550641070391483
12 norm of weights: 6.237088528006989
-----
Iteration Number: 2051
Loss: 56.7716363208972
12 norm of gradients: 0.5504060930095734
12 norm of weights: 6.236946145524166
-----
Iteration Number: 2052
Loss: 56.75669343138135
12 norm of gradients: 0.5501713409191098
12 norm of weights: 6.236803804853076
-----
Iteration Number: 2053
Loss: 56.741763139963865
12 norm of gradients: 0.5499368137624884
12 norm of weights: 6.236661505956736
-----
Iteration Number: 2054
Loss: 56.72684542942485
12 norm of gradients: 0.5497025111824506
12 norm of weights: 6.236519248798214
-----
Iteration Number: 2055
Loss: 56.711940282408506
12 norm of gradients: 0.5494684328220834
12 norm of weights: 6.236377033340641
-----
Iteration Number: 2056

```

Loss: 56.697047681761
l2 norm of gradients: 0.5492345783248233
l2 norm of weights: 6.236234859547206

Iteration Number: 2057
Loss: 56.68216761015423
l2 norm of gradients: 0.5490009473344578
l2 norm of weights: 6.236092727381157

Iteration Number: 2058
Loss: 56.6673000505362
l2 norm of gradients: 0.5487675394951285
l2 norm of weights: 6.235950636805803

Iteration Number: 2059
Loss: 56.65244498567891
l2 norm of gradients: 0.5485343544513331
l2 norm of weights: 6.235808587784505

Iteration Number: 2060
Loss: 56.63760239851272
l2 norm of gradients: 0.5483013918479277
l2 norm of weights: 6.235666580280689

Iteration Number: 2061
Loss: 56.62277227199111
l2 norm of gradients: 0.5480686513301293
l2 norm of weights: 6.235524614257838

Iteration Number: 2062
Loss: 56.607954589002254
l2 norm of gradients: 0.5478361325435182
l2 norm of weights: 6.235382689679493

Iteration Number: 2063
Loss: 56.59314933261966
l2 norm of gradients: 0.5476038351340397
l2 norm of weights: 6.235240806509248

Iteration Number: 2064
Loss: 56.57835648580262
l2 norm of gradients: 0.5473717587480074
l2 norm of weights: 6.235098964710761

Iteration Number: 2065
Loss: 56.56357603169073
l2 norm of gradients: 0.5471399030321049
l2 norm of weights: 6.234957164247745

```

-----
Iteration Number: 2066
Loss: 56.54880795330724
l2 norm of gradients: 0.5469082676333876
l2 norm of weights: 6.234815405083973
-----
Iteration Number: 2067
Loss: 56.53405223384259
l2 norm of gradients: 0.5466768521992859
l2 norm of weights: 6.234673687183271
-----
Iteration Number: 2068
Loss: 56.51930885638336
l2 norm of gradients: 0.5464456563776061
l2 norm of weights: 6.234532010509526
-----
Iteration Number: 2069
Loss: 56.504577804192756
l2 norm of gradients: 0.5462146798165342
l2 norm of weights: 6.23439037502668
-----
Iteration Number: 2070
Loss: 56.489859060458784
l2 norm of gradients: 0.5459839221646363
l2 norm of weights: 6.234248780698733
-----
Iteration Number: 2071
Loss: 56.47515260848791
l2 norm of gradients: 0.5457533830708622
l2 norm of weights: 6.2341072274897416
-----
Iteration Number: 2072
Loss: 56.46045843153961
l2 norm of gradients: 0.5455230621845465
l2 norm of weights: 6.2339657153638175
-----
Iteration Number: 2073
Loss: 56.44577651295021
l2 norm of gradients: 0.5452929591554113
l2 norm of weights: 6.233824244285131
-----
Iteration Number: 2074
Loss: 56.43110683607289
l2 norm of gradients: 0.545063073633568
l2 norm of weights: 6.233682814217908
-----
Iteration Number: 2075
Loss: 56.41644938433743

```

l2 norm of gradients: 0.5448334052695188
l2 norm of weights: 6.233541425126431

Iteration Number: 2076
Loss: 56.40180414112741
l2 norm of gradients: 0.5446039537141603
l2 norm of weights: 6.233400076975036

Iteration Number: 2077
Loss: 56.3871710899302
l2 norm of gradients: 0.544374718618784
l2 norm of weights: 6.233258769728119

Iteration Number: 2078
Loss: 56.372550214197055
l2 norm of gradients: 0.5441456996350784
l2 norm of weights: 6.233117503350129

Iteration Number: 2079
Loss: 56.35794149749834
l2 norm of gradients: 0.5439168964151322
l2 norm of weights: 6.23297627780557

Iteration Number: 2080
Loss: 56.34334492335587
l2 norm of gradients: 0.5436883086114345
l2 norm of weights: 6.232835093059005

Iteration Number: 2081
Loss: 56.32876047540299
l2 norm of gradients: 0.5434599358768782
l2 norm of weights: 6.232693949075048

Iteration Number: 2082
Loss: 56.31418813722102
l2 norm of gradients: 0.5432317778647614
l2 norm of weights: 6.232552845818371

Iteration Number: 2083
Loss: 56.29962789249523
l2 norm of gradients: 0.5430038342287884
l2 norm of weights: 6.2324117832537

Iteration Number: 2084
Loss: 56.2850797248734
l2 norm of gradients: 0.5427761046230735
l2 norm of weights: 6.232270761345816

Iteration Number: 2085
Loss: 56.270543618133615
l2 norm of gradients: 0.5425485887021406
l2 norm of weights: 6.232129780059555

Iteration Number: 2086
Loss: 56.256019555958595
l2 norm of gradients: 0.5423212861209271
l2 norm of weights: 6.231988839359809

Iteration Number: 2087
Loss: 56.24150752216556
l2 norm of gradients: 0.5420941965347842
l2 norm of weights: 6.231847939211521

Iteration Number: 2088
Loss: 56.22700750058318
l2 norm of gradients: 0.5418673195994792
l2 norm of weights: 6.23170707957969

Iteration Number: 2089
Loss: 56.21251947501506
l2 norm of gradients: 0.5416406549711977
l2 norm of weights: 6.2315662604293705

Iteration Number: 2090
Loss: 56.198043429388726
l2 norm of gradients: 0.5414142023065447
l2 norm of weights: 6.23142548172567

Iteration Number: 2091
Loss: 56.18357934759169
l2 norm of gradients: 0.5411879612625464
l2 norm of weights: 6.231284743433749

Iteration Number: 2092
Loss: 56.16912721358141
l2 norm of gradients: 0.5409619314966526
l2 norm of weights: 6.231144045518823

Iteration Number: 2093
Loss: 56.15468701130739
l2 norm of gradients: 0.5407361126667374
l2 norm of weights: 6.231003387946162

Iteration Number: 2094
Loss: 56.14025872477714
l2 norm of gradients: 0.5405105044311019

```
12 norm of weights: 6.230862770681086
-----
Iteration Number: 2095
Loss: 56.12584233808146
12 norm of gradients: 0.5402851064484746
12 norm of weights: 6.230722193688972
-----
Iteration Number: 2096
Loss: 56.11143783520494
12 norm of gradients: 0.5400599183780146
12 norm of weights: 6.230581656935248
-----
Iteration Number: 2097
Loss: 56.09704520029564
12 norm of gradients: 0.5398349398793119
12 norm of weights: 6.2304411603853955
-----
Iteration Number: 2098
Loss: 56.082664417483755
12 norm of gradients: 0.5396101706123897
12 norm of weights: 6.230300704004952
-----
Iteration Number: 2099
Loss: 56.06829547096173
12 norm of gradients: 0.5393856102377055
12 norm of weights: 6.230160287759501
-----
Iteration Number: 2100
Loss: 56.05393834484445
12 norm of gradients: 0.5391612584161534
12 norm of weights: 6.2300199116146855
-----
Iteration Number: 2101
Loss: 56.03959302342383
12 norm of gradients: 0.5389371148090649
12 norm of weights: 6.229879575536197
-----
Iteration Number: 2102
Loss: 56.02525949091763
12 norm of gradients: 0.5387131790782108
12 norm of weights: 6.22973927948978
-----
Iteration Number: 2103
Loss: 56.01093773161604
12 norm of gradients: 0.538489450885803
12 norm of weights: 6.229599023441235
-----
Iteration Number: 2104
```

Loss: 55.99662772988228
l2 norm of gradients: 0.5382659298944953
l2 norm of weights: 6.229458807356408

Iteration Number: 2105
Loss: 55.98232947003403
l2 norm of gradients: 0.5380426157673849
l2 norm of weights: 6.229318631201201

Iteration Number: 2106
Loss: 55.968042936429114
l2 norm of gradients: 0.537819508168015
l2 norm of weights: 6.229178494941567

Iteration Number: 2107
Loss: 55.95376811352319
l2 norm of gradients: 0.5375966067603751
l2 norm of weights: 6.229038398543513

Iteration Number: 2108
Loss: 55.93950498574677
l2 norm of gradients: 0.537373911208903
l2 norm of weights: 6.2288983419730934

Iteration Number: 2109
Loss: 55.92525353752659
l2 norm of gradients: 0.5371514211784855
l2 norm of weights: 6.2287583251964165

Iteration Number: 2110
Loss: 55.911013753422424
l2 norm of gradients: 0.5369291363344609
l2 norm of weights: 6.228618348179642

Iteration Number: 2111
Loss: 55.89678561795859
l2 norm of gradients: 0.5367070563426192
l2 norm of weights: 6.22847841088898

Iteration Number: 2112
Loss: 55.882569115670094
l2 norm of gradients: 0.5364851808692048
l2 norm of weights: 6.2283385132906925

Iteration Number: 2113
Loss: 55.868364231156775
l2 norm of gradients: 0.5362635095809165
l2 norm of weights: 6.22819865535109

Iteration Number: 2114
Loss: 55.854170949099895
l2 norm of gradients: 0.5360420421449096
l2 norm of weights: 6.228058837036537

Iteration Number: 2115
Loss: 55.83998925408078
l2 norm of gradients: 0.5358207782287973
l2 norm of weights: 6.227919058313447

Iteration Number: 2116
Loss: 55.82581913080873
l2 norm of gradients: 0.5355997175006513
l2 norm of weights: 6.227779319148284

Iteration Number: 2117
Loss: 55.8116605640062
l2 norm of gradients: 0.5353788596290041
l2 norm of weights: 6.227639619507563

Iteration Number: 2118
Loss: 55.79751353841997
l2 norm of gradients: 0.5351582042828494
l2 norm of weights: 6.227499959357847

Iteration Number: 2119
Loss: 55.78337803879039
l2 norm of gradients: 0.5349377511316437
l2 norm of weights: 6.227360338665753

Iteration Number: 2120
Loss: 55.76925405000249
l2 norm of gradients: 0.5347174998453075
l2 norm of weights: 6.227220757397945

Iteration Number: 2121
Loss: 55.75514155683471
l2 norm of gradients: 0.5344974500942266
l2 norm of weights: 6.227081215521139

Iteration Number: 2122
Loss: 55.74104054412567
l2 norm of gradients: 0.5342776015492531
l2 norm of weights: 6.2269417130020965

Iteration Number: 2123
Loss: 55.72695099682297

l2 norm of gradients: 0.5340579538817072
l2 norm of weights: 6.226802249807633

Iteration Number: 2124
Loss: 55.7128728998194
l2 norm of gradients: 0.5338385067633771
l2 norm of weights: 6.226662825904612

Iteration Number: 2125
Loss: 55.6988062381136
l2 norm of gradients: 0.5336192598665218
l2 norm of weights: 6.226523441259948

Iteration Number: 2126
Loss: 55.684750996612415
l2 norm of gradients: 0.5334002128638711
l2 norm of weights: 6.2263840958406

Iteration Number: 2127
Loss: 55.6707071604085
l2 norm of gradients: 0.5331813654286267
l2 norm of weights: 6.22624478961358

Iteration Number: 2128
Loss: 55.656674714517656
l2 norm of gradients: 0.5329627172344644
l2 norm of weights: 6.226105522545949

Iteration Number: 2129
Loss: 55.642653643984914
l2 norm of gradients: 0.5327442679555338
l2 norm of weights: 6.225966294604815

Iteration Number: 2130
Loss: 55.62864393392029
l2 norm of gradients: 0.5325260172664606
l2 norm of weights: 6.225827105757335

Iteration Number: 2131
Loss: 55.61464556947773
l2 norm of gradients: 0.5323079648423468
l2 norm of weights: 6.225687955970717

Iteration Number: 2132
Loss: 55.600658535821964
l2 norm of gradients: 0.5320901103587723
l2 norm of weights: 6.225548845212211

Iteration Number: 2133
Loss: 55.58668281810436
l2 norm of gradients: 0.5318724534917953
l2 norm of weights: 6.225409773449125

Iteration Number: 2134
Loss: 55.572718401547434
l2 norm of gradients: 0.5316549939179546
l2 norm of weights: 6.225270740648807

Iteration Number: 2135
Loss: 55.55876527143859
l2 norm of gradients: 0.5314377313142692
l2 norm of weights: 6.225131746778657

Iteration Number: 2136
Loss: 55.54482341300496
l2 norm of gradients: 0.53122066535824
l2 norm of weights: 6.224992791806121

Iteration Number: 2137
Loss: 55.530892811586554
l2 norm of gradients: 0.5310037957278507
l2 norm of weights: 6.224853875698695

Iteration Number: 2138
Loss: 55.5169734525048
l2 norm of gradients: 0.530787122101569
l2 norm of weights: 6.22471499842392

Iteration Number: 2139
Loss: 55.50306532112508
l2 norm of gradients: 0.5305706441583471
l2 norm of weights: 6.224576159949387

Iteration Number: 2140
Loss: 55.48916840280262
l2 norm of gradients: 0.5303543615776227
l2 norm of weights: 6.224437360242733

Iteration Number: 2141
Loss: 55.475282683037626
l2 norm of gradients: 0.5301382740393208
l2 norm of weights: 6.224298599271643

Iteration Number: 2142
Loss: 55.461408147168775
l2 norm of gradients: 0.5299223812238532

```

12 norm of weights: 6.2241598770038475
-----
Iteration Number: 2143
Loss: 55.44754478080606
12 norm of gradients: 0.5297066828121205
12 norm of weights: 6.224021193407127
-----
Iteration Number: 2144
Loss: 55.43369256934916
12 norm of gradients: 0.5294911784855126
12 norm of weights: 6.223882548449308
-----
Iteration Number: 2145
Loss: 55.41985149834609
12 norm of gradients: 0.5292758679259095
12 norm of weights: 6.22374394209826
-----
Iteration Number: 2146
Loss: 55.406021553371275
12 norm of gradients: 0.5290607508156827
12 norm of weights: 6.223605374321906
-----
Iteration Number: 2147
Loss: 55.39220272004843
12 norm of gradients: 0.5288458268376953
12 norm of weights: 6.223466845088209
-----
Iteration Number: 2148
Loss: 55.37839498392845
12 norm of gradients: 0.5286310956753032
12 norm of weights: 6.223328354365183
-----
Iteration Number: 2149
Loss: 55.364598330734445
12 norm of gradients: 0.5284165570123561
12 norm of weights: 6.223189902120887
-----
Iteration Number: 2150
Loss: 55.35081274607137
12 norm of gradients: 0.528202210533198
12 norm of weights: 6.223051488323425
-----
Iteration Number: 2151
Loss: 55.337038215696104
12 norm of gradients: 0.5279880559226685
12 norm of weights: 6.222913112940948
-----
Iteration Number: 2152

```

Loss: 55.32327472530868
l2 norm of gradients: 0.5277740928661024
l2 norm of weights: 6.222774775941654

Iteration Number: 2153
Loss: 55.30952226065641
l2 norm of gradients: 0.5275603210493326
l2 norm of weights: 6.222636477293784

Iteration Number: 2154
Loss: 55.295780807557854
l2 norm of gradients: 0.5273467401586885
l2 norm of weights: 6.222498216965628

Iteration Number: 2155
Loss: 55.28205035179468
l2 norm of gradients: 0.5271333498809985
l2 norm of weights: 6.222359994925521

Iteration Number: 2156
Loss: 55.268330879216705
l2 norm of gradients: 0.5269201499035897
l2 norm of weights: 6.222221811141839

Iteration Number: 2157
Loss: 55.25462237568717
l2 norm of gradients: 0.5267071399142894
l2 norm of weights: 6.222083665583011

Iteration Number: 2158
Loss: 55.24092482717048
l2 norm of gradients: 0.5264943196014253
l2 norm of weights: 6.221945558217505

Iteration Number: 2159
Loss: 55.22723821951946
l2 norm of gradients: 0.5262816886538263
l2 norm of weights: 6.2218074890138375

Iteration Number: 2160
Loss: 55.213562538664114
l2 norm of gradients: 0.5260692467608235
l2 norm of weights: 6.221669457940569

Iteration Number: 2161
Loss: 55.199897770628084
l2 norm of gradients: 0.5258569936122501
l2 norm of weights: 6.221531464966303

Iteration Number: 2162
Loss: 55.1862439014251
l2 norm of gradients: 0.5256449288984437
l2 norm of weights: 6.221393510059691

Iteration Number: 2163
Loss: 55.172600917085674
l2 norm of gradients: 0.5254330523102446
l2 norm of weights: 6.221255593189428

Iteration Number: 2164
Loss: 55.15896880363761
l2 norm of gradients: 0.5252213635389988
l2 norm of weights: 6.221117714324253

Iteration Number: 2165
Loss: 55.14534754718919
l2 norm of gradients: 0.5250098622765569
l2 norm of weights: 6.220979873432948

Iteration Number: 2166
Loss: 55.13173713386266
l2 norm of gradients: 0.5247985482152759
l2 norm of weights: 6.220842070484344

Iteration Number: 2167
Loss: 55.11813754978252
l2 norm of gradients: 0.5245874210480189
l2 norm of weights: 6.2207043054473115

Iteration Number: 2168
Loss: 55.104548781123654
l2 norm of gradients: 0.5243764804681567
l2 norm of weights: 6.220566578290767

Iteration Number: 2169
Loss: 55.090970814079306
l2 norm of gradients: 0.5241657261695672
l2 norm of weights: 6.22042888898367

Iteration Number: 2170
Loss: 55.077403634882955
l2 norm of gradients: 0.5239551578466372
l2 norm of weights: 6.220291237495027

Iteration Number: 2171
Loss: 55.06384722974566

l2 norm of gradients: 0.5237447751942619
l2 norm of weights: 6.220153623793884

Iteration Number: 2172
Loss: 55.0503015850422
l2 norm of gradients: 0.5235345779078464
l2 norm of weights: 6.220016047849332

Iteration Number: 2173
Loss: 55.03676668696376
l2 norm of gradients: 0.5233245656833057
l2 norm of weights: 6.219878509630507

Iteration Number: 2174
Loss: 55.0232425219004
l2 norm of gradients: 0.523114738217065
l2 norm of weights: 6.219741009106589

Iteration Number: 2175
Loss: 55.0097290761743
l2 norm of gradients: 0.5229050952060614
l2 norm of weights: 6.219603546246797

Iteration Number: 2176
Loss: 54.99622633617249
l2 norm of gradients: 0.5226956363477426
l2 norm of weights: 6.219466121020397

Iteration Number: 2177
Loss: 54.98273428834295
l2 norm of gradients: 0.5224863613400693
l2 norm of weights: 6.219328733396697

Iteration Number: 2178
Loss: 54.96925291904941
l2 norm of gradients: 0.5222772698815146
l2 norm of weights: 6.2191913833450485

Iteration Number: 2179
Loss: 54.955782214851034
l2 norm of gradients: 0.5220683616710643
l2 norm of weights: 6.219054070834844

Iteration Number: 2180
Loss: 54.94232216216099
l2 norm of gradients: 0.5218596364082184
l2 norm of weights: 6.21891679583552

Iteration Number: 2181
Loss: 54.928872747478934
l2 norm of gradients: 0.5216510937929906
l2 norm of weights: 6.218779558316557

Iteration Number: 2182
Loss: 54.91543395742992
l2 norm of gradients: 0.5214427335259093
l2 norm of weights: 6.218642358247475

Iteration Number: 2183
Loss: 54.90200577849516
l2 norm of gradients: 0.5212345553080181
l2 norm of weights: 6.21850519559784

Iteration Number: 2184
Loss: 54.88858819728374
l2 norm of gradients: 0.5210265588408756
l2 norm of weights: 6.218368070337256

Iteration Number: 2185
Loss: 54.87518120044106
l2 norm of gradients: 0.5208187438265566
l2 norm of weights: 6.218230982435373

Iteration Number: 2186
Loss: 54.861784774644576
l2 norm of gradients: 0.5206111099676519
l2 norm of weights: 6.218093931861881

Iteration Number: 2187
Loss: 54.848398906447635
l2 norm of gradients: 0.5204036569672694
l2 norm of weights: 6.217956918586512

Iteration Number: 2188
Loss: 54.835023582645796
l2 norm of gradients: 0.520196384529034
l2 norm of weights: 6.21781994257904

Iteration Number: 2189
Loss: 54.82165878992285
l2 norm of gradients: 0.5199892923570882
l2 norm of weights: 6.217683003809284

Iteration Number: 2190
Loss: 54.80830451503656
l2 norm of gradients: 0.5197823801560919

```
12 norm of weights: 6.2175461022470975
-----
Iteration Number: 2191
Loss: 54.794960744713386
12 norm of gradients: 0.5195756476312238
12 norm of weights: 6.217409237862382
-----
Iteration Number: 2192
Loss: 54.7816274658381
12 norm of gradients: 0.5193690944881812
12 norm of weights: 6.217272410625078
-----
Iteration Number: 2193
Loss: 54.7683046651757
12 norm of gradients: 0.5191627204331801
12 norm of weights: 6.217135620505167
-----
Iteration Number: 2194
Loss: 54.754992329583935
12 norm of gradients: 0.5189565251729563
12 norm of weights: 6.216998867472673
-----
Iteration Number: 2195
Loss: 54.74169044591465
12 norm of gradients: 0.5187505084147649
12 norm of weights: 6.216862151497659
-----
Iteration Number: 2196
Loss: 54.72839900111501
12 norm of gradients: 0.5185446698663811
12 norm of weights: 6.21672547255023
-----
Iteration Number: 2197
Loss: 54.71511798206685
12 norm of gradients: 0.5183390092361008
12 norm of weights: 6.216588830600535
-----
Iteration Number: 2198
Loss: 54.701847375731546
12 norm of gradients: 0.5181335262327403
12 norm of weights: 6.216452225618757
-----
Iteration Number: 2199
Loss: 54.688587169083945
12 norm of gradients: 0.5179282205656369
12 norm of weights: 6.216315657575126
-----
Iteration Number: 2200
```

```
Loss: 54.675337349142474
l2 norm of gradients: 0.5177230919446488
l2 norm of weights: 6.216179126439909
-----
Iteration Number: 2201
Loss: 54.66209790290168
l2 norm of gradients: 0.517518140080157
l2 norm of weights: 6.216042632183415
-----
Iteration Number: 2202
Loss: 54.6488688174321
l2 norm of gradients: 0.5173133646830625
l2 norm of weights: 6.215906174775992
-----
Iteration Number: 2203
Loss: 54.635650079777115
l2 norm of gradients: 0.5171087654647902
l2 norm of weights: 6.215769754188031
-----
Iteration Number: 2204
Loss: 54.622441677108974
l2 norm of gradients: 0.5169043421372862
l2 norm of weights: 6.21563337038996
-----
Iteration Number: 2205
Loss: 54.609243596446795
l2 norm of gradients: 0.5167000944130198
l2 norm of weights: 6.215497023352248
-----
Iteration Number: 2206
Loss: 54.59605582503253
l2 norm of gradients: 0.5164960220049831
l2 norm of weights: 6.2153607130454045
-----
Iteration Number: 2207
Loss: 54.58287835000726
l2 norm of gradients: 0.5162921246266912
l2 norm of weights: 6.2152244394399805
-----
Iteration Number: 2208
Loss: 54.56971115852455
l2 norm of gradients: 0.5160884019921828
l2 norm of weights: 6.215088202506562
-----
Iteration Number: 2209
Loss: 54.55655423787102
l2 norm of gradients: 0.5158848538160197
l2 norm of weights: 6.214952002215779
```

```

-----
Iteration Number: 2210
Loss: 54.54340757526545
l2 norm of gradients: 0.5156814798132883
l2 norm of weights: 6.2148158385383
-----
Iteration Number: 2211
Loss: 54.530271157950565
l2 norm of gradients: 0.5154782796995984
l2 norm of weights: 6.2146797114448304
-----
Iteration Number: 2212
Loss: 54.51714497329505
l2 norm of gradients: 0.5152752531910842
l2 norm of weights: 6.2145436209061184
-----
Iteration Number: 2213
Loss: 54.504029008564665
l2 norm of gradients: 0.5150724000044041
l2 norm of weights: 6.21440756689295
-----
Iteration Number: 2214
Loss: 54.49092325111374
l2 norm of gradients: 0.5148697198567418
l2 norm of weights: 6.21427154937615
-----
Iteration Number: 2215
Loss: 54.4778276882892
l2 norm of gradients: 0.5146672124658048
l2 norm of weights: 6.214135568326583
-----
Iteration Number: 2216
Loss: 54.464742307525974
l2 norm of gradients: 0.5144648775498264
l2 norm of weights: 6.213999623715151
-----
Iteration Number: 2217
Loss: 54.45166709620719
l2 norm of gradients: 0.5142627148275644
l2 norm of weights: 6.213863715512796
-----
Iteration Number: 2218
Loss: 54.438602041780484
l2 norm of gradients: 0.5140607240183024
l2 norm of weights: 6.213727843690499
-----
Iteration Number: 2219
Loss: 54.425547131708335

```

l2 norm of gradients: 0.5138589048418487
l2 norm of weights: 6.21359200821928

Iteration Number: 2220
Loss: 54.412502353540404
l2 norm of gradients: 0.5136572570185381
l2 norm of weights: 6.213456209070195

Iteration Number: 2221
Loss: 54.399467694683175
l2 norm of gradients: 0.5134557802692306
l2 norm of weights: 6.2133204462143405

Iteration Number: 2222
Loss: 54.38644314273172
l2 norm of gradients: 0.5132544743153117
l2 norm of weights: 6.213184719622852

Iteration Number: 2223
Loss: 54.373428685247326
l2 norm of gradients: 0.5130533388786932
l2 norm of weights: 6.213049029266901

Iteration Number: 2224
Loss: 54.36042430982008
l2 norm of gradients: 0.5128523736818135
l2 norm of weights: 6.212913375117697

Iteration Number: 2225
Loss: 54.34743000401692
l2 norm of gradients: 0.512651578447636
l2 norm of weights: 6.212777757146493

Iteration Number: 2226
Loss: 54.33444575549766
l2 norm of gradients: 0.5124509528996513
l2 norm of weights: 6.2126421753245715

Iteration Number: 2227
Loss: 54.3214715519187
l2 norm of gradients: 0.512250496761876
l2 norm of weights: 6.21250662962326

Iteration Number: 2228
Loss: 54.30850738093688
l2 norm of gradients: 0.5120502097588532
l2 norm of weights: 6.212371120013918

Iteration Number: 2229
Loss: 54.29555323026375
l2 norm of gradients: 0.5118500916156525
l2 norm of weights: 6.2122356464679465

Iteration Number: 2230
Loss: 54.282609087663566
l2 norm of gradients: 0.5116501420578704
l2 norm of weights: 6.212100208956785

Iteration Number: 2231
Loss: 54.269674940810454
l2 norm of gradients: 0.5114503608116295
l2 norm of weights: 6.2119648074519045

Iteration Number: 2232
Loss: 54.256750777522534
l2 norm of gradients: 0.5112507476035796
l2 norm of weights: 6.211829441924821

Iteration Number: 2233
Loss: 54.243836585606175
l2 norm of gradients: 0.511051302160897
l2 norm of weights: 6.211694112347082

Iteration Number: 2234
Loss: 54.23093235281494
l2 norm of gradients: 0.5108520242112852
l2 norm of weights: 6.211558818690274

Iteration Number: 2235
Loss: 54.21803806704693
l2 norm of gradients: 0.510652913482974
l2 norm of weights: 6.211423560926021

Iteration Number: 2236
Loss: 54.20515371616429
l2 norm of gradients: 0.5104539697047209
l2 norm of weights: 6.211288339025985

Iteration Number: 2237
Loss: 54.19227928798789
l2 norm of gradients: 0.5102551926058096
l2 norm of weights: 6.211153152961861

Iteration Number: 2238
Loss: 54.17941477050694
l2 norm of gradients: 0.5100565819160511


```

12 norm of weights: 6.211018002705385
-----
Iteration Number: 2239
Loss: 54.16656015163491
12 norm of gradients: 0.5098581373657833
12 norm of weights: 6.210882888228328
-----
Iteration Number: 2240
Loss: 54.153715419328336
12 norm of gradients: 0.5096598586858713
12 norm of weights: 6.2107478095024975
-----
Iteration Number: 2241
Loss: 54.140880561513356
12 norm of gradients: 0.5094617456077069
12 norm of weights: 6.210612766499737
-----
Iteration Number: 2242
Loss: 54.12805556623695
12 norm of gradients: 0.5092637978632087
12 norm of weights: 6.210477759191926
-----
Iteration Number: 2243
Loss: 54.115240421517335
12 norm of gradients: 0.5090660151848233
12 norm of weights: 6.210342787550985
-----
Iteration Number: 2244
Loss: 54.10243511537358
12 norm of gradients: 0.508868397305523
12 norm of weights: 6.210207851548862
-----
Iteration Number: 2245
Loss: 54.0896396359254
12 norm of gradients: 0.5086709439588077
12 norm of weights: 6.21007295115755
-----
Iteration Number: 2246
Loss: 54.07685397119914
12 norm of gradients: 0.5084736548787043
12 norm of weights: 6.209938086349072
-----
Iteration Number: 2247
Loss: 54.06407810936857
12 norm of gradients: 0.5082765297997663
12 norm of weights: 6.209803257095491
-----
Iteration Number: 2248

```

Loss: 54.05131203852594
l2 norm of gradients: 0.5080795684570744
l2 norm of weights: 6.209668463368901

Iteration Number: 2249
Loss: 54.038555746840856
l2 norm of gradients: 0.5078827705862357
l2 norm of weights: 6.209533705141437

Iteration Number: 2250
Loss: 54.02580922249227
l2 norm of gradients: 0.5076861359233844
l2 norm of weights: 6.2093989823852676

Iteration Number: 2251
Loss: 54.013072453728014
l2 norm of gradients: 0.5074896642051813
l2 norm of weights: 6.209264295072595

Iteration Number: 2252
Loss: 54.00034542871086
l2 norm of gradients: 0.5072933551688138
l2 norm of weights: 6.209129643175659

Iteration Number: 2253
Loss: 53.98762813566648
l2 norm of gradients: 0.5070972085519958
l2 norm of weights: 6.208995026666736

Iteration Number: 2254
Loss: 53.97492056295066
l2 norm of gradients: 0.5069012240929683
l2 norm of weights: 6.208860445518134

Iteration Number: 2255
Loss: 53.96222269878039
l2 norm of gradients: 0.5067054015304981
l2 norm of weights: 6.208725899702199

Iteration Number: 2256
Loss: 53.94953453148378
l2 norm of gradients: 0.5065097406038788
l2 norm of weights: 6.208591389191313

Iteration Number: 2257
Loss: 53.93685604942503
l2 norm of gradients: 0.5063142410529295
l2 norm of weights: 6.208456913957889

```

-----
Iteration Number: 2258
Loss: 53.92418724094689
l2 norm of gradients: 0.506118902617997
l2 norm of weights: 6.208322473974378
-----
Iteration Number: 2259
Loss: 53.911528094394185
l2 norm of gradients: 0.5059237250399524
l2 norm of weights: 6.208188069213267
-----
Iteration Number: 2260
Loss: 53.898878598192205
l2 norm of gradients: 0.5057287080601942
l2 norm of weights: 6.208053699647073
-----
Iteration Number: 2261
Loss: 53.88623874076419
l2 norm of gradients: 0.5055338514206461
l2 norm of weights: 6.207919365248353
-----
Iteration Number: 2262
Loss: 53.873608510555044
l2 norm of gradients: 0.505339154863758
l2 norm of weights: 6.207785065989697
-----
Iteration Number: 2263
Loss: 53.86098789603156
l2 norm of gradients: 0.5051446181325047
l2 norm of weights: 6.207650801843726
-----
Iteration Number: 2264
Loss: 53.84837688570323
l2 norm of gradients: 0.5049502409703875
l2 norm of weights: 6.2075165727831
-----
Iteration Number: 2265
Loss: 53.83577546800871
l2 norm of gradients: 0.5047560231214324
l2 norm of weights: 6.207382378780511
-----
Iteration Number: 2266
Loss: 53.82318363149386
l2 norm of gradients: 0.5045619643301913
l2 norm of weights: 6.207248219808687
-----
Iteration Number: 2267
Loss: 53.81060136475515

```

l2 norm of gradients: 0.5043680643417405
l2 norm of weights: 6.207114095840388

Iteration Number: 2268
Loss: 53.79802865632836
l2 norm of gradients: 0.5041743229016821
l2 norm of weights: 6.206980006848408

Iteration Number: 2269
Loss: 53.78546549481055
l2 norm of gradients: 0.5039807397561423
l2 norm of weights: 6.206845952805577

Iteration Number: 2270
Loss: 53.7729118688201
l2 norm of gradients: 0.5037873146517727
l2 norm of weights: 6.206711933684758

Iteration Number: 2271
Loss: 53.760367767007956
l2 norm of gradients: 0.5035940473357491
l2 norm of weights: 6.206577949458848

Iteration Number: 2272
Loss: 53.74783317800353
l2 norm of gradients: 0.5034009375557715
l2 norm of weights: 6.206444000100777

Iteration Number: 2273
Loss: 53.73530809050501
l2 norm of gradients: 0.5032079850600648
l2 norm of weights: 6.206310085583508

Iteration Number: 2274
Loss: 53.72279249321111
l2 norm of gradients: 0.5030151895973772
l2 norm of weights: 6.20617620588004

Iteration Number: 2275
Loss: 53.710286374812135
l2 norm of gradients: 0.5028225509169812
l2 norm of weights: 6.206042360963404

Iteration Number: 2276
Loss: 53.697789724048675
l2 norm of gradients: 0.5026300687686733
l2 norm of weights: 6.205908550806665

Iteration Number: 2277
Loss: 53.68530252973886
l2 norm of gradients: 0.5024377429027728
l2 norm of weights: 6.205774775382918

Iteration Number: 2278
Loss: 53.67282478060103
l2 norm of gradients: 0.5022455730701229
l2 norm of weights: 6.205641034665298

Iteration Number: 2279
Loss: 53.66035646545324
l2 norm of gradients: 0.5020535590220897
l2 norm of weights: 6.205507328626966

Iteration Number: 2280
Loss: 53.64789757312491
l2 norm of gradients: 0.5018617005105624
l2 norm of weights: 6.205373657241121

Iteration Number: 2281
Loss: 53.63544809246399
l2 norm of gradients: 0.5016699972879531
l2 norm of weights: 6.205240020480993

Iteration Number: 2282
Loss: 53.62300801233964
l2 norm of gradients: 0.501478449107196
l2 norm of weights: 6.205106418319843

Iteration Number: 2283
Loss: 53.61057732165535
l2 norm of gradients: 0.501287055721748
l2 norm of weights: 6.20497285073097

Iteration Number: 2284
Loss: 53.598156009245784
l2 norm of gradients: 0.5010958168855879
l2 norm of weights: 6.2048393176877

Iteration Number: 2285
Loss: 53.58574406410158
l2 norm of gradients: 0.5009047323532169
l2 norm of weights: 6.2047058191633955

Iteration Number: 2286
Loss: 53.57334147512854
l2 norm of gradients: 0.5007138018796572

```

12 norm of weights: 6.204572355131451
-----
Iteration Number: 2287
Loss: 53.56094823134029
12 norm of gradients: 0.500523025220453
12 norm of weights: 6.204438925565292
-----
Iteration Number: 2288
Loss: 53.54856432168479
12 norm of gradients: 0.5003324021316695
12 norm of weights: 6.204305530438375
-----
Iteration Number: 2289
Loss: 53.53618973519497
12 norm of gradients: 0.500141932369893
12 norm of weights: 6.2041721697241945
-----
Iteration Number: 2290
Loss: 53.523824460866074
12 norm of gradients: 0.4999516156922305
12 norm of weights: 6.204038843396272
-----
Iteration Number: 2291
Loss: 53.51146848776587
12 norm of gradients: 0.49976145185630955
12 norm of weights: 6.2039055514281625
-----
Iteration Number: 2292
Loss: 53.499121804961284
12 norm of gradients: 0.4995714406202781
12 norm of weights: 6.203772293793454
-----
Iteration Number: 2293
Loss: 53.486784401513944
12 norm of gradients: 0.49938158174280384
12 norm of weights: 6.2036390704657665
-----
Iteration Number: 2294
Loss: 53.4744562665724
12 norm of gradients: 0.49919187498307455
12 norm of weights: 6.203505881418751
-----
Iteration Number: 2295
Loss: 53.4621373892774
12 norm of gradients: 0.4990023201007973
12 norm of weights: 6.203372726626091
-----
Iteration Number: 2296

```

```
Loss: 53.44982775871695
12 norm of gradients: 0.4988129168561986
12 norm of weights: 6.2032396060615
-----
Iteration Number: 2297
Loss: 53.437527364071684
12 norm of gradients: 0.4986236650100238
12 norm of weights: 6.203106519698726
-----
Iteration Number: 2298
Loss: 53.42523619457286
12 norm of gradients: 0.49843456432353717
12 norm of weights: 6.202973467511547
-----
Iteration Number: 2299
Loss: 53.412954239393635
12 norm of gradients: 0.4982456145585211
12 norm of weights: 6.202840449473772
-----
Iteration Number: 2300
Loss: 53.40068148773177
12 norm of gradients: 0.4980568154772765
12 norm of weights: 6.202707465559245
-----
Iteration Number: 2301
Loss: 53.38841792892183
12 norm of gradients: 0.4978681668426219
12 norm of weights: 6.2025745157418335
-----
Iteration Number: 2302
Loss: 53.376163552164186
12 norm of gradients: 0.49767966841789363
12 norm of weights: 6.202441599995446
-----
Iteration Number: 2303
Loss: 53.36391834673044
12 norm of gradients: 0.4974913199669454
12 norm of weights: 6.202308718294015
-----
Iteration Number: 2304
Loss: 53.351682301961596
12 norm of gradients: 0.4973031212541477
12 norm of weights: 6.202175870611508
-----
Iteration Number: 2305
Loss: 53.33945540715707
12 norm of gradients: 0.49711507204438776
12 norm of weights: 6.202043056921922
```

```

-----
Iteration Number: 2306
Loss: 53.32723765168713
l2 norm of gradients: 0.4969271721030696
l2 norm of weights: 6.201910277199285
-----
Iteration Number: 2307
Loss: 53.31502902491355
l2 norm of gradients: 0.49673942119611286
l2 norm of weights: 6.2017775314176555
-----
Iteration Number: 2308
Loss: 53.302829516179266
l2 norm of gradients: 0.4965518190899536
l2 norm of weights: 6.201644819551125
-----
Iteration Number: 2309
Loss: 53.29063911491314
l2 norm of gradients: 0.4963643655515429
l2 norm of weights: 6.201512141573812
-----
Iteration Number: 2310
Loss: 53.27845781056969
l2 norm of gradients: 0.49617706034834724
l2 norm of weights: 6.20137949745987
-----
Iteration Number: 2311
Loss: 53.26628559252614
l2 norm of gradients: 0.4959899032483479
l2 norm of weights: 6.201246887183482
-----
Iteration Number: 2312
Loss: 53.25412245027201
l2 norm of gradients: 0.4958028940200408
l2 norm of weights: 6.2011143107188555
-----
Iteration Number: 2313
Loss: 53.24196837331882
l2 norm of gradients: 0.49561603243243607
l2 norm of weights: 6.200981768040238
-----
Iteration Number: 2314
Loss: 53.22982335108708
l2 norm of gradients: 0.49542931825505804
l2 norm of weights: 6.200849259121902
-----
Iteration Number: 2315
Loss: 53.21768737315853

```



```

12 norm of gradients: 0.49524275125794404
12 norm of weights: 6.200716783938151
-----
Iteration Number: 2316
Loss: 53.2055604290285
12 norm of gradients: 0.4950563312116451
12 norm of weights: 6.200584342463317
-----
Iteration Number: 2317
Loss: 53.19344250825666
12 norm of gradients: 0.494870057887225
12 norm of weights: 6.200451934671768
-----
Iteration Number: 2318
Loss: 53.18133360047693
12 norm of gradients: 0.49468393105626
12 norm of weights: 6.200319560537894
-----
Iteration Number: 2319
Loss: 53.16923369520799
12 norm of gradients: 0.49449795049083894
12 norm of weights: 6.2001872200361206
-----
Iteration Number: 2320
Loss: 53.15714278205412
12 norm of gradients: 0.4943121159635623
12 norm of weights: 6.200054913140902
-----
Iteration Number: 2321
Loss: 53.14506085070955
12 norm of gradients: 0.49412642724754197
12 norm of weights: 6.199922639826723
-----
Iteration Number: 2322
Loss: 53.13298789076163
12 norm of gradients: 0.49394088411640114
12 norm of weights: 6.199790400068096
-----
Iteration Number: 2323
Loss: 53.120923891911666
12 norm of gradients: 0.49375548634427385
12 norm of weights: 6.199658193839564
-----
Iteration Number: 2324
Loss: 53.10886884386797
12 norm of gradients: 0.4935702337058047
12 norm of weights: 6.199526021115702
-----

```

Iteration Number: 2325
Loss: 53.09682273627109
l2 norm of gradients: 0.4933851259761481
l2 norm of weights: 6.19939388187111

Iteration Number: 2326
Loss: 53.08478555885623
l2 norm of gradients: 0.4932001629309684
l2 norm of weights: 6.199261776080423

Iteration Number: 2327
Loss: 53.072757301409176
l2 norm of gradients: 0.4930153443464393
l2 norm of weights: 6.1991297037183

Iteration Number: 2328
Loss: 53.06073795365182
l2 norm of gradients: 0.4928306699992433
l2 norm of weights: 6.198997664759433

Iteration Number: 2329
Loss: 53.048727505337204
l2 norm of gradients: 0.4926461396665717
l2 norm of weights: 6.1988656591785425

Iteration Number: 2330
Loss: 53.03672594634226
l2 norm of gradients: 0.4924617531261241
l2 norm of weights: 6.198733686950377

Iteration Number: 2331
Loss: 53.02473326641837
l2 norm of gradients: 0.4922775101561076
l2 norm of weights: 6.1986017480497155

Iteration Number: 2332
Loss: 53.01274945539008
l2 norm of gradients: 0.4920934105352372
l2 norm of weights: 6.198469842451366

Iteration Number: 2333
Loss: 53.000774503137805
l2 norm of gradients: 0.4919094540427345
l2 norm of weights: 6.198337970130163

Iteration Number: 2334
Loss: 52.98880839953526
l2 norm of gradients: 0.4917256404583282

```

12 norm of weights: 6.198206131060975
-----
Iteration Number: 2335
Loss: 52.976851134436345
12 norm of gradients: 0.4915419695622532
12 norm of weights: 6.198074325218695
-----
Iteration Number: 2336
Loss: 52.96490269777122
12 norm of gradients: 0.4913584411352499
12 norm of weights: 6.197942552578246
-----
Iteration Number: 2337
Loss: 52.952963079458215
12 norm of gradients: 0.4911750549585646
12 norm of weights: 6.197810813114579
-----
Iteration Number: 2338
Loss: 52.94103226939802
12 norm of gradients: 0.4909918108139487
12 norm of weights: 6.197679106802676
-----
Iteration Number: 2339
Loss: 52.92911025761202
12 norm of gradients: 0.49080870848365804
12 norm of weights: 6.1975474336175465
-----
Iteration Number: 2340
Loss: 52.91719703405086
12 norm of gradients: 0.49062574775045265
12 norm of weights: 6.197415793534226
-----
Iteration Number: 2341
Loss: 52.9052925886928
12 norm of gradients: 0.49044292839759673
12 norm of weights: 6.197284186527782
-----
Iteration Number: 2342
Loss: 52.893396911550056
12 norm of gradients: 0.4902602502088578
12 norm of weights: 6.197152612573309
-----
Iteration Number: 2343
Loss: 52.881509992678716
12 norm of gradients: 0.4900777129685063
12 norm of weights: 6.19702107164593
-----
Iteration Number: 2344

```

```

Loss: 52.86963182208842
12 norm of gradients: 0.48989531646131534
12 norm of weights: 6.196889563720795
-----
Iteration Number: 2345
Loss: 52.85776238988292
12 norm of gradients: 0.48971306047255997
12 norm of weights: 6.196758088773085
-----
Iteration Number: 2346
Loss: 52.84590168611586
12 norm of gradients: 0.48953094478801745
12 norm of weights: 6.196626646778004
-----
Iteration Number: 2347
Loss: 52.83404970090716
12 norm of gradients: 0.4893489691939661
12 norm of weights: 6.19649523771079
-----
Iteration Number: 2348
Loss: 52.822206424374556
12 norm of gradients: 0.48916713347718493
12 norm of weights: 6.196363861546707
-----
Iteration Number: 2349
Loss: 52.81037184660361
12 norm of gradients: 0.48898543742495376
12 norm of weights: 6.196232518261044
-----
Iteration Number: 2350
Loss: 52.79854595778463
12 norm of gradients: 0.488803880825052
12 norm of weights: 6.19610120782912
-----
Iteration Number: 2351
Loss: 52.78672874809456
12 norm of gradients: 0.4886224634657591
12 norm of weights: 6.195969930226283
-----
Iteration Number: 2352
Loss: 52.77492020769957
12 norm of gradients: 0.48844118513585316
12 norm of weights: 6.1958386854279075
-----
Iteration Number: 2353
Loss: 52.763120326846305
12 norm of gradients: 0.4882600456246112
12 norm of weights: 6.195707473409395

```

```

-----
Iteration Number: 2354
Loss: 52.75132909567314
l2 norm of gradients: 0.4880790447218086
l2 norm of weights: 6.195576294146176
-----
Iteration Number: 2355
Loss: 52.739546504482846
l2 norm of gradients: 0.48789818221771825
l2 norm of weights: 6.195445147613707
-----
Iteration Number: 2356
Loss: 52.727772543541846
l2 norm of gradients: 0.48771745790311016
l2 norm of weights: 6.195314033787475
-----
Iteration Number: 2357
Loss: 52.71600720305325
l2 norm of gradients: 0.48753687156925185
l2 norm of weights: 6.195182952642988
-----
Iteration Number: 2358
Loss: 52.704250473399355
l2 norm of gradients: 0.4873564230079066
l2 norm of weights: 6.195051904155789
-----
Iteration Number: 2359
Loss: 52.69250234478334
l2 norm of gradients: 0.48717611201133404
l2 norm of weights: 6.194920888301442
-----
Iteration Number: 2360
Loss: 52.680762807583406
l2 norm of gradients: 0.48699593837228894
l2 norm of weights: 6.194789905055542
-----
Iteration Number: 2361
Loss: 52.66903185214815
l2 norm of gradients: 0.48681590188402135
l2 norm of weights: 6.19465895439371
-----
Iteration Number: 2362
Loss: 52.65730946883128
l2 norm of gradients: 0.4866360023402755
l2 norm of weights: 6.1945280362915955
-----
Iteration Number: 2363
Loss: 52.6455956479964

```

```

12 norm of gradients: 0.4864562395352901
12 norm of weights: 6.194397150724871
-----
Iteration Number: 2364
Loss: 52.63389038003579
12 norm of gradients: 0.4862766132637969
12 norm of weights: 6.19426629766924
-----
Iteration Number: 2365
Loss: 52.622193655345846
12 norm of gradients: 0.48609712332102123
12 norm of weights: 6.19413547710043
-----
Iteration Number: 2366
Loss: 52.61050546435241
12 norm of gradients: 0.4859177695026809
12 norm of weights: 6.194004688994199
-----
Iteration Number: 2367
Loss: 52.598825797503196
12 norm of gradients: 0.4857385516049855
12 norm of weights: 6.193873933326327
-----
Iteration Number: 2368
Loss: 52.58715464528409
12 norm of gradients: 0.4855594694246369
12 norm of weights: 6.193743210072625
-----
Iteration Number: 2369
Loss: 52.57549199812969
12 norm of gradients: 0.4853805227588273
12 norm of weights: 6.193612519208926
-----
Iteration Number: 2370
Loss: 52.56383784653118
12 norm of gradients: 0.48520171140524027
12 norm of weights: 6.193481860711095
-----
Iteration Number: 2371
Loss: 52.5521921810149
12 norm of gradients: 0.4850230351620491
12 norm of weights: 6.193351234555019
-----
Iteration Number: 2372
Loss: 52.54055499207594
12 norm of gradients: 0.48484449382791694
12 norm of weights: 6.193220640716614
-----

```

Iteration Number: 2373
Loss: 52.52892627028676
l2 norm of gradients: 0.48466608720199605
l2 norm of weights: 6.193090079171822

Iteration Number: 2374
Loss: 52.517306006196485
l2 norm of gradients: 0.48448781508392735
l2 norm of weights: 6.192959549896609

Iteration Number: 2375
Loss: 52.50569419037724
l2 norm of gradients: 0.4843096772738397
l2 norm of weights: 6.192829052866971

Iteration Number: 2376
Loss: 52.494090813370605
l2 norm of gradients: 0.4841316735723501
l2 norm of weights: 6.192698588058927

Iteration Number: 2377
Loss: 52.48249586585039
l2 norm of gradients: 0.483953803780562
l2 norm of weights: 6.192568155448525

Iteration Number: 2378
Loss: 52.47090933841179
l2 norm of gradients: 0.4837760677000661
l2 norm of weights: 6.192437755011836

Iteration Number: 2379
Loss: 52.45933122169106
l2 norm of gradients: 0.48359846513293864
l2 norm of weights: 6.19230738672496

Iteration Number: 2380
Loss: 52.44776150633282
l2 norm of gradients: 0.48342099588174203
l2 norm of weights: 6.192177050564019

Iteration Number: 2381
Loss: 52.436200183008914
l2 norm of gradients: 0.4832436597495232
l2 norm of weights: 6.192046746505165

Iteration Number: 2382
Loss: 52.42464724241754
l2 norm of gradients: 0.4830664565398139

```

12 norm of weights: 6.191916474524575
-----
Iteration Number: 2383
Loss: 52.413102675241284
12 norm of gradients: 0.4828893860566296
12 norm of weights: 6.19178623459845
-----
Iteration Number: 2384
Loss: 52.401566472197075
12 norm of gradients: 0.4827124481044699
12 norm of weights: 6.191656026703017
-----
Iteration Number: 2385
Loss: 52.390038624066584
12 norm of gradients: 0.48253564248831654
12 norm of weights: 6.191525850814531
-----
Iteration Number: 2386
Loss: 52.37851912150799
12 norm of gradients: 0.4823589690136341
12 norm of weights: 6.191395706909269
-----
Iteration Number: 2387
Loss: 52.36700795536928
12 norm of gradients: 0.4821824274863691
12 norm of weights: 6.191265594963537
-----
Iteration Number: 2388
Loss: 52.35550511640557
12 norm of gradients: 0.4820060177129492
12 norm of weights: 6.191135514953663
-----
Iteration Number: 2389
Loss: 52.34401059537875
12 norm of gradients: 0.4818297395002829
12 norm of weights: 6.191005466856004
-----
Iteration Number: 2390
Loss: 52.33252438319084
12 norm of gradients: 0.4816535926557589
12 norm of weights: 6.190875450646941
-----
Iteration Number: 2391
Loss: 52.321046470542754
12 norm of gradients: 0.48147757698724614
12 norm of weights: 6.190745466302879
-----
Iteration Number: 2392

```



```
Loss: 52.30957684838951
l2 norm of gradients: 0.4813016923030918
l2 norm of weights: 6.19061551380025
-----
Iteration Number: 2393
Loss: 52.298115507498814
l2 norm of gradients: 0.4811259384121223
l2 norm of weights: 6.190485593115509
-----
Iteration Number: 2394
Loss: 52.28666243881867
l2 norm of gradients: 0.48095031512364206
l2 norm of weights: 6.190355704225139
-----
Iteration Number: 2395
Loss: 52.27521763323663
l2 norm of gradients: 0.4807748222474328
l2 norm of weights: 6.190225847105647
-----
Iteration Number: 2396
Loss: 52.26378108161044
l2 norm of gradients: 0.4805994595937533
l2 norm of weights: 6.190096021733564
-----
Iteration Number: 2397
Loss: 52.2523527748748
l2 norm of gradients: 0.48042422697333875
l2 norm of weights: 6.189966228085447
-----
Iteration Number: 2398
Loss: 52.24093270397805
l2 norm of gradients: 0.48024912419739996
l2 norm of weights: 6.189836466137876
-----
Iteration Number: 2399
Loss: 52.22952085987454
l2 norm of gradients: 0.48007415107762347
l2 norm of weights: 6.189706735867459
-----
Iteration Number: 2400
Loss: 52.21811723353419
l2 norm of gradients: 0.47989930742616993
l2 norm of weights: 6.189577037250827
-----
Iteration Number: 2401
Loss: 52.20672181593553
l2 norm of gradients: 0.47972459305567455
l2 norm of weights: 6.189447370264635
```

```
-----  
Iteration Number: 2402  
Loss: 52.19533459803602  
l2 norm of gradients: 0.4795500077792459  
l2 norm of weights: 6.189317734885565  
-----  
Iteration Number: 2403  
Loss: 52.18395557090588  
l2 norm of gradients: 0.47937555141046534  
l2 norm of weights: 6.189188131090321  
-----  
Iteration Number: 2404  
Loss: 52.172584725542315  
l2 norm of gradients: 0.47920122376338714  
l2 norm of weights: 6.189058558855635  
-----  
Iteration Number: 2405  
Loss: 52.16122205302706  
l2 norm of gradients: 0.479027024652537  
l2 norm of weights: 6.188929018158258  
-----  
Iteration Number: 2406  
Loss: 52.14986754436672  
l2 norm of gradients: 0.4788529538929116  
l2 norm of weights: 6.188799508974971  
-----  
Iteration Number: 2407  
Loss: 52.138521190641626  
l2 norm of gradients: 0.47867901129997875  
l2 norm of weights: 6.188670031282577  
-----  
Iteration Number: 2408  
Loss: 52.12718298296079  
l2 norm of gradients: 0.47850519668967645  
l2 norm of weights: 6.188540585057903  
-----  
Iteration Number: 2409  
Loss: 52.11585291242756  
l2 norm of gradients: 0.47833150987841155  
l2 norm of weights: 6.188411170277801  
-----  
Iteration Number: 2410  
Loss: 52.104530970131776  
l2 norm of gradients: 0.4781579506830604  
l2 norm of weights: 6.1882817869191475  
-----  
Iteration Number: 2411  
Loss: 52.093217147259686
```

```

12 norm of gradients: 0.4779845189209673
12 norm of weights: 6.188152434958844
-----
Iteration Number: 2412
Loss: 52.08191143489538
12 norm of gradients: 0.47781121440994423
12 norm of weights: 6.188023114373812
-----
Iteration Number: 2413
Loss: 52.07061382425696
12 norm of gradients: 0.47763803696827034
12 norm of weights: 6.187893825141002
-----
Iteration Number: 2414
Loss: 52.05932430645329
12 norm of gradients: 0.4774649864146917
12 norm of weights: 6.1877645672373855
-----
Iteration Number: 2415
Loss: 52.04804287275173
12 norm of gradients: 0.4772920625684193
12 norm of weights: 6.187635340639961
-----
Iteration Number: 2416
Loss: 52.0367695143405
12 norm of gradients: 0.47711926524913034
12 norm of weights: 6.187506145325747
-----
Iteration Number: 2417
Loss: 52.025504222384086
12 norm of gradients: 0.4769465942769665
12 norm of weights: 6.18737698127179
-----
Iteration Number: 2418
Loss: 52.014246988198934
12 norm of gradients: 0.4767740494725332
12 norm of weights: 6.187247848455155
-----
Iteration Number: 2419
Loss: 52.00299780297711
12 norm of gradients: 0.4766016306568997
12 norm of weights: 6.187118746852937
-----
Iteration Number: 2420
Loss: 51.99175665798474
12 norm of gradients: 0.4764293376515978
12 norm of weights: 6.18698967644225
-----

```

Iteration Number: 2421
Loss: 51.9805235445385
l2 norm of gradients: 0.47625717027862197
l2 norm of weights: 6.186860637200234

Iteration Number: 2422
Loss: 51.969298453916196
l2 norm of gradients: 0.4760851283604279
l2 norm of weights: 6.186731629104052

Iteration Number: 2423
Loss: 51.95808137743632
l2 norm of gradients: 0.4759132117199323
l2 norm of weights: 6.18660265213089

Iteration Number: 2424
Loss: 51.94687230637107
l2 norm of gradients: 0.4757414201805125
l2 norm of weights: 6.186473706257959

Iteration Number: 2425
Loss: 51.93567123212466
l2 norm of gradients: 0.47556975356600556
l2 norm of weights: 6.186344791462492

Iteration Number: 2426
Loss: 51.92447814599124
l2 norm of gradients: 0.4753982117007074
l2 norm of weights: 6.186215907721746

Iteration Number: 2427
Loss: 51.91329303941885
l2 norm of gradients: 0.47522679440937277
l2 norm of weights: 6.186087055013

Iteration Number: 2428
Loss: 51.90211590369026
l2 norm of gradients: 0.475055501517214
l2 norm of weights: 6.18595823331356

Iteration Number: 2429
Loss: 51.89094673022724
l2 norm of gradients: 0.47488433284990095
l2 norm of weights: 6.185829442600751

Iteration Number: 2430
Loss: 51.879785510467926
l2 norm of gradients: 0.4747132882335598

```

12 norm of weights: 6.185700682851924
-----
Iteration Number: 2431
Loss: 51.86863223583829
12 norm of gradients: 0.47454236749477285
12 norm of weights: 6.185571954044453
-----
Iteration Number: 2432
Loss: 51.857486897732244
12 norm of gradients: 0.4743715704605778
12 norm of weights: 6.1854432561557315
-----
Iteration Number: 2433
Loss: 51.846349487635095
12 norm of gradients: 0.47420089695846673
12 norm of weights: 6.185314589163182
-----
Iteration Number: 2434
Loss: 51.83521999699759
12 norm of gradients: 0.4740303468163862
12 norm of weights: 6.185185953044246
-----
Iteration Number: 2435
Loss: 51.824098417300135
12 norm of gradients: 0.4738599198627355
12 norm of weights: 6.185057347776388
-----
Iteration Number: 2436
Loss: 51.812984740010265
12 norm of gradients: 0.4736896159263676
12 norm of weights: 6.184928773337099
-----
Iteration Number: 2437
Loss: 51.80187895671549
12 norm of gradients: 0.4735194348365868
12 norm of weights: 6.184800229703885
-----
Iteration Number: 2438
Loss: 51.79078105884131
12 norm of gradients: 0.4733493764231489
12 norm of weights: 6.184671716854286
-----
Iteration Number: 2439
Loss: 51.77969103794981
12 norm of gradients: 0.47317944051626104
12 norm of weights: 6.184543234765854
-----
Iteration Number: 2440

```

```

Loss: 51.768608885614654
l2 norm of gradients: 0.4730096269465799
l2 norm of weights: 6.184414783416172
-----
Iteration Number: 2441
Loss: 51.757534593373286
l2 norm of gradients: 0.4728399355452119
l2 norm of weights: 6.18428636278284
-----
Iteration Number: 2442
Loss: 51.74646815285054
l2 norm of gradients: 0.47267036614371205
l2 norm of weights: 6.184157972843485
-----
Iteration Number: 2443
Loss: 51.73540955554684
l2 norm of gradients: 0.47250091857408394
l2 norm of weights: 6.184029613575752
-----
Iteration Number: 2444
Loss: 51.72435879314167
l2 norm of gradients: 0.47233159266877817
l2 norm of weights: 6.183901284957312
-----
Iteration Number: 2445
Loss: 51.71331585719838
l2 norm of gradients: 0.47216238826069246
l2 norm of weights: 6.183772986965857
-----
Iteration Number: 2446
Loss: 51.702280739379525
l2 norm of gradients: 0.4719933051831705
l2 norm of weights: 6.183644719579102
-----
Iteration Number: 2447
Loss: 51.69125343134091
l2 norm of gradients: 0.4718243432700017
l2 norm of weights: 6.183516482774785
-----
Iteration Number: 2448
Loss: 51.68023392469168
l2 norm of gradients: 0.47165550235542003
l2 norm of weights: 6.183388276530665
-----
Iteration Number: 2449
Loss: 51.669222211142674
l2 norm of gradients: 0.4714867822741036
l2 norm of weights: 6.183260100824524

```

```

-----
Iteration Number: 2450
Loss: 51.65821828235652
l2 norm of gradients: 0.47131818286117394
l2 norm of weights: 6.183131955634165
-----
Iteration Number: 2451
Loss: 51.647222130028176
l2 norm of gradients: 0.47114970395219574
l2 norm of weights: 6.183003840937417
-----
Iteration Number: 2452
Loss: 51.636233745897215
l2 norm of gradients: 0.4709813453831751
l2 norm of weights: 6.182875756712126
-----
Iteration Number: 2453
Loss: 51.625253121676195
l2 norm of gradients: 0.47081310699056006
l2 norm of weights: 6.182747702936164
-----
Iteration Number: 2454
Loss: 51.614280249069964
l2 norm of gradients: 0.47064498861123916
l2 norm of weights: 6.1826196795874235
-----
Iteration Number: 2455
Loss: 51.60331511986598
l2 norm of gradients: 0.4704769900825412
l2 norm of weights: 6.182491686643819
-----
Iteration Number: 2456
Loss: 51.59235772583517
l2 norm of gradients: 0.47030911124223396
l2 norm of weights: 6.182363724083288
-----
Iteration Number: 2457
Loss: 51.5814080587023
l2 norm of gradients: 0.4701413519285239
l2 norm of weights: 6.182235791883787
-----
Iteration Number: 2458
Loss: 51.570466110312296
l2 norm of gradients: 0.4699737119800559
l2 norm of weights: 6.182107890023299
-----
Iteration Number: 2459
Loss: 51.55953187243175

```

```

12 norm of gradients: 0.4698061912359115
12 norm of weights: 6.181980018479825
-----
Iteration Number: 2460
Loss: 51.548605336918634
12 norm of gradients: 0.46963878953560917
12 norm of weights: 6.181852177231389
-----
Iteration Number: 2461
Loss: 51.53768649555092
12 norm of gradients: 0.4694715067191031
12 norm of weights: 6.181724366256038
-----
Iteration Number: 2462
Loss: 51.526775340216325
12 norm of gradients: 0.4693043426267828
12 norm of weights: 6.181596585531839
-----
Iteration Number: 2463
Loss: 51.515871862722335
12 norm of gradients: 0.46913729709947205
12 norm of weights: 6.181468835036882
-----
Iteration Number: 2464
Loss: 51.50497605500349
12 norm of gradients: 0.46897036997842845
12 norm of weights: 6.181341114749277
-----
Iteration Number: 2465
Loss: 51.49408790887633
12 norm of gradients: 0.4688035611053425
12 norm of weights: 6.181213424647157
-----
Iteration Number: 2466
Loss: 51.48320741626089
12 norm of gradients: 0.4686368703223374
12 norm of weights: 6.181085764708675
-----
Iteration Number: 2467
Loss: 51.472334569070796
12 norm of gradients: 0.4684702974719677
12 norm of weights: 6.180958134912009
-----
Iteration Number: 2468
Loss: 51.46146935923827
12 norm of gradients: 0.46830384239721873
12 norm of weights: 6.180830535235354
-----

```


Iteration Number: 2469
Loss: 51.45061177866777
l2 norm of gradients: 0.46813750494150624
l2 norm of weights: 6.180702965656929

Iteration Number: 2470
Loss: 51.43976181931285
l2 norm of gradients: 0.4679712849486756
l2 norm of weights: 6.180575426154973

Iteration Number: 2471
Loss: 51.42891947314688
l2 norm of gradients: 0.46780518226300055
l2 norm of weights: 6.180447916707748

Iteration Number: 2472
Loss: 51.4180847321194
l2 norm of gradients: 0.467639196729183
l2 norm of weights: 6.1803204372935365

Iteration Number: 2473
Loss: 51.40725758822667
l2 norm of gradients: 0.4674733281923525
l2 norm of weights: 6.180192987890642

Iteration Number: 2474
Loss: 51.39643803343884
l2 norm of gradients: 0.4673075764980646
l2 norm of weights: 6.1800655684773895

Iteration Number: 2475
Loss: 51.3856260598305
l2 norm of gradients: 0.4671419414923012
l2 norm of weights: 6.179938179032124

Iteration Number: 2476
Loss: 51.374821659350765
l2 norm of gradients: 0.4669764230214691
l2 norm of weights: 6.1798108195332135

Iteration Number: 2477
Loss: 51.36402482409651
l2 norm of gradients: 0.46681102093239957
l2 norm of weights: 6.179683489959046

Iteration Number: 2478
Loss: 51.35323554604677
l2 norm of gradients: 0.4666457350723476

```

12 norm of weights: 6.1795561902880305
-----
Iteration Number: 2479
Loss: 51.34245381731446
12 norm of gradients: 0.4664805652889909
12 norm of weights: 6.179428920498596
-----
Iteration Number: 2480
Loss: 51.331679629936964
12 norm of gradients: 0.4663155114304297
12 norm of weights: 6.179301680569196
-----
Iteration Number: 2481
Loss: 51.32091297602828
12 norm of gradients: 0.46615057334518545
12 norm of weights: 6.179174470478301
-----
Iteration Number: 2482
Loss: 51.31015384767483
12 norm of gradients: 0.46598575088220046
12 norm of weights: 6.179047290204406
-----
Iteration Number: 2483
Loss: 51.299402236951586
12 norm of gradients: 0.46582104389083706
12 norm of weights: 6.178920139726021
-----
Iteration Number: 2484
Loss: 51.288658136023365
12 norm of gradients: 0.46565645222087654
12 norm of weights: 6.178793019021683
-----
Iteration Number: 2485
Loss: 51.27792153703472
12 norm of gradients: 0.4654919757225188
12 norm of weights: 6.178665928069947
-----
Iteration Number: 2486
Loss: 51.267192432077906
12 norm of gradients: 0.4653276142463818
12 norm of weights: 6.178538866849388
-----
Iteration Number: 2487
Loss: 51.25647081332487
12 norm of gradients: 0.46516336764350025
12 norm of weights: 6.178411835338604
-----
Iteration Number: 2488

```

Loss: 51.24575667295172
l2 norm of gradients: 0.46499923576532504
l2 norm of weights: 6.178284833516211

Iteration Number: 2489
Loss: 51.23505000317269
l2 norm of gradients: 0.4648352184637226
l2 norm of weights: 6.178157861360848

Iteration Number: 2490
Loss: 51.224350796170825
l2 norm of gradients: 0.46467131559097413
l2 norm of weights: 6.178030918851173

Iteration Number: 2491
Loss: 51.21365904409237
l2 norm of gradients: 0.4645075269997748
l2 norm of weights: 6.177904005965864

Iteration Number: 2492
Loss: 51.20297473920516
l2 norm of gradients: 0.4643438525432333
l2 norm of weights: 6.1777771226836204

Iteration Number: 2493
Loss: 51.192297873748444
l2 norm of gradients: 0.46418029207487027
l2 norm of weights: 6.1776502689831645

Iteration Number: 2494
Loss: 51.181628439973764
l2 norm of gradients: 0.46401684544861826
l2 norm of weights: 6.177523444843233

Iteration Number: 2495
Loss: 51.17096643008373
l2 norm of gradients: 0.46385351251882123
l2 norm of weights: 6.177396650242588

Iteration Number: 2496
Loss: 51.16031183635785
l2 norm of gradients: 0.46369029314023247
l2 norm of weights: 6.17726988516001

Iteration Number: 2497
Loss: 51.14966465111517
l2 norm of gradients: 0.4635271871680154
l2 norm of weights: 6.177143149574301

```

-----
Iteration Number: 2498
Loss: 51.13902486658571
l2 norm of gradients: 0.4633641944577418
l2 norm of weights: 6.177016443464282
-----
Iteration Number: 2499
Loss: 51.12839247514711
l2 norm of gradients: 0.46320131486539123
l2 norm of weights: 6.176889766808793
-----
Iteration Number: 2500
Loss: 51.117767469027115
l2 norm of gradients: 0.4630385482473508
l2 norm of weights: 6.176763119586699
-----
Iteration Number: 2501
Loss: 51.10714984062702
l2 norm of gradients: 0.4628758944604134
l2 norm of weights: 6.176636501776879
-----
Iteration Number: 2502
Loss: 51.09653958223244
l2 norm of gradients: 0.46271335336177777
l2 norm of weights: 6.176509913358235
-----
Iteration Number: 2503
Loss: 51.08593668619381
l2 norm of gradients: 0.4625509248090476
l2 norm of weights: 6.17638335430969
-----
Iteration Number: 2504
Loss: 51.07534114491252
l2 norm of gradients: 0.46238860866023007
l2 norm of weights: 6.176256824610187
-----
Iteration Number: 2505
Loss: 51.064752950742296
l2 norm of gradients: 0.4622264047737363
l2 norm of weights: 6.176130324238685
-----
Iteration Number: 2506
Loss: 51.05417209602653
l2 norm of gradients: 0.4620643130083794
l2 norm of weights: 6.1760038531741674
-----
Iteration Number: 2507
Loss: 51.043598573216315

```

```

12 norm of gradients: 0.461902333223374
12 norm of weights: 6.1758774113956365
-----
Iteration Number: 2508
Loss: 51.03303237471096
12 norm of gradients: 0.4617404652783361
12 norm of weights: 6.175750998882114
-----
Iteration Number: 2509
Loss: 51.02247349290012
12 norm of gradients: 0.4615787090332816
12 norm of weights: 6.175624615612641
-----
Iteration Number: 2510
Loss: 51.01192192020549
12 norm of gradients: 0.46141706434862545
12 norm of weights: 6.175498261566278
-----
Iteration Number: 2511
Loss: 51.00137764909701
12 norm of gradients: 0.46125553108518164
12 norm of weights: 6.175371936722105
-----
Iteration Number: 2512
Loss: 50.99084067203468
12 norm of gradients: 0.4610941091041613
12 norm of weights: 6.175245641059227
-----
Iteration Number: 2513
Loss: 50.98031098146651
12 norm of gradients: 0.46093279826717315
12 norm of weights: 6.175119374556761
-----
Iteration Number: 2514
Loss: 50.9697885698762
12 norm of gradients: 0.46077159843622145
12 norm of weights: 6.174993137193848
-----
Iteration Number: 2515
Loss: 50.95927342973938
12 norm of gradients: 0.4606105094737063
12 norm of weights: 6.174866928949648
-----
Iteration Number: 2516
Loss: 50.94876555355104
12 norm of gradients: 0.460449531242422
12 norm of weights: 6.174740749803339
-----

```

Iteration Number: 2517
Loss: 50.938264933842504
l2 norm of gradients: 0.4602886636055569
l2 norm of weights: 6.174614599734121

Iteration Number: 2518
Loss: 50.92777156310841
l2 norm of gradients: 0.46012790642669205
l2 norm of weights: 6.174488478721211

Iteration Number: 2519
Loss: 50.917285433901846
l2 norm of gradients: 0.45996725956980106
l2 norm of weights: 6.17436238674385

Iteration Number: 2520
Loss: 50.90680653873275
l2 norm of gradients: 0.45980672289924845
l2 norm of weights: 6.174236323781291

Iteration Number: 2521
Loss: 50.89633487019536
l2 norm of gradients: 0.4596462962797894
l2 norm of weights: 6.1741102898128135

Iteration Number: 2522
Loss: 50.88587042086902
l2 norm of gradients: 0.4594859795765692
l2 norm of weights: 6.173984284817712

Iteration Number: 2523
Loss: 50.875413183274965
l2 norm of gradients: 0.4593257726551216
l2 norm of weights: 6.173858308775303

Iteration Number: 2524
Loss: 50.86496315002157
l2 norm of gradients: 0.4591656753813688
l2 norm of weights: 6.1737323616649205

Iteration Number: 2525
Loss: 50.854520313705756
l2 norm of gradients: 0.45900568762162053
l2 norm of weights: 6.173606443465918

Iteration Number: 2526
Loss: 50.84408466695846
l2 norm of gradients: 0.45884580924257223

```
12 norm of weights: 6.173480554157669
-----
Iteration Number: 2527
Loss: 50.83365620240821
12 norm of gradients: 0.458686040111306
12 norm of weights: 6.173354693719567
-----
Iteration Number: 2528
Loss: 50.82323491263966
12 norm of gradients: 0.4585263800952882
12 norm of weights: 6.173228862131022
-----
Iteration Number: 2529
Loss: 50.81282079032643
12 norm of gradients: 0.4583668290623695
12 norm of weights: 6.1731030593714635
-----
Iteration Number: 2530
Loss: 50.802413828100676
12 norm of gradients: 0.45820738688078394
12 norm of weights: 6.172977285420345
-----
Iteration Number: 2531
Loss: 50.7920140186756
12 norm of gradients: 0.4580480534191479
12 norm of weights: 6.172851540257131
-----
Iteration Number: 2532
Loss: 50.781621354665596
12 norm of gradients: 0.4578888285464593
12 norm of weights: 6.172725823861311
-----
Iteration Number: 2533
Loss: 50.77123582877743
12 norm of gradients: 0.457729712132097
12 norm of weights: 6.1726001362123935
-----
Iteration Number: 2534
Loss: 50.76085743374509
12 norm of gradients: 0.45757070404581973
12 norm of weights: 6.172474477289904
-----
Iteration Number: 2535
Loss: 50.75048616225821
12 norm of gradients: 0.4574118041577656
12 norm of weights: 6.172348847073384
-----
Iteration Number: 2536
```

```
Loss: 50.74012200698445
l2 norm of gradients: 0.45725301233845067
l2 norm of weights: 6.1722232455423995
-----
Iteration Number: 2537
Loss: 50.72976496072388
l2 norm of gradients: 0.457094328458769
l2 norm of weights: 6.172097672676534
-----
Iteration Number: 2538
Loss: 50.71941501616917
l2 norm of gradients: 0.45693575238999107
l2 norm of weights: 6.171972128455386
-----
Iteration Number: 2539
Loss: 50.709072166068445
l2 norm of gradients: 0.4567772840037631
l2 norm of weights: 6.171846612858578
-----
Iteration Number: 2540
Loss: 50.698736403261485
l2 norm of gradients: 0.4566189231721068
l2 norm of weights: 6.171721125865748
-----
Iteration Number: 2541
Loss: 50.6884077204036
l2 norm of gradients: 0.4564606697674174
l2 norm of weights: 6.171595667456554
-----
Iteration Number: 2542
Loss: 50.67808611036022
l2 norm of gradients: 0.4563025236624642
l2 norm of weights: 6.171470237610672
-----
Iteration Number: 2543
Loss: 50.66777156589123
l2 norm of gradients: 0.4561444847303887
l2 norm of weights: 6.171344836307797
-----
Iteration Number: 2544
Loss: 50.657464079821864
l2 norm of gradients: 0.45598655284470396
l2 norm of weights: 6.171219463527642
-----
Iteration Number: 2545
Loss: 50.64716364494203
l2 norm of gradients: 0.45582872787929424
l2 norm of weights: 6.171094119249941
```



```

-----
Iteration Number: 2546
Loss: 50.63687025405977
l2 norm of gradients: 0.45567100970841357
l2 norm of weights: 6.170968803454444
-----
Iteration Number: 2547
Loss: 50.62658390007132
l2 norm of gradients: 0.4555133982066855
l2 norm of weights: 6.170843516120921
-----
Iteration Number: 2548
Loss: 50.61630457578858
l2 norm of gradients: 0.45535589324910153
l2 norm of weights: 6.170718257229158
-----
Iteration Number: 2549
Loss: 50.60603227403624
l2 norm of gradients: 0.4551984947110209
l2 norm of weights: 6.170593026758963
-----
Iteration Number: 2550
Loss: 50.595766987726186
l2 norm of gradients: 0.45504120246816965
l2 norm of weights: 6.170467824690161
-----
Iteration Number: 2551
Loss: 50.58550870971475
l2 norm of gradients: 0.4548840163966393
l2 norm of weights: 6.170342651002596
-----
Iteration Number: 2552
Loss: 50.57525743288362
l2 norm of gradients: 0.45472693637288675
l2 norm of weights: 6.170217505676128
-----
Iteration Number: 2553
Loss: 50.565013150183766
l2 norm of gradients: 0.4545699622737329
l2 norm of weights: 6.170092388690638
-----
Iteration Number: 2554
Loss: 50.554775854435576
l2 norm of gradients: 0.45441309397636187
l2 norm of weights: 6.169967300026025
-----
Iteration Number: 2555
Loss: 50.5445455386058

```

```

12 norm of gradients: 0.45425633135832033
12 norm of weights: 6.169842239662205
-----
Iteration Number: 2556
Loss: 50.53432219559023
12 norm of gradients: 0.45409967429751663
12 norm of weights: 6.169717207579113
-----
Iteration Number: 2557
Loss: 50.524105818414206
12 norm of gradients: 0.45394312267221965
12 norm of weights: 6.169592203756703
-----
Iteration Number: 2558
Loss: 50.5138963998999
12 norm of gradients: 0.4537866763610585
12 norm of weights: 6.169467228174947
-----
Iteration Number: 2559
Loss: 50.5036939330905
12 norm of gradients: 0.4536303352430211
12 norm of weights: 6.169342280813832
-----
Iteration Number: 2560
Loss: 50.49349841095495
12 norm of gradients: 0.45347409919745374
12 norm of weights: 6.16921736165337
-----
Iteration Number: 2561
Loss: 50.483309826427174
12 norm of gradients: 0.45331796810406005
12 norm of weights: 6.169092470673583
-----
Iteration Number: 2562
Loss: 50.473128172534146
12 norm of gradients: 0.4531619418429
12 norm of weights: 6.168967607854519
-----
Iteration Number: 2563
Loss: 50.462953442231075
12 norm of gradients: 0.4530060202943894
12 norm of weights: 6.1688427731762365
-----
Iteration Number: 2564
Loss: 50.45278562855631
12 norm of gradients: 0.452850203339299
12 norm of weights: 6.168717966618818
-----

```

Iteration Number: 2565
Loss: 50.44262472453715
l2 norm of gradients: 0.452694490858753
l2 norm of weights: 6.168593188162362

Iteration Number: 2566
Loss: 50.432470723190214
l2 norm of gradients: 0.4525388827342294
l2 norm of weights: 6.168468437786984

Iteration Number: 2567
Loss: 50.42232361755005
l2 norm of gradients: 0.45238337884755764
l2 norm of weights: 6.168343715472818

Iteration Number: 2568
Loss: 50.412183400657455
l2 norm of gradients: 0.4522279790809192
l2 norm of weights: 6.1682190212000165

Iteration Number: 2569
Loss: 50.402050065608826
l2 norm of gradients: 0.4520726833168455
l2 norm of weights: 6.16809435494875

Iteration Number: 2570
Loss: 50.391923605421184
l2 norm of gradients: 0.4519174914382184
l2 norm of weights: 6.167969716699206

Iteration Number: 2571
Loss: 50.381804013214456
l2 norm of gradients: 0.45176240332826756
l2 norm of weights: 6.167845106431591

Iteration Number: 2572
Loss: 50.37169128202954
l2 norm of gradients: 0.4516074188705714
l2 norm of weights: 6.167720524126128

Iteration Number: 2573
Loss: 50.36158540502559
l2 norm of gradients: 0.4514525379490549
l2 norm of weights: 6.167595969763059

Iteration Number: 2574
Loss: 50.35148637526935
l2 norm of gradients: 0.4512977604479896

```

12 norm of weights: 6.167471443322643
-----
Iteration Number: 2575
Loss: 50.34139418585988
12 norm of gradients: 0.451143086251992
12 norm of weights: 6.167346944785156
-----
Iteration Number: 2576
Loss: 50.331308829957436
12 norm of gradients: 0.45098851524602335
12 norm of weights: 6.167222474130895
-----
Iteration Number: 2577
Loss: 50.321230300679595
12 norm of gradients: 0.45083404731538856
12 norm of weights: 6.16709803134017
-----
Iteration Number: 2578
Loss: 50.311158591190264
12 norm of gradients: 0.4506796823457352
12 norm of weights: 6.166973616393312
-----
Iteration Number: 2579
Loss: 50.30109369464658
12 norm of gradients: 0.4505254202230525
12 norm of weights: 6.16684922927067
-----
Iteration Number: 2580
Loss: 50.291035604203095
12 norm of gradients: 0.45037126083367085
12 norm of weights: 6.166724869952607
-----
Iteration Number: 2581
Loss: 50.280984313006726
12 norm of gradients: 0.45021720406426097
12 norm of weights: 6.166600538419508
-----
Iteration Number: 2582
Loss: 50.270939814290365
12 norm of gradients: 0.4500632498018325
12 norm of weights: 6.166476234651772
-----
Iteration Number: 2583
Loss: 50.26090210122032
12 norm of gradients: 0.44990939793373375
12 norm of weights: 6.166351958629818
-----
Iteration Number: 2584

```

Loss: 50.25087116698784
l2 norm of gradients: 0.4497556483476505
l2 norm of weights: 6.166227710334081

Iteration Number: 2585
Loss: 50.24084700484182
l2 norm of gradients: 0.4496020009316051
l2 norm of weights: 6.166103489745012

Iteration Number: 2586
Loss: 50.23082960798898
l2 norm of gradients: 0.4494484555739558
l2 norm of weights: 6.165979296843085

Iteration Number: 2587
Loss: 50.220818969636504
l2 norm of gradients: 0.4492950121633955
l2 norm of weights: 6.165855131608786

Iteration Number: 2588
Loss: 50.21081508307436
l2 norm of gradients: 0.44914167058895155
l2 norm of weights: 6.165730994022621

Iteration Number: 2589
Loss: 50.200817941494215
l2 norm of gradients: 0.4489884307399841
l2 norm of weights: 6.16560688406511

Iteration Number: 2590
Loss: 50.19082753818766
l2 norm of gradients: 0.4488352925061858
l2 norm of weights: 6.165482801716798

Iteration Number: 2591
Loss: 50.180843866447276
l2 norm of gradients: 0.4486822557775806
l2 norm of weights: 6.165358746958237

Iteration Number: 2592
Loss: 50.170866919518616
l2 norm of gradients: 0.4485293204445229
l2 norm of weights: 6.165234719770005

Iteration Number: 2593
Loss: 50.160896690681824
l2 norm of gradients: 0.4483764863976971
l2 norm of weights: 6.165110720132693

```

-----
Iteration Number: 2594
Loss: 50.1509331732362
l2 norm of gradients: 0.4482237535281161
l2 norm of weights: 6.1649867480269105
-----
Iteration Number: 2595
Loss: 50.14097636053593
l2 norm of gradients: 0.44807112172712066
l2 norm of weights: 6.164862803433282
-----
Iteration Number: 2596
Loss: 50.131026245841454
l2 norm of gradients: 0.44791859088637875
l2 norm of weights: 6.1647388863324535
-----
Iteration Number: 2597
Loss: 50.12108282248787
l2 norm of gradients: 0.4477661608978846
l2 norm of weights: 6.164614996705084
-----
Iteration Number: 2598
Loss: 50.111146083820046
l2 norm of gradients: 0.4476138316539573
l2 norm of weights: 6.164491134531852
-----
Iteration Number: 2599
Loss: 50.101216023169854
l2 norm of gradients: 0.44746160304724053
l2 norm of weights: 6.164367299793454
-----
Iteration Number: 2600
Loss: 50.091292633912694
l2 norm of gradients: 0.44730947497070167
l2 norm of weights: 6.164243492470599
-----
Iteration Number: 2601
Loss: 50.08137590938221
l2 norm of gradients: 0.44715744731763046
l2 norm of weights: 6.164119712544019
-----
Iteration Number: 2602
Loss: 50.071465842969346
l2 norm of gradients: 0.4470055199816385
l2 norm of weights: 6.163995959994459
-----
Iteration Number: 2603
Loss: 50.06156242801839

```

```

12 norm of gradients: 0.4468536928566583
12 norm of weights: 6.163872234802683
-----
Iteration Number: 2604
Loss: 50.05166565797936
12 norm of gradients: 0.44670196583694244
12 norm of weights: 6.16374853694947
-----
Iteration Number: 2605
Loss: 50.041775526177574
12 norm of gradients: 0.44655033881706224
12 norm of weights: 6.163624866415618
-----
Iteration Number: 2606
Loss: 50.03189202605088
12 norm of gradients: 0.4463988116919078
12 norm of weights: 6.163501223181942
-----
Iteration Number: 2607
Loss: 50.02201515106461
12 norm of gradients: 0.44624738435668587
12 norm of weights: 6.163377607229272
-----
Iteration Number: 2608
Loss: 50.012144894555654
12 norm of gradients: 0.4460960567069203
12 norm of weights: 6.163254018538457
-----
Iteration Number: 2609
Loss: 50.00228124997772
12 norm of gradients: 0.44594482863845014
12 norm of weights: 6.163130457090362
-----
Iteration Number: 2610
Loss: 49.992424210828744
12 norm of gradients: 0.44579370004742924
12 norm of weights: 6.163006922865868
-----
Iteration Number: 2611
Loss: 49.982573770526564
12 norm of gradients: 0.4456426708303253
12 norm of weights: 6.162883415845875
-----
Iteration Number: 2612
Loss: 49.97272992249816
12 norm of gradients: 0.4454917408839188
12 norm of weights: 6.162759936011297
-----

```

Iteration Number: 2613
Loss: 49.962892660233095
l2 norm of gradients: 0.4453409101053025
l2 norm of weights: 6.162636483343068

Iteration Number: 2614
Loss: 49.95306197724037
l2 norm of gradients: 0.4451901783918801
l2 norm of weights: 6.162513057822135

Iteration Number: 2615
Loss: 49.94323786697022
l2 norm of gradients: 0.44503954564136555
l2 norm of weights: 6.162389659429466

Iteration Number: 2616
Loss: 49.93342032291017
l2 norm of gradients: 0.4448890117517823
l2 norm of weights: 6.162266288146042

Iteration Number: 2617
Loss: 49.923609338617396
l2 norm of gradients: 0.44473857662146227
l2 norm of weights: 6.162142943952864

Iteration Number: 2618
Loss: 49.913804907518596
l2 norm of gradients: 0.4445882401490451
l2 norm of weights: 6.162019626830946

Iteration Number: 2619
Loss: 49.90400702324151
l2 norm of gradients: 0.44443800223347674
l2 norm of weights: 6.161896336761322

Iteration Number: 2620
Loss: 49.894215679198744
l2 norm of gradients: 0.44428786277400956
l2 norm of weights: 6.161773073725041

Iteration Number: 2621
Loss: 49.884430869037324
l2 norm of gradients: 0.44413782167020066
l2 norm of weights: 6.161649837703169

Iteration Number: 2622
Loss: 49.87465258622011
l2 norm of gradients: 0.44398787882191065


```

12 norm of weights: 6.161526628676788
-----
Iteration Number: 2623
Loss: 49.86488082434731
12 norm of gradients: 0.4438380341293045
12 norm of weights: 6.161403446626998
-----
Iteration Number: 2624
Loss: 49.8551155769569
12 norm of gradients: 0.4436882874928483
12 norm of weights: 6.161280291534913
-----
Iteration Number: 2625
Loss: 49.845356837666806
12 norm of gradients: 0.44353863881331
12 norm of weights: 6.1611571633816675
-----
Iteration Number: 2626
Loss: 49.835604599979085
12 norm of gradients: 0.44338908799175847
12 norm of weights: 6.161034062148409
-----
Iteration Number: 2627
Loss: 49.82585885758739
12 norm of gradients: 0.4432396349295615
12 norm of weights: 6.160910987816303
-----
Iteration Number: 2628
Loss: 49.81611960400532
12 norm of gradients: 0.4430902795283862
12 norm of weights: 6.1607879403665295
-----
Iteration Number: 2629
Loss: 49.80638683285259
12 norm of gradients: 0.4429410216901971
12 norm of weights: 6.160664919780289
-----
Iteration Number: 2630
Loss: 49.796660537780625
12 norm of gradients: 0.44279186131725606
12 norm of weights: 6.160541926038795
-----
Iteration Number: 2631
Loss: 49.78694071240308
12 norm of gradients: 0.44264279831212094
12 norm of weights: 6.1604189591232785
-----
Iteration Number: 2632

```

Loss: 49.77722735032343
l2 norm of gradients: 0.4424938325776445
l2 norm of weights: 6.160296019014988

Iteration Number: 2633
Loss: 49.76752044519864
l2 norm of gradients: 0.4423449640169741
l2 norm of weights: 6.1601731056951845

Iteration Number: 2634
Loss: 49.75781999066039
l2 norm of gradients: 0.4421961925335507
l2 norm of weights: 6.16005021914515

Iteration Number: 2635
Loss: 49.74812598042973
l2 norm of gradients: 0.44204751803110715
l2 norm of weights: 6.159927359346181

Iteration Number: 2636
Loss: 49.7384384080658
l2 norm of gradients: 0.4418989404136687
l2 norm of weights: 6.15980452627959

Iteration Number: 2637
Loss: 49.72875726734742
l2 norm of gradients: 0.44175045958555087
l2 norm of weights: 6.159681719926704

Iteration Number: 2638
Loss: 49.719082551867835
l2 norm of gradients: 0.4416020754513592
l2 norm of weights: 6.159558940268871

Iteration Number: 2639
Loss: 49.70941425539675
l2 norm of gradients: 0.44145378791598816
l2 norm of weights: 6.159436187287452

Iteration Number: 2640
Loss: 49.69975237157872
l2 norm of gradients: 0.4413055968846206
l2 norm of weights: 6.159313460963824

Iteration Number: 2641
Loss: 49.690096894165485
l2 norm of gradients: 0.441157502262726
l2 norm of weights: 6.159190761279379

```

-----
Iteration Number: 2642
Loss: 49.680447816805
l2 norm of gradients: 0.44100950395606103
l2 norm of weights: 6.159068088215531
-----
Iteration Number: 2643
Loss: 49.670805133281696
l2 norm of gradients: 0.44086160187066664
l2 norm of weights: 6.158945441753703
-----
Iteration Number: 2644
Loss: 49.661168837282716
l2 norm of gradients: 0.44071379591286963
l2 norm of weights: 6.158822821875339
-----
Iteration Number: 2645
Loss: 49.65153892258283
l2 norm of gradients: 0.4405660859892793
l2 norm of weights: 6.158700228561898
-----
Iteration Number: 2646
Loss: 49.641915382943914
l2 norm of gradients: 0.44041847200678874
l2 norm of weights: 6.158577661794854
-----
Iteration Number: 2647
Loss: 49.63229821204371
l2 norm of gradients: 0.44027095387257215
l2 norm of weights: 6.158455121555697
-----
Iteration Number: 2648
Loss: 49.622687403717265
l2 norm of gradients: 0.4401235314940851
l2 norm of weights: 6.158332607825934
-----
Iteration Number: 2649
Loss: 49.613082951712435
l2 norm of gradients: 0.4399762047790632
l2 norm of weights: 6.158210120587088
-----
Iteration Number: 2650
Loss: 49.603484849826756
l2 norm of gradients: 0.4398289736355215
l2 norm of weights: 6.158087659820699
-----
Iteration Number: 2651
Loss: 49.5938930917949

```

```

12 norm of gradients: 0.43968183797175303
12 norm of weights: 6.15796522550832
-----
Iteration Number: 2652
Loss: 49.58430767148935
12 norm of gradients: 0.43953479769632864
12 norm of weights: 6.157842817631523
-----
Iteration Number: 2653
Loss: 49.574728582652654
12 norm of gradients: 0.43938785271809555
12 norm of weights: 6.1577204361718945
-----
Iteration Number: 2654
Loss: 49.56515581912147
12 norm of gradients: 0.43924100294617674
12 norm of weights: 6.157598081111038
-----
Iteration Number: 2655
Loss: 49.55558937471546
12 norm of gradients: 0.4390942482899702
12 norm of weights: 6.157475752430571
-----
Iteration Number: 2656
Loss: 49.54602924325212
12 norm of gradients: 0.43894758865914757
12 norm of weights: 6.157353450112128
-----
Iteration Number: 2657
Loss: 49.53647541857132
12 norm of gradients: 0.43880102396365356
12 norm of weights: 6.157231174137362
-----
Iteration Number: 2658
Loss: 49.526927894566356
12 norm of gradients: 0.43865455411370535
12 norm of weights: 6.157108924487936
-----
Iteration Number: 2659
Loss: 49.517386664996764
12 norm of gradients: 0.438508179019791
12 norm of weights: 6.156986701145534
-----
Iteration Number: 2660
Loss: 49.50785172377533
12 norm of gradients: 0.43836189859266916
12 norm of weights: 6.156864504091855
-----

```

Iteration Number: 2661
Loss: 49.49832306474557
l2 norm of gradients: 0.4382157127433683
l2 norm of weights: 6.156742333308611

Iteration Number: 2662
Loss: 49.48880068184789
l2 norm of gradients: 0.43806962138318495
l2 norm of weights: 6.156620188777533

Iteration Number: 2663
Loss: 49.479284568864806
l2 norm of gradients: 0.4379236244236839
l2 norm of weights: 6.1564980704803665

Iteration Number: 2664
Loss: 49.46977471975384
l2 norm of gradients: 0.4377777217766964
l2 norm of weights: 6.156375978398874

Iteration Number: 2665
Loss: 49.460271128403924
l2 norm of gradients: 0.43763191335432017
l2 norm of weights: 6.156253912514829

Iteration Number: 2666
Loss: 49.45077378872849
l2 norm of gradients: 0.43748619906891734
l2 norm of weights: 6.156131872810027

Iteration Number: 2667
Loss: 49.44128269459612
l2 norm of gradients: 0.43734057883311545
l2 norm of weights: 6.156009859266276

Iteration Number: 2668
Loss: 49.43179783996594
l2 norm of gradients: 0.43719505255980406
l2 norm of weights: 6.155887871865399

Iteration Number: 2669
Loss: 49.42231921878194
l2 norm of gradients: 0.4370496201621362
l2 norm of weights: 6.155765910589239

Iteration Number: 2670
Loss: 49.412846824924685
l2 norm of gradients: 0.436904281553526

```

12 norm of weights: 6.155643975419649
-----
Iteration Number: 2671
Loss: 49.403380652387355
12 norm of gradients: 0.4367590366476484
12 norm of weights: 6.1555220663385
-----
Iteration Number: 2672
Loss: 49.39392069510044
12 norm of gradients: 0.43661388535843837
12 norm of weights: 6.15540018332768
-----
Iteration Number: 2673
Loss: 49.38446694706293
12 norm of gradients: 0.4364688276000896
12 norm of weights: 6.155278326369092
-----
Iteration Number: 2674
Loss: 49.3750194021649
12 norm of gradients: 0.4363238632870542
12 norm of weights: 6.155156495444652
-----
Iteration Number: 2675
Loss: 49.36557805445415
12 norm of gradients: 0.4361789923340412
12 norm of weights: 6.155034690536295
-----
Iteration Number: 2676
Loss: 49.35614289788662
12 norm of gradients: 0.4360342146560162
12 norm of weights: 6.154912911625971
-----
Iteration Number: 2677
Loss: 49.346713926441076
12 norm of gradients: 0.4358895301682001
12 norm of weights: 6.154791158695643
-----
Iteration Number: 2678
Loss: 49.33729113412456
12 norm of gradients: 0.43574493878606846
12 norm of weights: 6.154669431727291
-----
Iteration Number: 2679
Loss: 49.32787451495101
12 norm of gradients: 0.4356004404253505
12 norm of weights: 6.154547730702912
-----
Iteration Number: 2680

```

```

Loss: 49.31846406292753
l2 norm of gradients: 0.4354560350020286
l2 norm of weights: 6.154426055604519
-----
Iteration Number: 2681
Loss: 49.30905977205259
l2 norm of gradients: 0.4353117224323367
l2 norm of weights: 6.154304406414135
-----
Iteration Number: 2682
Loss: 49.29966163638812
l2 norm of gradients: 0.4351675026327601
l2 norm of weights: 6.154182783113805
-----
Iteration Number: 2683
Loss: 49.290269649954524
l2 norm of gradients: 0.43502337552003434
l2 norm of weights: 6.1540611856855865
-----
Iteration Number: 2684
Loss: 49.2808838067776
l2 norm of gradients: 0.43487934101114434
l2 norm of weights: 6.15393961411155
-----
Iteration Number: 2685
Loss: 49.271504100944064
l2 norm of gradients: 0.43473539902332325
l2 norm of weights: 6.153818068373787
-----
Iteration Number: 2686
Loss: 49.262130526465086
l2 norm of gradients: 0.43459154947405226
l2 norm of weights: 6.153696548454399
-----
Iteration Number: 2687
Loss: 49.25276307743976
l2 norm of gradients: 0.4344477922810591
l2 norm of weights: 6.1535750543355086
-----
Iteration Number: 2688
Loss: 49.24340174793116
l2 norm of gradients: 0.43430412736231744
l2 norm of weights: 6.153453585999247
-----
Iteration Number: 2689
Loss: 49.234046532023264
l2 norm of gradients: 0.4341605546360459
l2 norm of weights: 6.1533321434277655

```

```

-----
Iteration Number: 2690
Loss: 49.22469742378331
l2 norm of gradients: 0.43401707402070755
l2 norm of weights: 6.1532107266032305
-----
Iteration Number: 2691
Loss: 49.21535441732837
l2 norm of gradients: 0.43387368543500826
l2 norm of weights: 6.153089335507822
-----
Iteration Number: 2692
Loss: 49.20601750673614
l2 norm of gradients: 0.433730388797897
l2 norm of weights: 6.1529679701237345
-----
Iteration Number: 2693
Loss: 49.19668668613066
l2 norm of gradients: 0.4335871840285637
l2 norm of weights: 6.1528466304331815
-----
Iteration Number: 2694
Loss: 49.18736194962984
l2 norm of gradients: 0.4334440710464393
l2 norm of weights: 6.1527253164183895
-----
Iteration Number: 2695
Loss: 49.17804329137359
l2 norm of gradients: 0.4333010497711946
l2 norm of weights: 6.152604028061598
-----
Iteration Number: 2696
Loss: 49.16873070544252
l2 norm of gradients: 0.43315812012273935
l2 norm of weights: 6.1524827653450656
-----
Iteration Number: 2697
Loss: 49.15942418598808
l2 norm of gradients: 0.4330152820212213
l2 norm of weights: 6.152361528251065
-----
Iteration Number: 2698
Loss: 49.15012372719943
l2 norm of gradients: 0.4328725353870257
l2 norm of weights: 6.152240316761883
-----
Iteration Number: 2699
Loss: 49.14082932317792

```



```

12 norm of gradients: 0.43272988014077407
12 norm of weights: 6.1521191308598215
-----
Iteration Number: 2700
Loss: 49.13154096809851
12 norm of gradients: 0.43258731620332347
12 norm of weights: 6.1519979705272005
-----
Iteration Number: 2701
Loss: 49.12225865612845
12 norm of gradients: 0.43244484349576556
12 norm of weights: 6.151876835746352
-----
Iteration Number: 2702
Loss: 49.11298238146159
12 norm of gradients: 0.43230246193942634
12 norm of weights: 6.151755726499623
-----
Iteration Number: 2703
Loss: 49.10371213825044
12 norm of gradients: 0.4321601714558641
12 norm of weights: 6.151634642769379
-----
Iteration Number: 2704
Loss: 49.09444792067282
12 norm of gradients: 0.4320179719668697
12 norm of weights: 6.151513584537995
-----
Iteration Number: 2705
Loss: 49.08518972295408
12 norm of gradients: 0.4318758633944651
12 norm of weights: 6.151392551787869
-----
Iteration Number: 2706
Loss: 49.07593753928265
12 norm of gradients: 0.4317338456609029
12 norm of weights: 6.151271544501407
-----
Iteration Number: 2707
Loss: 49.06669136384782
12 norm of gradients: 0.431591918688665
12 norm of weights: 6.151150562661033
-----
Iteration Number: 2708
Loss: 49.05745119087742
12 norm of gradients: 0.43145008240046206
12 norm of weights: 6.151029606249185
-----

```

Iteration Number: 2709
Loss: 49.04821701463741
l2 norm of gradients: 0.43130833671923285
l2 norm of weights: 6.150908675248319

Iteration Number: 2710
Loss: 49.03898882928176
l2 norm of gradients: 0.431166681568143
l2 norm of weights: 6.150787769640902

Iteration Number: 2711
Loss: 49.02976662909059
l2 norm of gradients: 0.4310251168705841
l2 norm of weights: 6.15066688940942

Iteration Number: 2712
Loss: 49.02055040830716
l2 norm of gradients: 0.4308836425501734
l2 norm of weights: 6.15054603453637

Iteration Number: 2713
Loss: 49.0113401611761
l2 norm of gradients: 0.43074225853075265
l2 norm of weights: 6.150425205004266

Iteration Number: 2714
Loss: 49.00213588192633
l2 norm of gradients: 0.4306009647363869
l2 norm of weights: 6.150304400795638

Iteration Number: 2715
Loss: 48.99293756486357
l2 norm of gradients: 0.4304597610913644
l2 norm of weights: 6.150183621893028

Iteration Number: 2716
Loss: 48.98374520425567
l2 norm of gradients: 0.4303186475201949
l2 norm of weights: 6.150062868278997

Iteration Number: 2717
Loss: 48.97455879435062
l2 norm of gradients: 0.43017762394760967
l2 norm of weights: 6.149942139936117

Iteration Number: 2718
Loss: 48.9653783294816
l2 norm of gradients: 0.43003669029855984

```

12 norm of weights: 6.1498214368469775
-----
Iteration Number: 2719
Loss: 48.95620380385683
12 norm of gradients: 0.4298958464982166
12 norm of weights: 6.149700758994182
-----
Iteration Number: 2720
Loss: 48.947035211875395
12 norm of gradients: 0.429755092471969
12 norm of weights: 6.149580106360349
-----
Iteration Number: 2721
Loss: 48.937872547764734
12 norm of gradients: 0.42961442814542417
12 norm of weights: 6.149459478928112
-----
Iteration Number: 2722
Loss: 48.92871580588035
12 norm of gradients: 0.42947385344440636
12 norm of weights: 6.149338876680118
-----
Iteration Number: 2723
Loss: 48.919564980504155
12 norm of gradients: 0.4293333682949555
12 norm of weights: 6.149218299599031
-----
Iteration Number: 2724
Loss: 48.91042006598687
12 norm of gradients: 0.42919297262332695
12 norm of weights: 6.14909774766753
-----
Iteration Number: 2725
Loss: 48.901281056684944
12 norm of gradients: 0.4290526663559905
12 norm of weights: 6.148977220868306
-----
Iteration Number: 2726
Loss: 48.892147946879014
12 norm of gradients: 0.4289124494196295
12 norm of weights: 6.148856719184067
-----
Iteration Number: 2727
Loss: 48.883020730967864
12 norm of gradients: 0.42877232174114
12 norm of weights: 6.148736242597535
-----
Iteration Number: 2728

```

Loss: 48.87389940324579
l2 norm of gradients: 0.42863228324763
l2 norm of weights: 6.14861579109145

Iteration Number: 2729
Loss: 48.86478395814148
l2 norm of gradients: 0.4284923338664185
l2 norm of weights: 6.14849536464856

Iteration Number: 2730
Loss: 48.8556743899877
l2 norm of gradients: 0.42835247352503497
l2 norm of weights: 6.148374963251634

Iteration Number: 2731
Loss: 48.846570693130296
l2 norm of gradients: 0.42821270215121815
l2 norm of weights: 6.148254586883452

Iteration Number: 2732
Loss: 48.837472861969346
l2 norm of gradients: 0.4280730196729153
l2 norm of weights: 6.148134235526811

Iteration Number: 2733
Loss: 48.82838089091628
l2 norm of gradients: 0.42793342601828144
l2 norm of weights: 6.148013909164522

Iteration Number: 2734
Loss: 48.81929477432669
l2 norm of gradients: 0.427793921115679
l2 norm of weights: 6.147893607779411

Iteration Number: 2735
Loss: 48.81021450662924
l2 norm of gradients: 0.4276545048936759
l2 norm of weights: 6.147773331354317

Iteration Number: 2736
Loss: 48.801140082194955
l2 norm of gradients: 0.4275151772810458
l2 norm of weights: 6.1476530798720965

Iteration Number: 2737
Loss: 48.79207149547704
l2 norm of gradients: 0.4273759382067669
l2 norm of weights: 6.147532853315618

```

-----
Iteration Number: 2738
Loss: 48.783008740847755
l2 norm of gradients: 0.4272367876000206
l2 norm of weights: 6.147412651667767
-----
Iteration Number: 2739
Loss: 48.773951812768225
l2 norm of gradients: 0.4270977253901916
l2 norm of weights: 6.147292474911441
-----
Iteration Number: 2740
Loss: 48.76490070565275
l2 norm of gradients: 0.42695875150686635
l2 norm of weights: 6.1471723230295545
-----
Iteration Number: 2741
Loss: 48.75585541395455
l2 norm of gradients: 0.42681986587983295
l2 norm of weights: 6.147052196005036
-----
Iteration Number: 2742
Loss: 48.746815932094215
l2 norm of gradients: 0.4266810684390794
l2 norm of weights: 6.146932093820827
-----
Iteration Number: 2743
Loss: 48.73778225453594
l2 norm of gradients: 0.42654235911479355
l2 norm of weights: 6.146812016459887
-----
Iteration Number: 2744
Loss: 48.728754375763536
l2 norm of gradients: 0.42640373783736196
l2 norm of weights: 6.146691963905187
-----
Iteration Number: 2745
Loss: 48.7197322902065
l2 norm of gradients: 0.42626520453736944
l2 norm of weights: 6.146571936139712
-----
Iteration Number: 2746
Loss: 48.71071599233184
l2 norm of gradients: 0.42612675914559756
l2 norm of weights: 6.146451933146467
-----
Iteration Number: 2747
Loss: 48.701705476659

```

```

12 norm of gradients: 0.4259884015930243
12 norm of weights: 6.146331954908463
-----
Iteration Number: 2748
Loss: 48.69270073762024
12 norm of gradients: 0.42585013181082365
12 norm of weights: 6.146212001408735
-----
Iteration Number: 2749
Loss: 48.683701769721594
12 norm of gradients: 0.4257119497303639
12 norm of weights: 6.146092072630323
-----
Iteration Number: 2750
Loss: 48.67470856746233
12 norm of gradients: 0.42557385528320724
12 norm of weights: 6.145972168556291
-----
Iteration Number: 2751
Loss: 48.66572112536116
12 norm of gradients: 0.4254358484011095
12 norm of weights: 6.145852289169709
-----
Iteration Number: 2752
Loss: 48.65673943790746
12 norm of gradients: 0.42529792901601815
12 norm of weights: 6.1457324344536675
-----
Iteration Number: 2753
Loss: 48.64776349963022
12 norm of gradients: 0.4251600970600728
12 norm of weights: 6.145612604391269
-----
Iteration Number: 2754
Loss: 48.63879330503914
12 norm of gradients: 0.4250223524656036
12 norm of weights: 6.14549279896563
-----
Iteration Number: 2755
Loss: 48.629828848669625
12 norm of gradients: 0.4248846951651304
12 norm of weights: 6.145373018159883
-----
Iteration Number: 2756
Loss: 48.62087012504291
12 norm of gradients: 0.4247471250913626
12 norm of weights: 6.145253261957173
-----

```

Iteration Number: 2757
Loss: 48.611917128700306
l2 norm of gradients: 0.42460964217719765
l2 norm of weights: 6.145133530340661

Iteration Number: 2758
Loss: 48.60296985423666
l2 norm of gradients: 0.42447224635572073
l2 norm of weights: 6.145013823293522

Iteration Number: 2759
Loss: 48.59402829612554
l2 norm of gradients: 0.42433493756020385
l2 norm of weights: 6.144894140798946

Iteration Number: 2760
Loss: 48.58509244897078
l2 norm of gradients: 0.4241977157241049
l2 norm of weights: 6.144774482840137

Iteration Number: 2761
Loss: 48.576162307355524
l2 norm of gradients: 0.4240605807810668
l2 norm of weights: 6.144654849400312

Iteration Number: 2762
Loss: 48.567237865824886
l2 norm of gradients: 0.42392353266491706
l2 norm of weights: 6.144535240462705

Iteration Number: 2763
Loss: 48.55831911894841
l2 norm of gradients: 0.4237865713096669
l2 norm of weights: 6.144415656010562

Iteration Number: 2764
Loss: 48.549406061325406
l2 norm of gradients: 0.4236496966495102
l2 norm of weights: 6.144296096027143

Iteration Number: 2765
Loss: 48.54049868755099
l2 norm of gradients: 0.4235129086188229
l2 norm of weights: 6.144176560495726

Iteration Number: 2766
Loss: 48.53159699218867
l2 norm of gradients: 0.42337620715216234

```

12 norm of weights: 6.1440570493996
-----
Iteration Number: 2767
Loss: 48.52270096989265
12 norm of gradients: 0.4232395921842661
12 norm of weights: 6.143937562722069
-----
Iteration Number: 2768
Loss: 48.51381061526429
12 norm of gradients: 0.42310306365005185
12 norm of weights: 6.143818100446452
-----
Iteration Number: 2769
Loss: 48.504925922863464
12 norm of gradients: 0.4229666214846158
12 norm of weights: 6.14369866255608
-----
Iteration Number: 2770
Loss: 48.49604688736205
12 norm of gradients: 0.42283026562323267
12 norm of weights: 6.143579249034304
-----
Iteration Number: 2771
Loss: 48.48717350339615
12 norm of gradients: 0.4226939960013541
12 norm of weights: 6.143459859864482
-----
Iteration Number: 2772
Loss: 48.47830576555618
12 norm of gradients: 0.4225578125546089
12 norm of weights: 6.143340495029991
-----
Iteration Number: 2773
Loss: 48.469443668495686
12 norm of gradients: 0.4224217152188012
12 norm of weights: 6.143221154514221
-----
Iteration Number: 2774
Loss: 48.4605872068734
12 norm of gradients: 0.4222857039299105
12 norm of weights: 6.143101838300576
-----
Iteration Number: 2775
Loss: 48.451736375348894
12 norm of gradients: 0.4221497786240905
12 norm of weights: 6.142982546372474
-----
Iteration Number: 2776

```


Loss: 48.44289116854045
l2 norm of gradients: 0.4220139392376684
l2 norm of weights: 6.142863278713349

Iteration Number: 2777
Loss: 48.434051581155074
l2 norm of gradients: 0.42187818570714436
l2 norm of weights: 6.142744035306646

Iteration Number: 2778
Loss: 48.42521760784221
l2 norm of gradients: 0.4217425179691903
l2 norm of weights: 6.142624816135828

Iteration Number: 2779
Loss: 48.41638924326188
l2 norm of gradients: 0.42160693596064946
l2 norm of weights: 6.142505621184368

Iteration Number: 2780
Loss: 48.40756648212929
l2 norm of gradients: 0.4214714396185354
l2 norm of weights: 6.142386450435757

Iteration Number: 2781
Loss: 48.398749319117485
l2 norm of gradients: 0.4213360288800318
l2 norm of weights: 6.142267303873499

Iteration Number: 2782
Loss: 48.38993774890597
l2 norm of gradients: 0.4212007036824909
l2 norm of weights: 6.142148181481111

Iteration Number: 2783
Loss: 48.38113176620428
l2 norm of gradients: 0.42106546396343325
l2 norm of weights: 6.142029083242124

Iteration Number: 2784
Loss: 48.37233136573185
l2 norm of gradients: 0.420930309660547
l2 norm of weights: 6.141910009140086

Iteration Number: 2785
Loss: 48.363536542183766
l2 norm of gradients: 0.42079524071168667
l2 norm of weights: 6.141790959158556

```

-----
Iteration Number: 2786
Loss: 48.354747290263276
l2 norm of gradients: 0.4206602570548729
l2 norm of weights: 6.141671933281108
-----
Iteration Number: 2787
Loss: 48.34596360473895
l2 norm of gradients: 0.4205253586282917
l2 norm of weights: 6.141552931491332
-----
Iteration Number: 2788
Loss: 48.33718548031243
l2 norm of gradients: 0.4203905453702933
l2 norm of weights: 6.141433953772828
-----
Iteration Number: 2789
Loss: 48.328412911714295
l2 norm of gradients: 0.4202558172193915
l2 norm of weights: 6.1413150001092145
-----
Iteration Number: 2790
Loss: 48.31964589369701
l2 norm of gradients: 0.4201211741142632
l2 norm of weights: 6.1411960704841215
-----
Iteration Number: 2791
Loss: 48.31088442098876
l2 norm of gradients: 0.4199866159937477
l2 norm of weights: 6.141077164881194
-----
Iteration Number: 2792
Loss: 48.30212848832877
l2 norm of gradients: 0.4198521427968453
l2 norm of weights: 6.140958283284091
-----
Iteration Number: 2793
Loss: 48.29337809053443
l2 norm of gradients: 0.41971775446271736
l2 norm of weights: 6.140839425676484
-----
Iteration Number: 2794
Loss: 48.28463322235574
l2 norm of gradients: 0.4195834509306852
l2 norm of weights: 6.140720592042061
-----
Iteration Number: 2795
Loss: 48.27589387851914

```

```

12 norm of gradients: 0.4194492321402291
12 norm of weights: 6.140601782364523
-----
Iteration Number: 2796
Loss: 48.26716005381203
12 norm of gradients: 0.41931509803098815
12 norm of weights: 6.1404829966275845
-----
Iteration Number: 2797
Loss: 48.25843174305963
12 norm of gradients: 0.419181048542759
12 norm of weights: 6.1403642348149745
-----
Iteration Number: 2798
Loss: 48.24970894100505
12 norm of gradients: 0.41904708361549536
12 norm of weights: 6.140245496910436
-----
Iteration Number: 2799
Loss: 48.240991642454226
12 norm of gradients: 0.41891320318930747
12 norm of weights: 6.140126782897725
-----
Iteration Number: 2800
Loss: 48.23227984221993
12 norm of gradients: 0.4187794072044608
12 norm of weights: 6.140008092760614
-----
Iteration Number: 2801
Loss: 48.22357353506349
12 norm of gradients: 0.4186456956013756
12 norm of weights: 6.139889426482887
-----
Iteration Number: 2802
Loss: 48.21487271586051
12 norm of gradients: 0.41851206832062676
12 norm of weights: 6.139770784048342
-----
Iteration Number: 2803
Loss: 48.20617737938942
12 norm of gradients: 0.4183785253029421
12 norm of weights: 6.139652165440794
-----
Iteration Number: 2804
Loss: 48.19748752045212
12 norm of gradients: 0.4182450664892022
12 norm of weights: 6.139533570644068
-----

```

Iteration Number: 2805
Loss: 48.18880313392291
l2 norm of gradients: 0.41811169182043967
l2 norm of weights: 6.139414999642005

Iteration Number: 2806
Loss: 48.18012421461734
l2 norm of gradients: 0.41797840123783825
l2 norm of weights: 6.13929645241846

Iteration Number: 2807
Loss: 48.17145075735204
l2 norm of gradients: 0.4178451946827321
l2 norm of weights: 6.1391779289573

Iteration Number: 2808
Loss: 48.16278275699445
l2 norm of gradients: 0.41771207209660544
l2 norm of weights: 6.139059429242409

Iteration Number: 2809
Loss: 48.15412020839839
l2 norm of gradients: 0.4175790334210913
l2 norm of weights: 6.138940953257683

Iteration Number: 2810
Loss: 48.14546310640086
l2 norm of gradients: 0.41744607859797095
l2 norm of weights: 6.138822500987032

Iteration Number: 2811
Loss: 48.1368114458745
l2 norm of gradients: 0.4173132075691737
l2 norm of weights: 6.138704072414379

Iteration Number: 2812
Loss: 48.128165221701565
l2 norm of gradients: 0.41718042027677554
l2 norm of weights: 6.138585667523663

Iteration Number: 2813
Loss: 48.11952442874166
l2 norm of gradients: 0.41704771666299867
l2 norm of weights: 6.138467286298836

Iteration Number: 2814
Loss: 48.11088906185621
l2 norm of gradients: 0.4169150966702108

```

12 norm of weights: 6.138348928723862
-----
Iteration Number: 2815
Loss: 48.10225911594273
12 norm of gradients: 0.41678256024092464
12 norm of weights: 6.138230594782724
-----
Iteration Number: 2816
Loss: 48.09363458590438
12 norm of gradients: 0.41665010731779667
12 norm of weights: 6.13811228445941
-----
Iteration Number: 2817
Loss: 48.08501546661688
12 norm of gradients: 0.416517737843627
12 norm of weights: 6.137993997737931
-----
Iteration Number: 2818
Loss: 48.07640175297094
12 norm of gradients: 0.41638545176135844
12 norm of weights: 6.137875734602307
-----
Iteration Number: 2819
Loss: 48.06779343994005
12 norm of gradients: 0.4162532490140757
12 norm of weights: 6.137757495036572
-----
Iteration Number: 2820
Loss: 48.059190522335385
12 norm of gradients: 0.41612112954500474
12 norm of weights: 6.137639279024776
-----
Iteration Number: 2821
Loss: 48.05059299514419
12 norm of gradients: 0.4159890932975124
12 norm of weights: 6.13752108655098
-----
Iteration Number: 2822
Loss: 48.042000853281976
12 norm of gradients: 0.41585714021510506
12 norm of weights: 6.13740291759926
-----
Iteration Number: 2823
Loss: 48.033414091673684
12 norm of gradients: 0.41572527024142863
12 norm of weights: 6.137284772153706
-----
Iteration Number: 2824

```

```

Loss: 48.02483270522068
l2 norm of gradients: 0.4155934833202674
l2 norm of weights: 6.137166650198423
-----
Iteration Number: 2825
Loss: 48.01625668890603
l2 norm of gradients: 0.41546177939554346
l2 norm of weights: 6.137048551717525
-----
Iteration Number: 2826
Loss: 48.007686037626016
l2 norm of gradients: 0.415330158411316
l2 norm of weights: 6.136930476695146
-----
Iteration Number: 2827
Loss: 47.99912074639627
l2 norm of gradients: 0.41519862031178106
l2 norm of weights: 6.13681242511543
-----
Iteration Number: 2828
Loss: 47.99056081011155
l2 norm of gradients: 0.41506716504126995
l2 norm of weights: 6.136694396962535
-----
Iteration Number: 2829
Loss: 47.982006223790876
l2 norm of gradients: 0.41493579254424934
l2 norm of weights: 6.1365763922206344
-----
Iteration Number: 2830
Loss: 47.973456982340664
l2 norm of gradients: 0.4148045027653205
l2 norm of weights: 6.136458410873914
-----
Iteration Number: 2831
Loss: 47.96491308076519
l2 norm of gradients: 0.41467329564921795
l2 norm of weights: 6.1363404529065715
-----
Iteration Number: 2832
Loss: 47.95637451403581
l2 norm of gradients: 0.4145421711408097
l2 norm of weights: 6.136222518302822
-----
Iteration Number: 2833
Loss: 47.947841277161054
l2 norm of gradients: 0.4144111291850958
l2 norm of weights: 6.136104607046892

```

```

-----
Iteration Number: 2834
Loss: 47.93931336507909
l2 norm of gradients: 0.41428016972720844
l2 norm of weights: 6.135986719123023
-----
Iteration Number: 2835
Loss: 47.93079077283014
l2 norm of gradients: 0.4141492927124105
l2 norm of weights: 6.135868854515468
-----
Iteration Number: 2836
Loss: 47.9222734953833
l2 norm of gradients: 0.4140184980860954
l2 norm of weights: 6.135751013208496
-----
Iteration Number: 2837
Loss: 47.91376152775981
l2 norm of gradients: 0.41388778579378616
l2 norm of weights: 6.135633195186387
-----
Iteration Number: 2838
Loss: 47.90525486494495
l2 norm of gradients: 0.4137571557811348
l2 norm of weights: 6.135515400433438
-----
Iteration Number: 2839
Loss: 47.896753501969705
l2 norm of gradients: 0.41362660799392187
l2 norm of weights: 6.1353976289339585
-----
Iteration Number: 2840
Loss: 47.88825743389448
l2 norm of gradients: 0.4134961423780556
l2 norm of weights: 6.135279880672269
-----
Iteration Number: 2841
Loss: 47.8797666556781
l2 norm of gradients: 0.41336575887957105
l2 norm of weights: 6.135162155632706
-----
Iteration Number: 2842
Loss: 47.87128116238553
l2 norm of gradients: 0.4132354574446299
l2 norm of weights: 6.135044453799622
-----
Iteration Number: 2843
Loss: 47.86280094904794

```

```

12 norm of gradients: 0.41310523801951954
12 norm of weights: 6.134926775157376
-----
Iteration Number: 2844
Loss: 47.85432601071086
12 norm of gradients: 0.41297510055065234
12 norm of weights: 6.134809119690347
-----
Iteration Number: 2845
Loss: 47.84585634240717
12 norm of gradients: 0.41284504498456503
12 norm of weights: 6.134691487382927
-----
Iteration Number: 2846
Loss: 47.83739193921175
12 norm of gradients: 0.4127150712679182
12 norm of weights: 6.134573878219518
-----
Iteration Number: 2847
Loss: 47.82893279617922
12 norm of gradients: 0.4125851793474955
12 norm of weights: 6.134456292184538
-----
Iteration Number: 2848
Loss: 47.82047890834343
12 norm of gradients: 0.4124553691702032
12 norm of weights: 6.134338729262419
-----
Iteration Number: 2849
Loss: 47.81203027079436
12 norm of gradients: 0.4123256406830691
12 norm of weights: 6.134221189437604
-----
Iteration Number: 2850
Loss: 47.803586878620976
12 norm of gradients: 0.41219599383324246
12 norm of weights: 6.134103672694554
-----
Iteration Number: 2851
Loss: 47.79514872688814
12 norm of gradients: 0.41206642856799286
12 norm of weights: 6.133986179017739
-----
Iteration Number: 2852
Loss: 47.78671581065278
12 norm of gradients: 0.4119369448347097
12 norm of weights: 6.133868708391644
-----

```


Iteration Number: 2853
Loss: 47.778288125028226
l2 norm of gradients: 0.411807542580902
l2 norm of weights: 6.133751260800768

Iteration Number: 2854
Loss: 47.76986566513312
l2 norm of gradients: 0.41167822175419694
l2 norm of weights: 6.133633836229624

Iteration Number: 2855
Loss: 47.76144842602721
l2 norm of gradients: 0.41154898230233977
l2 norm of weights: 6.133516434662738

Iteration Number: 2856
Loss: 47.75303640281362
l2 norm of gradients: 0.41141982417319334
l2 norm of weights: 6.133399056084649

Iteration Number: 2857
Loss: 47.74462959062608
l2 norm of gradients: 0.411290747314737
l2 norm of weights: 6.133281700479909

Iteration Number: 2858
Loss: 47.73622798459097
l2 norm of gradients: 0.41116175167506586
l2 norm of weights: 6.1331643678330865

Iteration Number: 2859
Loss: 47.727831579778986
l2 norm of gradients: 0.41103283720239114
l2 norm of weights: 6.133047058128758

Iteration Number: 2860
Loss: 47.71944037136292
l2 norm of gradients: 0.41090400384503833
l2 norm of weights: 6.13292977135152

Iteration Number: 2861
Loss: 47.71105435445748
l2 norm of gradients: 0.4107752515514473
l2 norm of weights: 6.132812507485977

Iteration Number: 2862
Loss: 47.70267352418014
l2 norm of gradients: 0.41064658027017115

```

12 norm of weights: 6.13269526651675
-----
Iteration Number: 2863
Loss: 47.69429787568595
12 norm of gradients: 0.4105179899498764
12 norm of weights: 6.132578048428473
-----
Iteration Number: 2864
Loss: 47.68592740413501
12 norm of gradients: 0.4103894805393418
12 norm of weights: 6.1324608532057905
-----
Iteration Number: 2865
Loss: 47.67756210465869
12 norm of gradients: 0.41026105198745744
12 norm of weights: 6.132343680833366
-----
Iteration Number: 2866
Loss: 47.66920197240191
12 norm of gradients: 0.41013270424322473
12 norm of weights: 6.1322265312958715
-----
Iteration Number: 2867
Loss: 47.660847002565845
12 norm of gradients: 0.4100044372557555
12 norm of weights: 6.132109404577995
-----
Iteration Number: 2868
Loss: 47.652497190269784
12 norm of gradients: 0.40987625097427155
12 norm of weights: 6.131992300664438
-----
Iteration Number: 2869
Loss: 47.64415253072553
12 norm of gradients: 0.4097481453481036
12 norm of weights: 6.1318752195399115
-----
Iteration Number: 2870
Loss: 47.63581301909971
12 norm of gradients: 0.4096201203266912
12 norm of weights: 6.131758161189145
-----
Iteration Number: 2871
Loss: 47.62747865055505
12 norm of gradients: 0.4094921758595818
12 norm of weights: 6.131641125596879
-----
Iteration Number: 2872

```

```

Loss: 47.61914942026411
l2 norm of gradients: 0.40936431189643047
l2 norm of weights: 6.131524112747868
-----
Iteration Number: 2873
Loss: 47.61082532348357
l2 norm of gradients: 0.4092365283869988
l2 norm of weights: 6.131407122626878
-----
Iteration Number: 2874
Loss: 47.602506355343515
l2 norm of gradients: 0.4091088252811546
l2 norm of weights: 6.131290155218691
-----
Iteration Number: 2875
Loss: 47.59419251105473
l2 norm of gradients: 0.4089812025288717
l2 norm of weights: 6.131173210508102
-----
Iteration Number: 2876
Loss: 47.58588378586976
l2 norm of gradients: 0.4088536600802283
l2 norm of weights: 6.1310562884799165
-----
Iteration Number: 2877
Loss: 47.57758017494753
l2 norm of gradients: 0.40872619788540726
l2 norm of weights: 6.130939389118955
-----
Iteration Number: 2878
Loss: 47.569281673523676
l2 norm of gradients: 0.4085988158946955
l2 norm of weights: 6.130822512410055
-----
Iteration Number: 2879
Loss: 47.56098827685467
l2 norm of gradients: 0.40847151405848264
l2 norm of weights: 6.130705658338061
-----
Iteration Number: 2880
Loss: 47.552699980102396
l2 norm of gradients: 0.4083442923272612
l2 norm of weights: 6.130588826887835
-----
Iteration Number: 2881
Loss: 47.54441677853569
l2 norm of gradients: 0.40821715065162584
l2 norm of weights: 6.13047201804425

```

```

-----
Iteration Number: 2882
Loss: 47.536138667381245
l2 norm of gradients: 0.40809008898227234
l2 norm of weights: 6.130355231792194
-----
Iteration Number: 2883
Loss: 47.527865641885754
l2 norm of gradients: 0.40796310726999746
l2 norm of weights: 6.130238468116569
-----
Iteration Number: 2884
Loss: 47.51959769728224
l2 norm of gradients: 0.4078362054656984
l2 norm of weights: 6.130121727002286
-----
Iteration Number: 2885
Loss: 47.51133482885997
l2 norm of gradients: 0.40770938352037145
l2 norm of weights: 6.130005008434273
-----
Iteration Number: 2886
Loss: 47.50307703181309
l2 norm of gradients: 0.4075826413851129
l2 norm of weights: 6.1298883123974734
-----
Iteration Number: 2887
Loss: 47.494824301455914
l2 norm of gradients: 0.40745597901111674
l2 norm of weights: 6.1297716388768375
-----
Iteration Number: 2888
Loss: 47.48657663302859
l2 norm of gradients: 0.40732939634967513
l2 norm of weights: 6.129654987857333
-----
Iteration Number: 2889
Loss: 47.47833402180278
l2 norm of gradients: 0.4072028933521778
l2 norm of weights: 6.1295383593239405
-----
Iteration Number: 2890
Loss: 47.47009646307938
l2 norm of gradients: 0.40707646997011104
l2 norm of weights: 6.129421753261654
-----
Iteration Number: 2891
Loss: 47.46186395209481

```

```

12 norm of gradients: 0.40695012615505743
12 norm of weights: 6.129305169655479
-----
Iteration Number: 2892
Loss: 47.453636484152156
12 norm of gradients: 0.40682386185869507
12 norm of weights: 6.129188608490435
-----
Iteration Number: 2893
Loss: 47.445414054574016
12 norm of gradients: 0.40669767703279713
12 norm of weights: 6.129072069751557
-----
Iteration Number: 2894
Loss: 47.43719665859246
12 norm of gradients: 0.40657157162923163
12 norm of weights: 6.128955553423889
-----
Iteration Number: 2895
Loss: 47.42898429157083
12 norm of gradients: 0.40644554559996005
12 norm of weights: 6.128839059492493
-----
Iteration Number: 2896
Loss: 47.42077694877457
12 norm of gradients: 0.4063195988970374
12 norm of weights: 6.128722587942439
-----
Iteration Number: 2897
Loss: 47.41257462552308
12 norm of gradients: 0.40619373147261173
12 norm of weights: 6.128606138758815
-----
Iteration Number: 2898
Loss: 47.40437731714277
12 norm of gradients: 0.40606794327892287
12 norm of weights: 6.128489711926718
-----
Iteration Number: 2899
Loss: 47.39618501894315
12 norm of gradients: 0.40594223426830267
12 norm of weights: 6.128373307431262
-----
Iteration Number: 2900
Loss: 47.38799772622763
12 norm of gradients: 0.40581660439317396
12 norm of weights: 6.128256925257571
-----

```

Iteration Number: 2901
Loss: 47.37981543434337
l2 norm of gradients: 0.40569105360605034
l2 norm of weights: 6.128140565390784

Iteration Number: 2902
Loss: 47.37163813865652
l2 norm of gradients: 0.40556558185953534
l2 norm of weights: 6.128024227816053

Iteration Number: 2903
Loss: 47.363465834442216
l2 norm of gradients: 0.40544018910632157
l2 norm of weights: 6.127907912518542

Iteration Number: 2904
Loss: 47.355298517079596
l2 norm of gradients: 0.40531487529919114
l2 norm of weights: 6.127791619483428

Iteration Number: 2905
Loss: 47.34713618193351
l2 norm of gradients: 0.4051896403910143
l2 norm of weights: 6.127675348695905

Iteration Number: 2906
Loss: 47.33897882431057
l2 norm of gradients: 0.405064484334749
l2 norm of weights: 6.127559100141175

Iteration Number: 2907
Loss: 47.33082643961958
l2 norm of gradients: 0.40493940708344056
l2 norm of weights: 6.127442873804455

Iteration Number: 2908
Loss: 47.32267902317737
l2 norm of gradients: 0.40481440859022094
l2 norm of weights: 6.1273266696709765

Iteration Number: 2909
Loss: 47.31453657036684
l2 norm of gradients: 0.4046894888083087
l2 norm of weights: 6.127210487725982

Iteration Number: 2910
Loss: 47.30639907657402
l2 norm of gradients: 0.40456464769100736

```

12 norm of weights: 6.127094327954729
-----
Iteration Number: 2911
Loss: 47.298266537132356
12 norm of gradients: 0.40443988519170615
12 norm of weights: 6.126978190342486
-----
Iteration Number: 2912
Loss: 47.290138947485715
12 norm of gradients: 0.40431520126387865
12 norm of weights: 6.126862074874536
-----
Iteration Number: 2913
Loss: 47.28201630297999
12 norm of gradients: 0.40419059586108236
12 norm of weights: 6.126745981536176
-----
Iteration Number: 2914
Loss: 47.27389859901926
12 norm of gradients: 0.4040660689369586
12 norm of weights: 6.126629910312713
-----
Iteration Number: 2915
Loss: 47.26578583098275
12 norm of gradients: 0.40394162044523146
12 norm of weights: 6.126513861189468
-----
Iteration Number: 2916
Loss: 47.257677994286496
12 norm of gradients: 0.40381725033970745
12 norm of weights: 6.1263978341517795
-----
Iteration Number: 2917
Loss: 47.24957508431846
12 norm of gradients: 0.403692958574275
12 norm of weights: 6.126281829184993
-----
Iteration Number: 2918
Loss: 47.241477096498286
12 norm of gradients: 0.4035687451029043
12 norm of weights: 6.1261658462744695
-----
Iteration Number: 2919
Loss: 47.233384026258776
12 norm of gradients: 0.40344460987964603
12 norm of weights: 6.1260498854055845
-----
Iteration Number: 2920

```

```

Loss: 47.225295868991715
12 norm of gradients: 0.40332055285863117
12 norm of weights: 6.125933946563723
-----
Iteration Number: 2921
Loss: 47.21721262012912
12 norm of gradients: 0.40319657399407094
12 norm of weights: 6.125818029734286
-----
Iteration Number: 2922
Loss: 47.209134275101356
12 norm of gradients: 0.4030726732402555
12 norm of weights: 6.125702134902687
-----
Iteration Number: 2923
Loss: 47.20106082933138
12 norm of gradients: 0.40294885055155405
12 norm of weights: 6.125586262054352
-----
Iteration Number: 2924
Loss: 47.19299227826747
12 norm of gradients: 0.40282510588241405
12 norm of weights: 6.125470411174718
-----
Iteration Number: 2925
Loss: 47.18492861733466
12 norm of gradients: 0.4027014391873605
12 norm of weights: 6.125354582249241
-----
Iteration Number: 2926
Loss: 47.17686984199794
12 norm of gradients: 0.402577850420996
12 norm of weights: 6.125238775263382
-----
Iteration Number: 2927
Loss: 47.1688159477139
12 norm of gradients: 0.40245433953799953
12 norm of weights: 6.125122990202621
-----
Iteration Number: 2928
Loss: 47.16076692992613
12 norm of gradients: 0.40233090649312686
12 norm of weights: 6.125007227052449
-----
Iteration Number: 2929
Loss: 47.15272278408172
12 norm of gradients: 0.40220755124120905
12 norm of weights: 6.1248914857983685

```



```

-----
Iteration Number: 2930
Loss: 47.14468350566417
l2 norm of gradients: 0.40208427373715233
l2 norm of weights: 6.1247757664258975
-----
Iteration Number: 2931
Loss: 47.13664909013545
l2 norm of gradients: 0.4019610739359382
l2 norm of weights: 6.124660068920567
-----
Iteration Number: 2932
Loss: 47.128619532965004
l2 norm of gradients: 0.4018379517926219
l2 norm of weights: 6.124544393267916
-----
Iteration Number: 2933
Loss: 47.120594829644624
l2 norm of gradients: 0.4017149072623325
l2 norm of weights: 6.124428739453504
-----
Iteration Number: 2934
Loss: 47.112574975655484
l2 norm of gradients: 0.4015919403002726
l2 norm of weights: 6.124313107462898
-----
Iteration Number: 2935
Loss: 47.104559966474376
l2 norm of gradients: 0.40146905086171725
l2 norm of weights: 6.124197497281679
-----
Iteration Number: 2936
Loss: 47.09654979760014
l2 norm of gradients: 0.4013462389020138
l2 norm of weights: 6.124081908895443
-----
Iteration Number: 2937
Loss: 47.088544464513475
l2 norm of gradients: 0.40122350437658155
l2 norm of weights: 6.123966342289795
-----
Iteration Number: 2938
Loss: 47.080543962762654
l2 norm of gradients: 0.4011008472409106
l2 norm of weights: 6.123850797450356
-----
Iteration Number: 2939
Loss: 47.07254828780224

```

```

12 norm of gradients: 0.40097826745056264
12 norm of weights: 6.123735274362761
-----
Iteration Number: 2940
Loss: 47.06455743517128
12 norm of gradients: 0.40085576496116887
12 norm of weights: 6.1236197730126545
-----
Iteration Number: 2941
Loss: 47.056571400367595
12 norm of gradients: 0.4007333397284306
12 norm of weights: 6.123504293385695
-----
Iteration Number: 2942
Loss: 47.04859017892183
12 norm of gradients: 0.4006109917081186
12 norm of weights: 6.123388835467554
-----
Iteration Number: 2943
Loss: 47.04061376634821
12 norm of gradients: 0.4004887208560725
12 norm of weights: 6.123273399243918
-----
Iteration Number: 2944
Loss: 47.03264215818423
12 norm of gradients: 0.4003665271281999
12 norm of weights: 6.123157984700482
-----
Iteration Number: 2945
Loss: 47.02467534994976
12 norm of gradients: 0.4002444104804769
12 norm of weights: 6.1230425918229585
-----
Iteration Number: 2946
Loss: 47.01671333722102
12 norm of gradients: 0.4001223708689464
12 norm of weights: 6.12292722059707
-----
Iteration Number: 2947
Loss: 47.00875611549329
12 norm of gradients: 0.4000004082497188
12 norm of weights: 6.12281187100855
-----
Iteration Number: 2948
Loss: 47.00080368031808
12 norm of gradients: 0.3998785225789706
12 norm of weights: 6.122696543043151
-----

```

Iteration Number: 2949
Loss: 46.99285602727537
l2 norm of gradients: 0.3997567138129445
l2 norm of weights: 6.122581236686633

Iteration Number: 2950
Loss: 46.984913151885934
l2 norm of gradients: 0.3996349819079484
l2 norm of weights: 6.122465951924771

Iteration Number: 2951
Loss: 46.97697504974748
l2 norm of gradients: 0.3995133268203559
l2 norm of weights: 6.122350688743351

Iteration Number: 2952
Loss: 46.96904171640317
l2 norm of gradients: 0.3993917485066045
l2 norm of weights: 6.122235447128174

Iteration Number: 2953
Loss: 46.96111314741418
l2 norm of gradients: 0.3992702469231964
l2 norm of weights: 6.122120227065052

Iteration Number: 2954
Loss: 46.95318933836793
l2 norm of gradients: 0.39914882202669716
l2 norm of weights: 6.122005028539812

Iteration Number: 2955
Loss: 46.9452702848214
l2 norm of gradients: 0.39902747377373554
l2 norm of weights: 6.121889851538293

Iteration Number: 2956
Loss: 46.93735598238384
l2 norm of gradients: 0.3989062021210034
l2 norm of weights: 6.1217746960463435

Iteration Number: 2957
Loss: 46.929446426620615
l2 norm of gradients: 0.3987850070252543
l2 norm of weights: 6.121659562049829

Iteration Number: 2958
Loss: 46.92154161313751
l2 norm of gradients: 0.39866388844330447

```

12 norm of weights: 6.1215444495346265
-----
Iteration Number: 2959
Loss: 46.91364153751784
12 norm of gradients: 0.39854284633203063
12 norm of weights: 6.121429358486625
-----
Iteration Number: 2960
Loss: 46.905746195369346
12 norm of gradients: 0.39842188064837114
12 norm of weights: 6.121314288891727
-----
Iteration Number: 2961
Loss: 46.89785558228486
12 norm of gradients: 0.3983009913493246
12 norm of weights: 6.121199240735848
-----
Iteration Number: 2962
Loss: 46.88996969388523
12 norm of gradients: 0.3981801783919495
12 norm of weights: 6.121084214004916
-----
Iteration Number: 2963
Loss: 46.882088525767465
12 norm of gradients: 0.39805944173336416
12 norm of weights: 6.120969208684869
-----
Iteration Number: 2964
Loss: 46.874212073570035
12 norm of gradients: 0.39793878133074606
12 norm of weights: 6.120854224761663
-----
Iteration Number: 2965
Loss: 46.866340332898154
12 norm of gradients: 0.39781819714133104
12 norm of weights: 6.120739262221262
-----
Iteration Number: 2966
Loss: 46.85847329936743
12 norm of gradients: 0.39769768912241393
12 norm of weights: 6.120624321049647
-----
Iteration Number: 2967
Loss: 46.85061096864806
12 norm of gradients: 0.3975772572313466
12 norm of weights: 6.120509401232805
-----
Iteration Number: 2968

```

Loss: 46.842753336324435
l2 norm of gradients: 0.3974569014255388
l2 norm of weights: 6.120394502756746

Iteration Number: 2969
Loss: 46.8349003980922
l2 norm of gradients: 0.39733662166245703
l2 norm of weights: 6.120279625607482

Iteration Number: 2970
Loss: 46.827052149521506
l2 norm of gradients: 0.39721641789962436
l2 norm of weights: 6.120164769771044

Iteration Number: 2971
Loss: 46.81920858633114
l2 norm of gradients: 0.3970962900946201
l2 norm of weights: 6.120049935233474

Iteration Number: 2972
Loss: 46.8113697041304
l2 norm of gradients: 0.39697623820507894
l2 norm of weights: 6.119935121980827

Iteration Number: 2973
Loss: 46.80353549857918
l2 norm of gradients: 0.3968562621886912
l2 norm of weights: 6.11982032999917

Iteration Number: 2974
Loss: 46.79570596535922
l2 norm of gradients: 0.3967363620032015
l2 norm of weights: 6.119705559274584

Iteration Number: 2975
Loss: 46.78788110012416
l2 norm of gradients: 0.39661653760640925
l2 norm of weights: 6.11959080979316

Iteration Number: 2976
Loss: 46.780060898519785
l2 norm of gradients: 0.3964967889561677
l2 norm of weights: 6.119476081541006

Iteration Number: 2977
Loss: 46.77224535627115
l2 norm of gradients: 0.3963771160103835
l2 norm of weights: 6.119361374504236

```

-----
Iteration Number: 2978
Loss: 46.76443446901379
l2 norm of gradients: 0.3962575187270165
l2 norm of weights: 6.119246688668985
-----
Iteration Number: 2979
Loss: 46.75662823241996
l2 norm of gradients: 0.3961379970640794
l2 norm of weights: 6.119132024021393
-----
Iteration Number: 2980
Loss: 46.748826642212386
l2 norm of gradients: 0.39601855097963706
l2 norm of weights: 6.11901738054762
-----
Iteration Number: 2981
Loss: 46.74102969406653
l2 norm of gradients: 0.39589918043180605
l2 norm of weights: 6.118902758233829
-----
Iteration Number: 2982
Loss: 46.733237383694856
l2 norm of gradients: 0.39577988537875464
l2 norm of weights: 6.118788157066206
-----
Iteration Number: 2983
Loss: 46.725449706735034
l2 norm of gradients: 0.39566066577870185
l2 norm of weights: 6.1186735770309415
-----
Iteration Number: 2984
Loss: 46.717666658952204
l2 norm of gradients: 0.3955415215899178
l2 norm of weights: 6.118559018114243
-----
Iteration Number: 2985
Loss: 46.70988823604412
l2 norm of gradients: 0.3954224527707222
l2 norm of weights: 6.11844448030233
-----
Iteration Number: 2986
Loss: 46.702114433704054
l2 norm of gradients: 0.39530345927948507
l2 norm of weights: 6.1183299635814326
-----
Iteration Number: 2987
Loss: 46.69434524766634

```

```

12 norm of gradients: 0.3951845410746256
12 norm of weights: 6.118215467937796
-----
Iteration Number: 2988
Loss: 46.686580673641515
12 norm of gradients: 0.39506569811461206
12 norm of weights: 6.118100993357677
-----
Iteration Number: 2989
Loss: 46.678820707341046
12 norm of gradients: 0.39494693035796136
12 norm of weights: 6.117986539827344
-----
Iteration Number: 2990
Loss: 46.671065344507326
12 norm of gradients: 0.39482823776323844
12 norm of weights: 6.1178721073330795
-----
Iteration Number: 2991
Loss: 46.66331458087651
12 norm of gradients: 0.3947096202890562
12 norm of weights: 6.117757695861178
-----
Iteration Number: 2992
Loss: 46.655568412192174
12 norm of gradients: 0.394591077894075
12 norm of weights: 6.1176433053979435
-----
Iteration Number: 2993
Loss: 46.64782683417142
12 norm of gradients: 0.39447261053700194
12 norm of weights: 6.117528935929699
-----
Iteration Number: 2994
Loss: 46.64008984257402
12 norm of gradients: 0.394354218176591
12 norm of weights: 6.117414587442775
-----
Iteration Number: 2995
Loss: 46.632357433146645
12 norm of gradients: 0.39423590077164233
12 norm of weights: 6.117300259923517
-----
Iteration Number: 2996
Loss: 46.62462960164316
12 norm of gradients: 0.39411765828100165
12 norm of weights: 6.1171859533582795
-----

```

Iteration Number: 2997
Loss: 46.616906343843624
l2 norm of gradients: 0.3939994906635607
l2 norm of weights: 6.117071667733434

Iteration Number: 2998
Loss: 46.60918765547413
l2 norm of gradients: 0.39388139787825555
l2 norm of weights: 6.116957403035361

Iteration Number: 2999
Loss: 46.601473532311715
l2 norm of gradients: 0.3937633798840677
l2 norm of weights: 6.116843159250457

Iteration Number: 3000
Loss: 46.59376397016377
l2 norm of gradients: 0.3936454366400222
l2 norm of weights: 6.116728936365127

Iteration Number: 3001
Loss: 46.58605896474652
l2 norm of gradients: 0.3935275681051885
l2 norm of weights: 6.1166147343657915

Iteration Number: 3002
Loss: 46.57835851187578
l2 norm of gradients: 0.39340977423867957
l2 norm of weights: 6.1165005532388825

Iteration Number: 3003
Loss: 46.57066260732238
l2 norm of gradients: 0.3932920549996511
l2 norm of weights: 6.116386392970844

Iteration Number: 3004
Loss: 46.56297124686627
l2 norm of gradients: 0.39317441034730216
l2 norm of weights: 6.116272253548132

Iteration Number: 3005
Loss: 46.55528442633297
l2 norm of gradients: 0.39305684024087384
l2 norm of weights: 6.116158134957217

Iteration Number: 3006
Loss: 46.54760214145495
l2 norm of gradients: 0.3929393446396492


```

12 norm of weights: 6.116044037184581
-----
Iteration Number: 3007
Loss: 46.53992438812333
12 norm of gradients: 0.39282192350295303
12 norm of weights: 6.115929960216718
-----
Iteration Number: 3008
Loss: 46.532251162056724
12 norm of gradients: 0.3927045767901515
12 norm of weights: 6.115815904040134
-----
Iteration Number: 3009
Loss: 46.52458245907274
12 norm of gradients: 0.3925873044606518
12 norm of weights: 6.115701868641348
-----
Iteration Number: 3010
Loss: 46.51691827504942
12 norm of gradients: 0.39247010647390135
12 norm of weights: 6.115587854006892
-----
Iteration Number: 3011
Loss: 46.5092586057342
12 norm of gradients: 0.39235298278938774
12 norm of weights: 6.115473860123311
-----
Iteration Number: 3012
Loss: 46.50160344699585
12 norm of gradients: 0.39223593336663887
12 norm of weights: 6.11535988697716
-----
Iteration Number: 3013
Loss: 46.49395279461609
12 norm of gradients: 0.39211895816522135
12 norm of weights: 6.1152459345550065
-----
Iteration Number: 3014
Loss: 46.48630664444332
12 norm of gradients: 0.3920020571447417
12 norm of weights: 6.115132002843434
-----
Iteration Number: 3015
Loss: 46.47866499231522
12 norm of gradients: 0.39188523026484434
12 norm of weights: 6.115018091829036
-----
Iteration Number: 3016

```

```
Loss: 46.47102783405628
l2 norm of gradients: 0.3917684774852128
l2 norm of weights: 6.114904201498417
-----
Iteration Number: 3017
Loss: 46.46339516552926
l2 norm of gradients: 0.3916517987655681
l2 norm of weights: 6.114790331838195
-----
Iteration Number: 3018
Loss: 46.45576698255898
l2 norm of gradients: 0.39153519406566933
l2 norm of weights: 6.114676482835003
-----
Iteration Number: 3019
Loss: 46.44814328098573
l2 norm of gradients: 0.3914186633453125
l2 norm of weights: 6.114562654475482
-----
Iteration Number: 3020
Loss: 46.44052405668068
l2 norm of gradients: 0.3913022065643309
l2 norm of weights: 6.114448846746289
-----
Iteration Number: 3021
Loss: 46.43290930550049
l2 norm of gradients: 0.39118582368259436
l2 norm of weights: 6.11433505963409
-----
Iteration Number: 3022
Loss: 46.42529902331061
l2 norm of gradients: 0.3910695146600088
l2 norm of weights: 6.114221293125566
-----
Iteration Number: 3023
Loss: 46.41769320593591
l2 norm of gradients: 0.3909532794565161
l2 norm of weights: 6.11410754720741
-----
Iteration Number: 3024
Loss: 46.41009184930132
l2 norm of gradients: 0.390837118032094
l2 norm of weights: 6.113993821866326
-----
Iteration Number: 3025
Loss: 46.4024949492492
l2 norm of gradients: 0.39072103034675515
l2 norm of weights: 6.113880117089032
```

```

-----
Iteration Number: 3026
Loss: 46.39490250166223
l2 norm of gradients: 0.3906050163605473
l2 norm of weights: 6.113766432862256
-----
Iteration Number: 3027
Loss: 46.387314502424225
l2 norm of gradients: 0.39048907603355254
l2 norm of weights: 6.113652769172742
-----
Iteration Number: 3028
Loss: 46.37973094740501
l2 norm of gradients: 0.3903732093258875
l2 norm of weights: 6.113539126007241
-----
Iteration Number: 3029
Loss: 46.37215183253283
l2 norm of gradients: 0.39025741619770227
l2 norm of weights: 6.113425503352523
-----
Iteration Number: 3030
Loss: 46.36457715366341
l2 norm of gradients: 0.3901416966091808
l2 norm of weights: 6.113311901195364
-----
Iteration Number: 3031
Loss: 46.35700690669016
l2 norm of gradients: 0.3900260505205401
l2 norm of weights: 6.113198319522556
-----
Iteration Number: 3032
Loss: 46.34944108754543
l2 norm of gradients: 0.3899104778920301
l2 norm of weights: 6.113084758320902
-----
Iteration Number: 3033
Loss: 46.341879692112684
l2 norm of gradients: 0.38979497868393337
l2 norm of weights: 6.1129712175772175
-----
Iteration Number: 3034
Loss: 46.33432271631539
l2 norm of gradients: 0.38967955285656447
l2 norm of weights: 6.112857697278331
-----
Iteration Number: 3035
Loss: 46.32677015606304

```

```

12 norm of gradients: 0.38956420037026995
12 norm of weights: 6.112744197411081
-----
Iteration Number: 3036
Loss: 46.31922200724414
12 norm of gradients: 0.38944892118542807
12 norm of weights: 6.112630717962321
-----
Iteration Number: 3037
Loss: 46.31167826583332
12 norm of gradients: 0.3893337152624481
12 norm of weights: 6.112517258918916
-----
Iteration Number: 3038
Loss: 46.304138927721475
12 norm of gradients: 0.3892185825617703
12 norm of weights: 6.112403820267742
-----
Iteration Number: 3039
Loss: 46.29660398883973
12 norm of gradients: 0.38910352304386575
12 norm of weights: 6.112290401995689
-----
Iteration Number: 3040
Loss: 46.28907344512227
12 norm of gradients: 0.3889885366692356
12 norm of weights: 6.1121770040896575
-----
Iteration Number: 3041
Loss: 46.28154729253194
12 norm of gradients: 0.38887362339841086
12 norm of weights: 6.11206362653656
-----
Iteration Number: 3042
Loss: 46.27402552696974
12 norm of gradients: 0.3887587831919524
12 norm of weights: 6.111950269323325
-----
Iteration Number: 3043
Loss: 46.266508144393704
12 norm of gradients: 0.3886440160104504
12 norm of weights: 6.111836932436888
-----
Iteration Number: 3044
Loss: 46.25899514077251
12 norm of gradients: 0.3885293218145237
12 norm of weights: 6.1117236158642
-----

```

Iteration Number: 3045
Loss: 46.251486512030255
l2 norm of gradients: 0.3884147005648205
l2 norm of weights: 6.111610319592224

Iteration Number: 3046
Loss: 46.243982254155966
l2 norm of gradients: 0.3883001522220168
l2 norm of weights: 6.111497043607933

Iteration Number: 3047
Loss: 46.23648236309063
l2 norm of gradients: 0.38818567674681714
l2 norm of weights: 6.1113837878983155

Iteration Number: 3048
Loss: 46.22898683478847
l2 norm of gradients: 0.3880712740999534
l2 norm of weights: 6.111270552450369

Iteration Number: 3049
Loss: 46.22149566523981
l2 norm of gradients: 0.3879569442421852
l2 norm of weights: 6.111157337251106

Iteration Number: 3050
Loss: 46.21400885040929
l2 norm of gradients: 0.38784268713429937
l2 norm of weights: 6.111044142287548

Iteration Number: 3051
Loss: 46.20652638628757
l2 norm of gradients: 0.3877285027371097
l2 norm of weights: 6.110930967546732

Iteration Number: 3052
Loss: 46.19904826882927
l2 norm of gradients: 0.38761439101145584
l2 norm of weights: 6.110817813015705

Iteration Number: 3053
Loss: 46.19157449404614
l2 norm of gradients: 0.3875003519182048
l2 norm of weights: 6.110704678681527

Iteration Number: 3054
Loss: 46.18410505788657
l2 norm of gradients: 0.3873863854182489

```

12 norm of weights: 6.110591564531269
-----
Iteration Number: 3055
Loss: 46.17663995637706
12 norm of gradients: 0.3872724914725061
12 norm of weights: 6.1104784705520165
-----
Iteration Number: 3056
Loss: 46.16917918552998
12 norm of gradients: 0.3871586700419199
12 norm of weights: 6.110365396730864
-----
Iteration Number: 3057
Loss: 46.16172274128251
12 norm of gradients: 0.38704492108745897
12 norm of weights: 6.1102523430549205
-----
Iteration Number: 3058
Loss: 46.15427061970677
12 norm of gradients: 0.3869312445701169
12 norm of weights: 6.110139309511307
-----
Iteration Number: 3059
Loss: 46.14682281676678
12 norm of gradients: 0.38681764045091127
12 norm of weights: 6.110026296087154
-----
Iteration Number: 3060
Loss: 46.13937932850675
12 norm of gradients: 0.38670410869088434
12 norm of weights: 6.109913302769608
-----
Iteration Number: 3061
Loss: 46.13194015091613
12 norm of gradients: 0.38659064925110204
12 norm of weights: 6.109800329545825
-----
Iteration Number: 3062
Loss: 46.12450528001479
12 norm of gradients: 0.38647726209265426
12 norm of weights: 6.1096873764029755
-----
Iteration Number: 3063
Loss: 46.117074711831094
12 norm of gradients: 0.38636394717665395
12 norm of weights: 6.109574443328237
-----
Iteration Number: 3064

```

Loss: 46.10964844240206
l2 norm of gradients: 0.3862507044642374
l2 norm of weights: 6.109461530308804

Iteration Number: 3065
Loss: 46.10222646776008
l2 norm of gradients: 0.38613753391656336
l2 norm of weights: 6.109348637331882

Iteration Number: 3066
Loss: 46.094808783918474
l2 norm of gradients: 0.3860244354948134
l2 norm of weights: 6.109235764384687

Iteration Number: 3067
Loss: 46.08739538694339
l2 norm of gradients: 0.38591140916019123
l2 norm of weights: 6.109122911454449

Iteration Number: 3068
Loss: 46.07998627286436
l2 norm of gradients: 0.3857984548739226
l2 norm of weights: 6.109010078528409

Iteration Number: 3069
Loss: 46.07258143773228
l2 norm of gradients: 0.38568557259725483
l2 norm of weights: 6.10889726559382

Iteration Number: 3070
Loss: 46.06518087759783
l2 norm of gradients: 0.38557276229145665
l2 norm of weights: 6.1087844726379465

Iteration Number: 3071
Loss: 46.05778458849049
l2 norm of gradients: 0.38546002391781814
l2 norm of weights: 6.108671699648067

Iteration Number: 3072
Loss: 46.050392566496235
l2 norm of gradients: 0.38534735743765
l2 norm of weights: 6.1085589466114705

Iteration Number: 3073
Loss: 46.04300480768088
l2 norm of gradients: 0.3852347628122838
l2 norm of weights: 6.1084462135154585

```

-----
Iteration Number: 3074
Loss: 46.035621308089056
l2 norm of gradients: 0.3851222400030712
l2 norm of weights: 6.108333500347342
-----
Iteration Number: 3075
Loss: 46.02824206381995
l2 norm of gradients: 0.3850097889713841
l2 norm of weights: 6.108220807094451
-----
Iteration Number: 3076
Loss: 46.02086707090268
l2 norm of gradients: 0.3848974096786143
l2 norm of weights: 6.108108133744118
-----
Iteration Number: 3077
Loss: 46.013496325458085
l2 norm of gradients: 0.3847851020861727
l2 norm of weights: 6.1079954802836935
-----
Iteration Number: 3078
Loss: 46.006129823544725
l2 norm of gradients: 0.3846728661554901
l2 norm of weights: 6.107882846700541
-----
Iteration Number: 3079
Loss: 45.998767561258155
l2 norm of gradients: 0.384560701848016
l2 norm of weights: 6.107770232982031
-----
Iteration Number: 3080
Loss: 45.991409534670396
l2 norm of gradients: 0.3844486091252185
l2 norm of weights: 6.107657639115551
-----
Iteration Number: 3081
Loss: 45.984055739893925
l2 norm of gradients: 0.3843365879485845
l2 norm of weights: 6.107545065088496
-----
Iteration Number: 3082
Loss: 45.97670617300913
l2 norm of gradients: 0.38422463827961917
l2 norm of weights: 6.107432510888278
-----
Iteration Number: 3083
Loss: 45.96936083014598

```



```

12 norm of gradients: 0.3841127600798455
12 norm of weights: 6.107319976502315
-----
Iteration Number: 3084
Loss: 45.96201970736119
12 norm of gradients: 0.3840009533108042
12 norm of weights: 6.107207461918041
-----
Iteration Number: 3085
Loss: 45.95468280077827
12 norm of gradients: 0.38388921793405373
12 norm of weights: 6.107094967122902
-----
Iteration Number: 3086
Loss: 45.94735010654577
12 norm of gradients: 0.38377755391116974
12 norm of weights: 6.106982492104354
-----
Iteration Number: 3087
Loss: 45.940021620734285
12 norm of gradients: 0.3836659612037445
12 norm of weights: 6.106870036849866
-----
Iteration Number: 3088
Loss: 45.932697339496485
12 norm of gradients: 0.38355443977338755
12 norm of weights: 6.106757601346918
-----
Iteration Number: 3089
Loss: 45.92537725890727
12 norm of gradients: 0.38344298958172474
12 norm of weights: 6.1066451855830035
-----
Iteration Number: 3090
Loss: 45.918061375159354
12 norm of gradients: 0.38333161059039805
12 norm of weights: 6.106532789545628
-----
Iteration Number: 3091
Loss: 45.91074968433099
12 norm of gradients: 0.3832203027610655
12 norm of weights: 6.106420413222306
-----
Iteration Number: 3092
Loss: 45.90344218257408
12 norm of gradients: 0.383109066055401
12 norm of weights: 6.106308056600566
-----

```

Iteration Number: 3093
Loss: 45.89613886603043
l2 norm of gradients: 0.3829979004350943
l2 norm of weights: 6.10619571966795

Iteration Number: 3094
Loss: 45.8888397308332
l2 norm of gradients: 0.3828868058618497
l2 norm of weights: 6.106083402412008

Iteration Number: 3095
Loss: 45.88154477312388
l2 norm of gradients: 0.3827757822973873
l2 norm of weights: 6.105971104820304

Iteration Number: 3096
Loss: 45.874253989071285
l2 norm of gradients: 0.38266482970344146
l2 norm of weights: 6.105858826880415

Iteration Number: 3097
Loss: 45.86696737481306
l2 norm of gradients: 0.3825539480417617
l2 norm of weights: 6.105746568579929

Iteration Number: 3098
Loss: 45.85968492650953
l2 norm of gradients: 0.3824431372741114
l2 norm of weights: 6.105634329906444

Iteration Number: 3099
Loss: 45.852406640344455
l2 norm of gradients: 0.38233239736226826
l2 norm of weights: 6.105522110847572

Iteration Number: 3100
Loss: 45.845132512418914
l2 norm of gradients: 0.3822217282680241
l2 norm of weights: 6.105409911390936

Iteration Number: 3101
Loss: 45.83786253896802
l2 norm of gradients: 0.38211112995318397
l2 norm of weights: 6.105297731524171

Iteration Number: 3102
Loss: 45.83059671612479
l2 norm of gradients: 0.38200060237956673

```

12 norm of weights: 6.105185571234924
-----
Iteration Number: 3103
Loss: 45.82333504009171
12 norm of gradients: 0.38189014550900435
12 norm of weights: 6.1050734305108545
-----
Iteration Number: 3104
Loss: 45.816077507001886
12 norm of gradients: 0.3817797593033417
12 norm of weights: 6.1049613093396315
-----
Iteration Number: 3105
Loss: 45.80882411307926
12 norm of gradients: 0.38166944372443656
12 norm of weights: 6.104849207708938
-----
Iteration Number: 3106
Loss: 45.8015748545001
12 norm of gradients: 0.38155919873415917
12 norm of weights: 6.104737125606467
-----
Iteration Number: 3107
Loss: 45.794329727455036
12 norm of gradients: 0.3814490242943922
12 norm of weights: 6.104625063019926
-----
Iteration Number: 3108
Loss: 45.787088728145385
12 norm of gradients: 0.3813389203670304
12 norm of weights: 6.1045130199370305
-----
Iteration Number: 3109
Loss: 45.779851852750056
12 norm of gradients: 0.38122888691398027
12 norm of weights: 6.104400996345512
-----
Iteration Number: 3110
Loss: 45.77261909748798
12 norm of gradients: 0.3811189238971604
12 norm of weights: 6.104288992233111
-----
Iteration Number: 3111
Loss: 45.765390458532195
12 norm of gradients: 0.3810090312785002
12 norm of weights: 6.104177007587579
-----
Iteration Number: 3112

```

Loss: 45.758165932133714
l2 norm of gradients: 0.38089920901994095
l2 norm of weights: 6.104065042396682

Iteration Number: 3113
Loss: 45.75094551448798
l2 norm of gradients: 0.3807894570834346
l2 norm of weights: 6.103953096648198

Iteration Number: 3114
Loss: 45.74372920180859
l2 norm of gradients: 0.38067977543094406
l2 norm of weights: 6.103841170329912

Iteration Number: 3115
Loss: 45.736516990305475
l2 norm of gradients: 0.3805701640244429
l2 norm of weights: 6.103729263429625

Iteration Number: 3116
Loss: 45.72930887621049
l2 norm of gradients: 0.3804606228259152
l2 norm of weights: 6.103617375935148

Iteration Number: 3117
Loss: 45.722104855762794
l2 norm of gradients: 0.3803511517973549
l2 norm of weights: 6.103505507834306

Iteration Number: 3118
Loss: 45.714904925181614
l2 norm of gradients: 0.3802417509007661
l2 norm of weights: 6.103393659114932

Iteration Number: 3119
Loss: 45.70770908069335
l2 norm of gradients: 0.380132420098163
l2 norm of weights: 6.103281829764875

Iteration Number: 3120
Loss: 45.7005173185516
l2 norm of gradients: 0.38002315935156905
l2 norm of weights: 6.103170019771991

Iteration Number: 3121
Loss: 45.69332963499065
l2 norm of gradients: 0.3799139686230173
l2 norm of weights: 6.1030582291241515

```

-----
Iteration Number: 3122
Loss: 45.68614602624419
l2 norm of gradients: 0.3798048478745497
l2 norm of weights: 6.102946457809239
-----
Iteration Number: 3123
Loss: 45.6789664886048
l2 norm of gradients: 0.37969579706821754
l2 norm of weights: 6.102834705815144
-----
Iteration Number: 3124
Loss: 45.671791018260365
l2 norm of gradients: 0.37958681616608075
l2 norm of weights: 6.102722973129775
-----
Iteration Number: 3125
Loss: 45.66461961151847
l2 norm of gradients: 0.37947790513020785
l2 norm of weights: 6.102611259741047
-----
Iteration Number: 3126
Loss: 45.6574522646182
l2 norm of gradients: 0.37936906392267605
l2 norm of weights: 6.1024995656368874
-----
Iteration Number: 3127
Loss: 45.6502889738217
l2 norm of gradients: 0.37926029250557025
l2 norm of weights: 6.10238789080524
-----
Iteration Number: 3128
Loss: 45.643129735387106
l2 norm of gradients: 0.3791515908409838
l2 norm of weights: 6.1022762352340525
-----
Iteration Number: 3129
Loss: 45.635974545615746
l2 norm of gradients: 0.3790429588910181
l2 norm of weights: 6.10216459891129
-----
Iteration Number: 3130
Loss: 45.62882340075167
l2 norm of gradients: 0.3789343966177813
l2 norm of weights: 6.102052981824928
-----
Iteration Number: 3131
Loss: 45.62167629709805

```

```

12 norm of gradients: 0.37882590398339
12 norm of weights: 6.101941383962952
-----
Iteration Number: 3132
Loss: 45.614533230919285
12 norm of gradients: 0.37871748094996766
12 norm of weights: 6.101829805313361
-----
Iteration Number: 3133
Loss: 45.60739419849987
12 norm of gradients: 0.37860912747964465
12 norm of weights: 6.1017182458641654
-----
Iteration Number: 3134
Loss: 45.60025919613758
12 norm of gradients: 0.3785008435345586
12 norm of weights: 6.101606705603385
-----
Iteration Number: 3135
Loss: 45.593128220125465
12 norm of gradients: 0.3783926290768537
12 norm of weights: 6.101495184519052
-----
Iteration Number: 3136
Loss: 45.58600126675976
12 norm of gradients: 0.37828448406868076
12 norm of weights: 6.101383682599215
-----
Iteration Number: 3137
Loss: 45.578878332328436
12 norm of gradients: 0.37817640847219686
12 norm of weights: 6.101272199831927
-----
Iteration Number: 3138
Loss: 45.5717594131486
12 norm of gradients: 0.37806840224956534
12 norm of weights: 6.101160736205257
-----
Iteration Number: 3139
Loss: 45.56464450549623
12 norm of gradients: 0.37796046536295547
12 norm of weights: 6.101049291707283
-----
Iteration Number: 3140
Loss: 45.55753360572806
12 norm of gradients: 0.37785259777454255
12 norm of weights: 6.100937866326097
-----

```

Iteration Number: 3141
Loss: 45.55042671013519
l2 norm of gradients: 0.37774479944650735
l2 norm of weights: 6.1008264600498014

Iteration Number: 3142
Loss: 45.54332381503972
l2 norm of gradients: 0.3776370703410361
l2 norm of weights: 6.100715072866509

Iteration Number: 3143
Loss: 45.53622491674664
l2 norm of gradients: 0.3775294104203206
l2 norm of weights: 6.100603704764347

Iteration Number: 3144
Loss: 45.52913001158869
l2 norm of gradients: 0.37742181964655763
l2 norm of weights: 6.100492355731451

Iteration Number: 3145
Loss: 45.52203909591576
l2 norm of gradients: 0.3773142979819489
l2 norm of weights: 6.100381025755971

Iteration Number: 3146
Loss: 45.51495216603514
l2 norm of gradients: 0.377206845388701
l2 norm of weights: 6.100269714826066

Iteration Number: 3147
Loss: 45.50786921831116
l2 norm of gradients: 0.37709946182902526
l2 norm of weights: 6.100158422929907

Iteration Number: 3148
Loss: 45.50079024903949
l2 norm of gradients: 0.37699214726513736
l2 norm of weights: 6.100047150055678

Iteration Number: 3149
Loss: 45.49371525459277
l2 norm of gradients: 0.37688490165925714
l2 norm of weights: 6.099935896191574

Iteration Number: 3150
Loss: 45.48664423129495
l2 norm of gradients: 0.37677772497360906

```

12 norm of weights: 6.099824661325799
-----
Iteration Number: 3151
Loss: 45.47957717553605
12 norm of gradients: 0.3766706171704213
12 norm of weights: 6.099713445446573
-----
Iteration Number: 3152
Loss: 45.47251408361935
12 norm of gradients: 0.37656357821192565
12 norm of weights: 6.099602248542124
-----
Iteration Number: 3153
Loss: 45.4654549519413
12 norm of gradients: 0.37645660806035797
12 norm of weights: 6.099491070600692
-----
Iteration Number: 3154
Loss: 45.45839977684561
12 norm of gradients: 0.3763497066779574
12 norm of weights: 6.099379911610528
-----
Iteration Number: 3155
Loss: 45.45134855469941
12 norm of gradients: 0.37624287402696655
12 norm of weights: 6.099268771559899
-----
Iteration Number: 3156
Loss: 45.4443012818507
12 norm of gradients: 0.3761361100696311
12 norm of weights: 6.099157650437075
-----
Iteration Number: 3157
Loss: 45.43725795471595
12 norm of gradients: 0.3760294147682
12 norm of weights: 6.099046548230346
-----
Iteration Number: 3158
Loss: 45.430218569611945
12 norm of gradients: 0.37592278808492496
12 norm of weights: 6.098935464928008
-----
Iteration Number: 3159
Loss: 45.4231831229781
12 norm of gradients: 0.37581622998206016
12 norm of weights: 6.09882440051837
-----
Iteration Number: 3160

```


Loss: 45.4161516111423
l2 norm of gradients: 0.37570974042186306
l2 norm of weights: 6.098713354989753

Iteration Number: 3161
Loss: 45.409124030530926
l2 norm of gradients: 0.37560331936659286
l2 norm of weights: 6.098602328330489

Iteration Number: 3162
Loss: 45.402100377496375
l2 norm of gradients: 0.3754969667785117
l2 norm of weights: 6.09849132052892

Iteration Number: 3163
Loss: 45.39508064847253
l2 norm of gradients: 0.37539068261988306
l2 norm of weights: 6.098380331573402

Iteration Number: 3164
Loss: 45.38806483981442
l2 norm of gradients: 0.37528446685297323
l2 norm of weights: 6.098269361452301

Iteration Number: 3165
Loss: 45.38105294792932
l2 norm of gradients: 0.3751783194400501
l2 norm of weights: 6.098158410153994

Iteration Number: 3166
Loss: 45.37404496924024
l2 norm of gradients: 0.375072240343383
l2 norm of weights: 6.098047477666871

Iteration Number: 3167
Loss: 45.3670409001412
l2 norm of gradients: 0.3749662295252432
l2 norm of weights: 6.09793656397933

Iteration Number: 3168
Loss: 45.36004073703798
l2 norm of gradients: 0.3748602869479032
l2 norm of weights: 6.097825669079783

Iteration Number: 3169
Loss: 45.35304447636989
l2 norm of gradients: 0.374754412573637
l2 norm of weights: 6.097714792956655

```

-----
Iteration Number: 3170
Loss: 45.34605211449854
l2 norm of gradients: 0.3746486063647194
l2 norm of weights: 6.097603935598378
-----
Iteration Number: 3171
Loss: 45.339063647893994
l2 norm of gradients: 0.37454286828342676
l2 norm of weights: 6.097493096993398
-----
Iteration Number: 3172
Loss: 45.33207907297276
l2 norm of gradients: 0.3744371982920359
l2 norm of weights: 6.097382277130173
-----
Iteration Number: 3173
Loss: 45.32509838612868
l2 norm of gradients: 0.37433159635282437
l2 norm of weights: 6.09727147599717
-----
Iteration Number: 3174
Loss: 45.31812158385001
l2 norm of gradients: 0.37422606242807066
l2 norm of weights: 6.097160693582869
-----
Iteration Number: 3175
Loss: 45.31114866251769
l2 norm of gradients: 0.3741205964800536
l2 norm of weights: 6.0970499298757606
-----
Iteration Number: 3176
Loss: 45.30417961859122
l2 norm of gradients: 0.3740151984710523
l2 norm of weights: 6.096939184864347
-----
Iteration Number: 3177
Loss: 45.29721444851877
l2 norm of gradients: 0.3739098683633463
l2 norm of weights: 6.096828458537141
-----
Iteration Number: 3178
Loss: 45.2902531487195
l2 norm of gradients: 0.373804606119215
l2 norm of weights: 6.096717750882668
-----
Iteration Number: 3179
Loss: 45.28329571566882

```

```

12 norm of gradients: 0.37369941170093773
12 norm of weights: 6.096607061889464
-----
Iteration Number: 3180
Loss: 45.27634214582161
12 norm of gradients: 0.3735942850707942
12 norm of weights: 6.096496391546077
-----
Iteration Number: 3181
Loss: 45.26939243559555
12 norm of gradients: 0.3734892261910633
12 norm of weights: 6.096385739841062
-----
Iteration Number: 3182
Loss: 45.2624465814925
12 norm of gradients: 0.37338423502402346
12 norm of weights: 6.0962751067629934
-----
Iteration Number: 3183
Loss: 45.25550457993629
12 norm of gradients: 0.3732793115319532
12 norm of weights: 6.096164492300448
-----
Iteration Number: 3184
Loss: 45.2485664274151
12 norm of gradients: 0.37317445567712976
12 norm of weights: 6.0960538964420214
-----
Iteration Number: 3185
Loss: 45.241632120402215
12 norm of gradients: 0.37306966742183006
12 norm of weights: 6.095943319176316
-----
Iteration Number: 3186
Loss: 45.23470165534818
12 norm of gradients: 0.3729649467283297
12 norm of weights: 6.095832760491945
-----
Iteration Number: 3187
Loss: 45.227775028740446
12 norm of gradients: 0.3728602935589037
12 norm of weights: 6.095722220377535
-----
Iteration Number: 3188
Loss: 45.22085223707403
12 norm of gradients: 0.37275570787582585
12 norm of weights: 6.095611698821724
-----

```

Iteration Number: 3189
Loss: 45.21393327681712
l2 norm of gradients: 0.3726511896413686
l2 norm of weights: 6.095501195813159

Iteration Number: 3190
Loss: 45.207018144457294
l2 norm of gradients: 0.3725467388178031
l2 norm of weights: 6.0953907113405

Iteration Number: 3191
Loss: 45.20010683646676
l2 norm of gradients: 0.3724423553673989
l2 norm of weights: 6.095280245392418

Iteration Number: 3192
Loss: 45.19319934936393
l2 norm of gradients: 0.37233803925242437
l2 norm of weights: 6.0951697979575945

Iteration Number: 3193
Loss: 45.18629567962747
l2 norm of gradients: 0.3722337904351457
l2 norm of weights: 6.095059369024724

Iteration Number: 3194
Loss: 45.17939582377666
l2 norm of gradients: 0.37212960887782776
l2 norm of weights: 6.094948958582507

Iteration Number: 3195
Loss: 45.1724997783041
l2 norm of gradients: 0.37202549454273326
l2 norm of weights: 6.094838566619662

Iteration Number: 3196
Loss: 45.1656075397124
l2 norm of gradients: 0.37192144739212285
l2 norm of weights: 6.094728193124913

Iteration Number: 3197
Loss: 45.15871910449912
l2 norm of gradients: 0.3718174673882551
l2 norm of weights: 6.094617838087001

Iteration Number: 3198
Loss: 45.15183446921561
l2 norm of gradients: 0.37171355449338656

```
12 norm of weights: 6.094507501494672
-----
Iteration Number: 3199
Loss: 45.14495363033233
12 norm of gradients: 0.3716097086697712
12 norm of weights: 6.094397183336686
-----
Iteration Number: 3200
Loss: 45.13807658442114
12 norm of gradients: 0.3715059298796607
12 norm of weights: 6.094286883601815
-----
Iteration Number: 3201
Loss: 45.1312033279534
12 norm of gradients: 0.3714022180853042
12 norm of weights: 6.094176602278841
-----
Iteration Number: 3202
Loss: 45.12433385750059
12 norm of gradients: 0.37129857324894827
12 norm of weights: 6.094066339356556
-----
Iteration Number: 3203
Loss: 45.11746816955441
12 norm of gradients: 0.3711949953328364
12 norm of weights: 6.093956094823766
-----
Iteration Number: 3204
Loss: 45.1106062606885
12 norm of gradients: 0.3710914842992098
12 norm of weights: 6.093845868669285
-----
Iteration Number: 3205
Loss: 45.10374812741259
12 norm of gradients: 0.3709880401103062
12 norm of weights: 6.09373566088194
-----
Iteration Number: 3206
Loss: 45.096893766282896
12 norm of gradients: 0.3708846627283608
12 norm of weights: 6.093625471450569
-----
Iteration Number: 3207
Loss: 45.090043173824334
12 norm of gradients: 0.3707813521156054
12 norm of weights: 6.093515300364017
-----
Iteration Number: 3208
```

```
Loss: 45.08319634659267
l2 norm of gradients: 0.37067810823426856
l2 norm of weights: 6.093405147611149
-----
Iteration Number: 3209
Loss: 45.07635328114973
l2 norm of gradients: 0.3705749310465758
l2 norm of weights: 6.093295013180833
-----
Iteration Number: 3210
Loss: 45.069513974044575
l2 norm of gradients: 0.3704718205147488
l2 norm of weights: 6.09318489706195
-----
Iteration Number: 3211
Loss: 45.06267842183554
l2 norm of gradients: 0.37036877660100614
l2 norm of weights: 6.093074799243395
-----
Iteration Number: 3212
Loss: 45.055846621065555
l2 norm of gradients: 0.3702657992675627
l2 norm of weights: 6.092964719714068
-----
Iteration Number: 3213
Loss: 45.0490185683258
l2 norm of gradients: 0.37016288847662965
l2 norm of weights: 6.092854658462888
-----
Iteration Number: 3214
Loss: 45.04219426016493
l2 norm of gradients: 0.3700600441904144
l2 norm of weights: 6.092744615478779
-----
Iteration Number: 3215
Loss: 45.03537369315603
l2 norm of gradients: 0.36995726637112075
l2 norm of weights: 6.092634590750678
-----
Iteration Number: 3216
Loss: 45.02855686387781
l2 norm of gradients: 0.36985455498094805
l2 norm of weights: 6.092524584267532
-----
Iteration Number: 3217
Loss: 45.021743768929696
l2 norm of gradients: 0.3697519099820921
l2 norm of weights: 6.0924145960183
```

```

-----
Iteration Number: 3218
Loss: 45.01493440485897
l2 norm of gradients: 0.36964933133674466
l2 norm of weights: 6.092304625991953
-----
Iteration Number: 3219
Loss: 45.00812876825543
l2 norm of gradients: 0.36954681900709274
l2 norm of weights: 6.09219467417747
-----
Iteration Number: 3220
Loss: 45.00132685572177
l2 norm of gradients: 0.3694443729553197
l2 norm of weights: 6.092084740563844
-----
Iteration Number: 3221
Loss: 44.994528663855284
l2 norm of gradients: 0.3693419931436042
l2 norm of weights: 6.091974825140077
-----
Iteration Number: 3222
Loss: 44.987734189215985
l2 norm of gradients: 0.36923967953412073
l2 norm of weights: 6.091864927895182
-----
Iteration Number: 3223
Loss: 44.98094342844422
l2 norm of gradients: 0.36913743208903904
l2 norm of weights: 6.091755048818185
-----
Iteration Number: 3224
Loss: 44.974156378103
l2 norm of gradients: 0.3690352507705244
l2 norm of weights: 6.09164518789812
-----
Iteration Number: 3225
Loss: 44.96737303483952
l2 norm of gradients: 0.3689331355407375
l2 norm of weights: 6.091535345124035
-----
Iteration Number: 3226
Loss: 44.96059339520664
l2 norm of gradients: 0.36883108636183426
l2 norm of weights: 6.091425520484986
-----
Iteration Number: 3227
Loss: 44.95381745587901

```

```

12 norm of gradients: 0.36872910319596564
12 norm of weights: 6.091315713970041
-----
Iteration Number: 3228
Loss: 44.94704521341124
12 norm of gradients: 0.3686271860052779
12 norm of weights: 6.091205925568281
-----
Iteration Number: 3229
Loss: 44.94027666446944
12 norm of gradients: 0.36852533475191246
12 norm of weights: 6.091096155268794
-----
Iteration Number: 3230
Loss: 44.933511805634296
12 norm of gradients: 0.36842354939800537
12 norm of weights: 6.090986403060681
-----
Iteration Number: 3231
Loss: 44.92675063355609
12 norm of gradients: 0.3683218299056879
12 norm of weights: 6.0908766689330545
-----
Iteration Number: 3232
Loss: 44.91999314485855
12 norm of gradients: 0.36822017623708597
12 norm of weights: 6.090766952875037
-----
Iteration Number: 3233
Loss: 44.913239336182784
12 norm of gradients: 0.3681185883543204
12 norm of weights: 6.090657254875762
-----
Iteration Number: 3234
Loss: 44.90648920414704
12 norm of gradients: 0.36801706621950675
12 norm of weights: 6.090547574924375
-----
Iteration Number: 3235
Loss: 44.899742745386305
12 norm of gradients: 0.36791560979475496
12 norm of weights: 6.090437913010028
-----
Iteration Number: 3236
Loss: 44.892999956562605
12 norm of gradients: 0.3678142190421698
12 norm of weights: 6.0903282691218905
-----

```


Iteration Number: 3237
Loss: 44.88626083430829
l2 norm of gradients: 0.3677128939238506
l2 norm of weights: 6.090218643249138

Iteration Number: 3238
Loss: 44.87952537525587
l2 norm of gradients: 0.3676116344018909
l2 norm of weights: 6.090109035380957

Iteration Number: 3239
Loss: 44.87279357608086
l2 norm of gradients: 0.3675104404383786
l2 norm of weights: 6.089999445506548

Iteration Number: 3240
Loss: 44.86606543343577
l2 norm of gradients: 0.36740931199539617
l2 norm of weights: 6.089889873615119

Iteration Number: 3241
Loss: 44.859340943967545
l2 norm of gradients: 0.3673082490350203
l2 norm of weights: 6.08978031969589

Iteration Number: 3242
Loss: 44.85262010433318
l2 norm of gradients: 0.36720725151932176
l2 norm of weights: 6.089670783738094

Iteration Number: 3243
Loss: 44.84590291121017
l2 norm of gradients: 0.3671063194103654
l2 norm of weights: 6.08956126573097

Iteration Number: 3244
Loss: 44.83918936125947
l2 norm of gradients: 0.36700545267021056
l2 norm of weights: 6.089451765663772

Iteration Number: 3245
Loss: 44.832479451159536
l2 norm of gradients: 0.36690465126091004
l2 norm of weights: 6.089342283525762

Iteration Number: 3246
Loss: 44.825773177571705
l2 norm of gradients: 0.3668039151445111

```

12 norm of weights: 6.089232819306216
-----
Iteration Number: 3247
Loss: 44.81907053718751
12 norm of gradients: 0.3667032442830545
12 norm of weights: 6.0891233729944165
-----
Iteration Number: 3248
Loss: 44.81237152667678
12 norm of gradients: 0.36660263863857523
12 norm of weights: 6.089013944579659
-----
Iteration Number: 3249
Loss: 44.80567614272031
12 norm of gradients: 0.3665020981731021
12 norm of weights: 6.088904534051253
-----
Iteration Number: 3250
Loss: 44.79898438203633
12 norm of gradients: 0.3664016228486573
12 norm of weights: 6.088795141398512
-----
Iteration Number: 3251
Loss: 44.7922962412884
12 norm of gradients: 0.36630121262725696
12 norm of weights: 6.088685766610764
-----
Iteration Number: 3252
Loss: 44.785611717158844
12 norm of gradients: 0.36620086747091113
12 norm of weights: 6.08857640967735
-----
Iteration Number: 3253
Loss: 44.77893080637042
12 norm of gradients: 0.3661005873416231
12 norm of weights: 6.088467070587615
-----
Iteration Number: 3254
Loss: 44.77225350560321
12 norm of gradients: 0.3660003722013897
12 norm of weights: 6.088357749330922
-----
Iteration Number: 3255
Loss: 44.76557981156763
12 norm of gradients: 0.3659002220122016
12 norm of weights: 6.088248445896641
-----
Iteration Number: 3256

```

```

Loss: 44.758909720993586
l2 norm of gradients: 0.3658001367360427
l2 norm of weights: 6.088139160274151
-----
Iteration Number: 3257
Loss: 44.752243230554086
l2 norm of gradients: 0.36570011633489063
l2 norm of weights: 6.088029892452846
-----
Iteration Number: 3258
Loss: 44.74558033698199
l2 norm of gradients: 0.3656001607707159
l2 norm of weights: 6.087920642422129
-----
Iteration Number: 3259
Loss: 44.73892103701956
l2 norm of gradients: 0.3655002700054827
l2 norm of weights: 6.087811410171411
-----
Iteration Number: 3260
Loss: 44.732265327332875
l2 norm of gradients: 0.3654004440011484
l2 norm of weights: 6.087702195690116
-----
Iteration Number: 3261
Loss: 44.725613204683164
l2 norm of gradients: 0.365300682719664
l2 norm of weights: 6.08759299896768
-----
Iteration Number: 3262
Loss: 44.71896466576608
l2 norm of gradients: 0.36520098612297297
l2 norm of weights: 6.087483819993546
-----
Iteration Number: 3263
Loss: 44.71231970736757
l2 norm of gradients: 0.3651013541730127
l2 norm of weights: 6.087374658757172
-----
Iteration Number: 3264
Loss: 44.70567832615798
l2 norm of gradients: 0.3650017868317132
l2 norm of weights: 6.0872655152480215
-----
Iteration Number: 3265
Loss: 44.69904051890746
l2 norm of gradients: 0.36490228406099795
l2 norm of weights: 6.087156389455575

```

```

-----
Iteration Number: 3266
Loss: 44.692406282354774
l2 norm of gradients: 0.36480284582278316
l2 norm of weights: 6.087047281369316
-----
Iteration Number: 3267
Loss: 44.685775613224905
l2 norm of gradients: 0.3647034720789784
l2 norm of weights: 6.086938190978746
-----
Iteration Number: 3268
Loss: 44.67914850827286
l2 norm of gradients: 0.36460416279148583
l2 norm of weights: 6.0868291182733705
-----
Iteration Number: 3269
Loss: 44.672524964248794
l2 norm of gradients: 0.36450491792220097
l2 norm of weights: 6.086720063242711
-----
Iteration Number: 3270
Loss: 44.665904977923034
l2 norm of gradients: 0.3644057374330121
l2 norm of weights: 6.086611025876297
-----
Iteration Number: 3271
Loss: 44.65928854601908
l2 norm of gradients: 0.36430662128580027
l2 norm of weights: 6.086502006163668
-----
Iteration Number: 3272
Loss: 44.65267566531278
l2 norm of gradients: 0.36420756944243954
l2 norm of weights: 6.086393004094376
-----
Iteration Number: 3273
Loss: 44.646066332588454
l2 norm of gradients: 0.3641085818647967
l2 norm of weights: 6.086284019657981
-----
Iteration Number: 3274
Loss: 44.63946054456081
l2 norm of gradients: 0.3640096585147314
l2 norm of weights: 6.086175052844056
-----
Iteration Number: 3275
Loss: 44.63285829804244

```

```

12 norm of gradients: 0.36391079935409615
12 norm of weights: 6.0860661036421835
-----
Iteration Number: 3276
Loss: 44.62625958976055
12 norm of gradients: 0.36381200434473604
12 norm of weights: 6.085957172041956
-----
Iteration Number: 3277
Loss: 44.6196644165435
12 norm of gradients: 0.3637132734484889
12 norm of weights: 6.085848258032978
-----
Iteration Number: 3278
Loss: 44.613072775123364
12 norm of gradients: 0.3636146066271853
12 norm of weights: 6.085739361604861
-----
Iteration Number: 3279
Loss: 44.606484662314394
12 norm of gradients: 0.3635160038426486
12 norm of weights: 6.0856304827472325
-----
Iteration Number: 3280
Loss: 44.59990007487426
12 norm of gradients: 0.3634174650566943
12 norm of weights: 6.085521621449726
-----
Iteration Number: 3281
Loss: 44.59331900961998
12 norm of gradients: 0.36331899023113107
12 norm of weights: 6.085412777701987
-----
Iteration Number: 3282
Loss: 44.5867414633313
12 norm of gradients: 0.3632205793277599
12 norm of weights: 6.085303951493672
-----
Iteration Number: 3283
Loss: 44.58016743277954
12 norm of gradients: 0.3631222323083744
12 norm of weights: 6.085195142814448
-----
Iteration Number: 3284
Loss: 44.57359691480064
12 norm of gradients: 0.3630239491347607
12 norm of weights: 6.085086351653989
-----

```

Iteration Number: 3285
Loss: 44.56702990614511
l2 norm of gradients: 0.36292572976869747
l2 norm of weights: 6.084977578001985

Iteration Number: 3286
Loss: 44.56046640367457
l2 norm of gradients: 0.3628275741719558
l2 norm of weights: 6.084868821848134

Iteration Number: 3287
Loss: 44.55390640415666
l2 norm of gradients: 0.36272948230629953
l2 norm of weights: 6.084760083182141

Iteration Number: 3288
Loss: 44.5473499044105
l2 norm of gradients: 0.36263145413348463
l2 norm of weights: 6.0846513619937275

Iteration Number: 3289
Loss: 44.540796901247454
l2 norm of gradients: 0.36253348961525966
l2 norm of weights: 6.084542658272622

Iteration Number: 3290
Loss: 44.53424739150726
l2 norm of gradients: 0.3624355887133656
l2 norm of weights: 6.084433972008563

Iteration Number: 3291
Loss: 44.52770137196919
l2 norm of gradients: 0.36233775138953594
l2 norm of weights: 6.084325303191301

Iteration Number: 3292
Loss: 44.52115883947951
l2 norm of gradients: 0.3622399776054965
l2 norm of weights: 6.084216651810596

Iteration Number: 3293
Loss: 44.5146197908741
l2 norm of gradients: 0.36214226732296534
l2 norm of weights: 6.084108017856218

Iteration Number: 3294
Loss: 44.50808422296167
l2 norm of gradients: 0.362044620503653

```
12 norm of weights: 6.083999401317949
-----
Iteration Number: 3295
Loss: 44.50155213257877
12 norm of gradients: 0.36194703710926246
12 norm of weights: 6.0838908021855795
-----
Iteration Number: 3296
Loss: 44.495023516565595
12 norm of gradients: 0.361849517101489
12 norm of weights: 6.083782220448912
-----
Iteration Number: 3297
Loss: 44.48849837175924
12 norm of gradients: 0.36175206044202024
12 norm of weights: 6.083673656097756
-----
Iteration Number: 3298
Loss: 44.48197669501736
12 norm of gradients: 0.36165466709253596
12 norm of weights: 6.083565109121936
-----
Iteration Number: 3299
Loss: 44.47545848313989
12 norm of gradients: 0.36155733701470855
12 norm of weights: 6.083456579511284
-----
Iteration Number: 3300
Loss: 44.46894373303448
12 norm of gradients: 0.3614600701702025
12 norm of weights: 6.083348067255644
-----
Iteration Number: 3301
Loss: 44.46243244149765
12 norm of gradients: 0.36136286652067445
12 norm of weights: 6.083239572344867
-----
Iteration Number: 3302
Loss: 44.45592460542567
12 norm of gradients: 0.36126572602777385
12 norm of weights: 6.083131094768818
-----
Iteration Number: 3303
Loss: 44.449420221637865
12 norm of gradients: 0.36116864865314185
12 norm of weights: 6.083022634517371
-----
Iteration Number: 3304
```

```
Loss: 44.44291928703938
l2 norm of gradients: 0.3610716343584123
l2 norm of weights: 6.082914191580409
-----
Iteration Number: 3305
Loss: 44.43642179846861
l2 norm of gradients: 0.36097468310521086
l2 norm of weights: 6.082805765947828
-----
Iteration Number: 3306
Loss: 44.42992775276501
l2 norm of gradients: 0.3608777948551562
l2 norm of weights: 6.082697357609533
-----
Iteration Number: 3307
Loss: 44.42343714685844
l2 norm of gradients: 0.36078096956985833
l2 norm of weights: 6.082588966555435
-----
Iteration Number: 3308
Loss: 44.416949977568
l2 norm of gradients: 0.3606842072109202
l2 norm of weights: 6.082480592775463
-----
Iteration Number: 3309
Loss: 44.4104662418132
l2 norm of gradients: 0.360587507739937
l2 norm of weights: 6.082372236259552
-----
Iteration Number: 3310
Loss: 44.403985936436484
l2 norm of gradients: 0.36049087111849565
l2 norm of weights: 6.082263896997646
-----
Iteration Number: 3311
Loss: 44.39750905835956
l2 norm of gradients: 0.36039429730817574
l2 norm of weights: 6.082155574979702
-----
Iteration Number: 3312
Loss: 44.39103560442899
l2 norm of gradients: 0.3602977862705492
l2 norm of weights: 6.0820472701956865
-----
Iteration Number: 3313
Loss: 44.38456557155759
l2 norm of gradients: 0.3602013379671798
l2 norm of weights: 6.081938982635575
```



```

-----
Iteration Number: 3314
Loss: 44.37809895663065
l2 norm of gradients: 0.360104952359624
l2 norm of weights: 6.081830712289354
-----
Iteration Number: 3315
Loss: 44.37163575651938
l2 norm of gradients: 0.3600086294094302
l2 norm of weights: 6.08172245914702
-----
Iteration Number: 3316
Loss: 44.365175968182434
l2 norm of gradients: 0.35991236907813934
l2 norm of weights: 6.08161422319858
-----
Iteration Number: 3317
Loss: 44.35871958845382
l2 norm of gradients: 0.35981617132728433
l2 norm of weights: 6.081506004434052
-----
Iteration Number: 3318
Loss: 44.35226661429172
l2 norm of gradients: 0.3597200361183906
l2 norm of weights: 6.081397802843463
-----
Iteration Number: 3319
Loss: 44.345817042581146
l2 norm of gradients: 0.35962396341297564
l2 norm of weights: 6.081289618416849
-----
Iteration Number: 3320
Loss: 44.339370870206125
l2 norm of gradients: 0.35952795317254943
l2 norm of weights: 6.08118145114426
-----
Iteration Number: 3321
Loss: 44.33292809412117
l2 norm of gradients: 0.3594320053586141
l2 norm of weights: 6.081073301015751
-----
Iteration Number: 3322
Loss: 44.32648871122683
l2 norm of gradients: 0.359336119932664
l2 norm of weights: 6.080965168021391
-----
Iteration Number: 3323
Loss: 44.320052718446554

```

```

12 norm of gradients: 0.35924029685618614
12 norm of weights: 6.080857052151258
-----
Iteration Number: 3324
Loss: 44.31362011271738
12 norm of gradients: 0.35914453609065944
12 norm of weights: 6.08074895339544
-----
Iteration Number: 3325
Loss: 44.30719089091467
12 norm of gradients: 0.3590488375975551
12 norm of weights: 6.080640871744035
-----
Iteration Number: 3326
Loss: 44.30076505002076
12 norm of gradients: 0.35895320133833714
12 norm of weights: 6.080532807187151
-----
Iteration Number: 3327
Loss: 44.29434258694029
12 norm of gradients: 0.35885762727446135
12 norm of weights: 6.080424759714907
-----
Iteration Number: 3328
Loss: 44.287923498631194
12 norm of gradients: 0.3587621153673762
12 norm of weights: 6.080316729317431
-----
Iteration Number: 3329
Loss: 44.281507782005974
12 norm of gradients: 0.3586666655785226
12 norm of weights: 6.080208715984861
-----
Iteration Number: 3330
Loss: 44.27509543403748
12 norm of gradients: 0.3585712778693335
12 norm of weights: 6.080100719707347
-----
Iteration Number: 3331
Loss: 44.268686451628675
12 norm of gradients: 0.35847595220123446
12 norm of weights: 6.079992740475046
-----
Iteration Number: 3332
Loss: 44.26228083176213
12 norm of gradients: 0.35838068853564337
12 norm of weights: 6.079884778278127
-----

```

Iteration Number: 3333
Loss: 44.255878571369145
l2 norm of gradients: 0.3582854868339706
l2 norm of weights: 6.079776833106769

Iteration Number: 3334
Loss: 44.24947966740985
l2 norm of gradients: 0.3581903470576188
l2 norm of weights: 6.079668904951161

Iteration Number: 3335
Loss: 44.24308411685869
l2 norm of gradients: 0.35809526916798323
l2 norm of weights: 6.079560993801502

Iteration Number: 3336
Loss: 44.236691916651026
l2 norm of gradients: 0.3580002531264515
l2 norm of weights: 6.079453099648

Iteration Number: 3337
Loss: 44.230303063748984
l2 norm of gradients: 0.3579052988944037
l2 norm of weights: 6.079345222480874

Iteration Number: 3338
Loss: 44.223917555128644
l2 norm of gradients: 0.3578104064332126
l2 norm of weights: 6.079237362290352

Iteration Number: 3339
Loss: 44.21753538775778
l2 norm of gradients: 0.35771557570424306
l2 norm of weights: 6.079129519066674

Iteration Number: 3340
Loss: 44.211156558618555
l2 norm of gradients: 0.35762080666885293
l2 norm of weights: 6.079021692800088

Iteration Number: 3341
Loss: 44.20478106468464
l2 norm of gradients: 0.35752609928839235
l2 norm of weights: 6.078913883480854

Iteration Number: 3342
Loss: 44.19840890291339
l2 norm of gradients: 0.3574314535242041

```

12 norm of weights: 6.0788060910992385
-----
Iteration Number: 3343
Loss: 44.192040070296436
12 norm of gradients: 0.3573368693376236
12 norm of weights: 6.078698315645521
-----
Iteration Number: 3344
Loss: 44.18567456381373
12 norm of gradients: 0.3572423466899787
12 norm of weights: 6.07859055710999
-----
Iteration Number: 3345
Loss: 44.17931238047005
12 norm of gradients: 0.3571478855425901
12 norm of weights: 6.078482815482945
-----
Iteration Number: 3346
Loss: 44.17295351725165
12 norm of gradients: 0.35705348585677094
12 norm of weights: 6.078375090754694
-----
Iteration Number: 3347
Loss: 44.16659797113809
12 norm of gradients: 0.35695914759382713
12 norm of weights: 6.078267382915555
-----
Iteration Number: 3348
Loss: 44.16024573913764
12 norm of gradients: 0.3568648707150574
12 norm of weights: 6.078159691955855
-----
Iteration Number: 3349
Loss: 44.15389681824848
12 norm of gradients: 0.35677065518175316
12 norm of weights: 6.078052017865934
-----
Iteration Number: 3350
Loss: 44.1475512054623
12 norm of gradients: 0.3566765009551983
12 norm of weights: 6.077944360636139
-----
Iteration Number: 3351
Loss: 44.14120889780774
12 norm of gradients: 0.3565824079966699
12 norm of weights: 6.077836720256829
-----
Iteration Number: 3352

```

Loss: 44.134869892272896
l2 norm of gradients: 0.3564883762674375
l2 norm of weights: 6.077729096718369

Iteration Number: 3353
Loss: 44.128534185872574
l2 norm of gradients: 0.3563944057287637
l2 norm of weights: 6.077621490011141

Iteration Number: 3354
Loss: 44.12220177562282
l2 norm of gradients: 0.3563004963419037
l2 norm of weights: 6.077513900125529

Iteration Number: 3355
Loss: 44.11587265854996
l2 norm of gradients: 0.35620664806810576
l2 norm of weights: 6.077406327051931

Iteration Number: 3356
Loss: 44.10954683166519
l2 norm of gradients: 0.35611286086861105
l2 norm of weights: 6.077298770780756

Iteration Number: 3357
Loss: 44.1032242919984
l2 norm of gradients: 0.3560191347046537
l2 norm of weights: 6.077191231302419

Iteration Number: 3358
Loss: 44.09690503658568
l2 norm of gradients: 0.35592546953746057
l2 norm of weights: 6.077083708607348

Iteration Number: 3359
Loss: 44.09058906242258
l2 norm of gradients: 0.3558318653282519
l2 norm of weights: 6.076976202685978

Iteration Number: 3360
Loss: 44.08427636658159
l2 norm of gradients: 0.35573832203824063
l2 norm of weights: 6.076868713528759

Iteration Number: 3361
Loss: 44.07796694607287
l2 norm of gradients: 0.3556448396286328
l2 norm of weights: 6.076761241126144

```
-----
Iteration Number: 3362
Loss: 44.07166079794918
l2 norm of gradients: 0.35555141806062784
l2 norm of weights: 6.0766537854686
-----
Iteration Number: 3363
Loss: 44.06535791925084
l2 norm of gradients: 0.35545805729541774
l2 norm of weights: 6.0765463465466025
-----
Iteration Number: 3364
Loss: 44.059058306997194
l2 norm of gradients: 0.35536475729418826
l2 norm of weights: 6.076438924350638
-----
Iteration Number: 3365
Loss: 44.05276195828427
l2 norm of gradients: 0.35527151801811785
l2 norm of weights: 6.076331518871202
-----
Iteration Number: 3366
Loss: 44.04646887012743
l2 norm of gradients: 0.35517833942837845
l2 norm of weights: 6.076224130098799
-----
Iteration Number: 3367
Loss: 44.04017903958946
l2 norm of gradients: 0.35508522148613525
l2 norm of weights: 6.076116758023943
-----
Iteration Number: 3368
Loss: 44.03389246373182
l2 norm of gradients: 0.3549921641525465
l2 norm of weights: 6.07600940263716
-----
Iteration Number: 3369
Loss: 44.027609139622115
l2 norm of gradients: 0.3548991673887641
l2 norm of weights: 6.075902063928983
-----
Iteration Number: 3370
Loss: 44.0213290642845
l2 norm of gradients: 0.3548062311559328
l2 norm of weights: 6.075794741889958
-----
Iteration Number: 3371
Loss: 44.01505223484532
```

```

12 norm of gradients: 0.3547133554151915
12 norm of weights: 6.075687436510636
-----
Iteration Number: 3372
Loss: 44.008778648311264
12 norm of gradients: 0.35462054012767186
12 norm of weights: 6.075580147781583
-----
Iteration Number: 3373
Loss: 44.002508301788254
12 norm of gradients: 0.35452778525449924
12 norm of weights: 6.075472875693371
-----
Iteration Number: 3374
Loss: 43.99624119235433
12 norm of gradients: 0.3544350907567926
12 norm of weights: 6.0753656202365836
-----
Iteration Number: 3375
Loss: 43.98997731708305
12 norm of gradients: 0.35434245659566427
12 norm of weights: 6.075258381401811
-----
Iteration Number: 3376
Loss: 43.98371667304432
12 norm of gradients: 0.3542498827322203
12 norm of weights: 6.075151159179659
-----
Iteration Number: 3377
Loss: 43.97745925731607
12 norm of gradients: 0.3541573691275601
12 norm of weights: 6.075043953560737
-----
Iteration Number: 3378
Loss: 43.97120506701212
12 norm of gradients: 0.3540649157427773
12 norm of weights: 6.074936764535667
-----
Iteration Number: 3379
Loss: 43.96495409921576
12 norm of gradients: 0.3539725225389585
12 norm of weights: 6.074829592095079
-----
Iteration Number: 3380
Loss: 43.95870635099685
12 norm of gradients: 0.35388018947718475
12 norm of weights: 6.074722436229616
-----

```

Iteration Number: 3381
Loss: 43.952461819480654
l2 norm of gradients: 0.35378791651853014
l2 norm of weights: 6.0746152969299265

Iteration Number: 3382
Loss: 43.946220501752954
l2 norm of gradients: 0.35369570362406316
l2 norm of weights: 6.074508174186672

Iteration Number: 3383
Loss: 43.93998239488678
l2 norm of gradients: 0.35360355075484606
l2 norm of weights: 6.07440106799052

Iteration Number: 3384
Loss: 43.93374749604391
l2 norm of gradients: 0.35351145787193483
l2 norm of weights: 6.074293978332151

Iteration Number: 3385
Loss: 43.92751580227811
l2 norm of gradients: 0.35341942493637923
l2 norm of weights: 6.074186905202254

Iteration Number: 3386
Loss: 43.92128731074073
l2 norm of gradients: 0.35332745190922366
l2 norm of weights: 6.074079848591526

Iteration Number: 3387
Loss: 43.915062018518796
l2 norm of gradients: 0.353235538751506
l2 norm of weights: 6.0739728084906774

Iteration Number: 3388
Loss: 43.90883992272819
l2 norm of gradients: 0.3531436854242582
l2 norm of weights: 6.0738657848904225

Iteration Number: 3389
Loss: 43.90262102050152
l2 norm of gradients: 0.35305189188850666
l2 norm of weights: 6.073758777781491

Iteration Number: 3390
Loss: 43.89640530896313
l2 norm of gradients: 0.3529601581052717


```

12 norm of weights: 6.073651787154619
-----
Iteration Number: 3391
Loss: 43.89019278522418
12 norm of gradients: 0.3528684840355681
12 norm of weights: 6.07354481300055
-----
Iteration Number: 3392
Loss: 43.88398344643873
12 norm of gradients: 0.3527768696404045
12 norm of weights: 6.073437855310042
-----
Iteration Number: 3393
Loss: 43.87777728971212
12 norm of gradients: 0.3526853148807842
12 norm of weights: 6.07333091407386
-----
Iteration Number: 3394
Loss: 43.871574312171745
12 norm of gradients: 0.3525938197177046
12 norm of weights: 6.073223989282778
-----
Iteration Number: 3395
Loss: 43.865374510973254
12 norm of gradients: 0.3525023841121577
12 norm of weights: 6.073117080927581
-----
Iteration Number: 3396
Loss: 43.85917788324593
12 norm of gradients: 0.3524110080251299
12 norm of weights: 6.07301018899906
-----
Iteration Number: 3397
Loss: 43.8529844261506
12 norm of gradients: 0.3523196914176021
12 norm of weights: 6.0729033134880215
-----
Iteration Number: 3398
Loss: 43.84679413680924
12 norm of gradients: 0.3522284342505496
12 norm of weights: 6.072796454385277
-----
Iteration Number: 3399
Loss: 43.84060701238324
12 norm of gradients: 0.35213723648494244
12 norm of weights: 6.072689611681648
-----
Iteration Number: 3400

```

Loss: 43.834423050043675
l2 norm of gradients: 0.35204609808174514
l2 norm of weights: 6.072582785367965

Iteration Number: 3401
Loss: 43.82824224690293
l2 norm of gradients: 0.35195501900191717
l2 norm of weights: 6.072475975435071

Iteration Number: 3402
Loss: 43.822064600130545
l2 norm of gradients: 0.3518639992064124
l2 norm of weights: 6.072369181873817

Iteration Number: 3403
Loss: 43.81589010690692
l2 norm of gradients: 0.35177303865618
l2 norm of weights: 6.072262404675058

Iteration Number: 3404
Loss: 43.80971876439109
l2 norm of gradients: 0.35168213731216336
l2 norm of weights: 6.072155643829668

Iteration Number: 3405
Loss: 43.80355056974226
l2 norm of gradients: 0.35159129513530113
l2 norm of weights: 6.072048899328523

Iteration Number: 3406
Loss: 43.79738552011111
l2 norm of gradients: 0.35150051208652683
l2 norm of weights: 6.0719421711625134

Iteration Number: 3407
Loss: 43.79122361269447
l2 norm of gradients: 0.3514097881267691
l2 norm of weights: 6.071835459322535

Iteration Number: 3408
Loss: 43.78506484467019
l2 norm of gradients: 0.3513191232169513
l2 norm of weights: 6.071728763799494

Iteration Number: 3409
Loss: 43.7789092131896
l2 norm of gradients: 0.35122851731799226
l2 norm of weights: 6.071622084584307

```

-----
Iteration Number: 3410
Loss: 43.77275671546276
l2 norm of gradients: 0.3511379703908057
l2 norm of weights: 6.071515421667901
-----
Iteration Number: 3411
Loss: 43.76660734865925
l2 norm of gradients: 0.3510474823963005
l2 norm of weights: 6.071408775041209
-----
Iteration Number: 3412
Loss: 43.7604611099579
l2 norm of gradients: 0.35095705329538124
l2 norm of weights: 6.071302144695176
-----
Iteration Number: 3413
Loss: 43.75431799656335
l2 norm of gradients: 0.3508666830489474
l2 norm of weights: 6.071195530620756
-----
Iteration Number: 3414
Loss: 43.74817800565273
l2 norm of gradients: 0.3507763716178938
l2 norm of weights: 6.07108893280891
-----
Iteration Number: 3415
Loss: 43.742041134428504
l2 norm of gradients: 0.350686118963111
l2 norm of weights: 6.0709823512506125
-----
Iteration Number: 3416
Loss: 43.73590738010069
l2 norm of gradients: 0.35059592504548487
l2 norm of weights: 6.0708757859368445
-----
Iteration Number: 3417
Loss: 43.72977673983817
l2 norm of gradients: 0.3505057898258966
l2 norm of weights: 6.070769236858595
-----
Iteration Number: 3418
Loss: 43.723649210892226
l2 norm of gradients: 0.35041571326522336
l2 norm of weights: 6.070662704006867
-----
Iteration Number: 3419
Loss: 43.71752479041603

```

```

12 norm of gradients: 0.35032569532433766
12 norm of weights: 6.0705561873726674
-----
Iteration Number: 3420
Loss: 43.71140347565121
12 norm of gradients: 0.350235735964108
12 norm of weights: 6.070449686947017
-----
Iteration Number: 3421
Loss: 43.70528526381912
12 norm of gradients: 0.3501458351453982
12 norm of weights: 6.070343202720941
-----
Iteration Number: 3422
Loss: 43.69917015209942
12 norm of gradients: 0.35005599282906846
12 norm of weights: 6.070236734685478
-----
Iteration Number: 3423
Loss: 43.6930581377482
12 norm of gradients: 0.3499662089759743
12 norm of weights: 6.070130282831676
-----
Iteration Number: 3424
Loss: 43.68694921795714
12 norm of gradients: 0.34987648354696765
12 norm of weights: 6.070023847150589
-----
Iteration Number: 3425
Loss: 43.68084338996673
12 norm of gradients: 0.3497868165028961
12 norm of weights: 6.069917427633281
-----
Iteration Number: 3426
Loss: 43.67474065098674
12 norm of gradients: 0.34969720780460334
12 norm of weights: 6.069811024270828
-----
Iteration Number: 3427
Loss: 43.66864099827711
12 norm of gradients: 0.3496076574129294
12 norm of weights: 6.069704637054312
-----
Iteration Number: 3428
Loss: 43.66254442902856
12 norm of gradients: 0.34951816528871005
12 norm of weights: 6.069598265974827
-----

```

Iteration Number: 3429
Loss: 43.65645094051503
l2 norm of gradients: 0.34942873139277786
l2 norm of weights: 6.069491911023475

Iteration Number: 3430
Loss: 43.65036052995651
l2 norm of gradients: 0.34933935568596103
l2 norm of weights: 6.069385572191363

Iteration Number: 3431
Loss: 43.64427319460039
l2 norm of gradients: 0.3492500381290847
l2 norm of weights: 6.0692792494696155

Iteration Number: 3432
Loss: 43.638188931683636
l2 norm of gradients: 0.34916077868296996
l2 norm of weights: 6.0691729428493595

Iteration Number: 3433
Loss: 43.632107738444645
l2 norm of gradients: 0.3490715773084347
l2 norm of weights: 6.069066652321735

Iteration Number: 3434
Loss: 43.626029612156835
l2 norm of gradients: 0.3489824339662932
l2 norm of weights: 6.068960377877888

Iteration Number: 3435
Loss: 43.61995455004578
l2 norm of gradients: 0.3488933486173562
l2 norm of weights: 6.068854119508976

Iteration Number: 3436
Loss: 43.61388254938224
l2 norm of gradients: 0.3488043212224314
l2 norm of weights: 6.068747877206164

Iteration Number: 3437
Loss: 43.6078136074244
l2 norm of gradients: 0.3487153517423229
l2 norm of weights: 6.068641650960629

Iteration Number: 3438
Loss: 43.60174772144286
l2 norm of gradients: 0.3486264401378316

```

12 norm of weights: 6.068535440763553
-----
Iteration Number: 3439
Loss: 43.59568488868116
12 norm of gradients: 0.3485375863697557
12 norm of weights: 6.068429246606129
-----
Iteration Number: 3440
Loss: 43.58962510640627
12 norm of gradients: 0.34844879039888954
12 norm of weights: 6.068323068479562
-----
Iteration Number: 3441
Loss: 43.58356837190007
12 norm of gradients: 0.34836005218602506
12 norm of weights: 6.06821690637506
-----
Iteration Number: 3442
Loss: 43.57751468243002
12 norm of gradients: 0.3482713716919508
12 norm of weights: 6.0681107602838456
-----
Iteration Number: 3443
Loss: 43.571464035270864
12 norm of gradients: 0.3481827488774528
12 norm of weights: 6.068004630197148
-----
Iteration Number: 3444
Loss: 43.565416427698175
12 norm of gradients: 0.34809418370331396
12 norm of weights: 6.067898516106205
-----
Iteration Number: 3445
Loss: 43.55937185699365
12 norm of gradients: 0.34800567613031436
12 norm of weights: 6.0677924180022655
-----
Iteration Number: 3446
Loss: 43.55333032044826
12 norm of gradients: 0.34791722611923154
12 norm of weights: 6.0676863358765845
-----
Iteration Number: 3447
Loss: 43.54729181534199
12 norm of gradients: 0.3478288336308402
12 norm of weights: 6.067580269720429
-----
Iteration Number: 3448

```

```
Loss: 43.541256338951854
12 norm of gradients: 0.3477404986259128
12 norm of weights: 6.0674742195250735
-----
Iteration Number: 3449
Loss: 43.53522388858052
12 norm of gradients: 0.34765222106521887
12 norm of weights: 6.067368185281801
-----
Iteration Number: 3450
Loss: 43.52919446153154
12 norm of gradients: 0.3475640009095256
12 norm of weights: 6.067262166981904
-----
Iteration Number: 3451
Loss: 43.52316805508737
12 norm of gradients: 0.3474758381195979
12 norm of weights: 6.067156164616687
-----
Iteration Number: 3452
Loss: 43.51714466655317
12 norm of gradients: 0.3473877326561983
12 norm of weights: 6.067050178177457
-----
Iteration Number: 3453
Loss: 43.51112429324095
12 norm of gradients: 0.34729968448008697
12 norm of weights: 6.066944207655536
-----
Iteration Number: 3454
Loss: 43.50510693246223
12 norm of gradients: 0.347211693552022
12 norm of weights: 6.066838253042252
-----
Iteration Number: 3455
Loss: 43.49909258149349
12 norm of gradients: 0.34712375983275945
12 norm of weights: 6.066732314328942
-----
Iteration Number: 3456
Loss: 43.493081237666274
12 norm of gradients: 0.34703588328305285
12 norm of weights: 6.066626391506954
-----
Iteration Number: 3457
Loss: 43.48707289829517
12 norm of gradients: 0.3469480638636543
12 norm of weights: 6.066520484567643
```

```

-----
Iteration Number: 3458
Loss: 43.481067560689816
l2 norm of gradients: 0.3468603015353138
l2 norm of weights: 6.066414593502373
-----
Iteration Number: 3459
Loss: 43.475065222187055
l2 norm of gradients: 0.34677259625877915
l2 norm of weights: 6.066308718302517
-----
Iteration Number: 3460
Loss: 43.46906588010584
l2 norm of gradients: 0.3466849479947967
l2 norm of weights: 6.066202858959459
-----
Iteration Number: 3461
Loss: 43.4630695317484
l2 norm of gradients: 0.34659735670411124
l2 norm of weights: 6.066097015464588
-----
Iteration Number: 3462
Loss: 43.45707617446272
l2 norm of gradients: 0.34650982234746536
l2 norm of weights: 6.065991187809306
-----
Iteration Number: 3463
Loss: 43.45108580558561
l2 norm of gradients: 0.3464223448856005
l2 norm of weights: 6.0658853759850215
-----
Iteration Number: 3464
Loss: 43.44509842241947
l2 norm of gradients: 0.34633492427925633
l2 norm of weights: 6.065779579983151
-----
Iteration Number: 3465
Loss: 43.43911402233772
l2 norm of gradients: 0.3462475604891712
l2 norm of weights: 6.065673799795124
-----
Iteration Number: 3466
Loss: 43.433132602663555
l2 norm of gradients: 0.3461602534760821
l2 norm of weights: 6.065568035412373
-----
Iteration Number: 3467
Loss: 43.427154160728556

```



```

12 norm of gradients: 0.3460730032007246
12 norm of weights: 6.065462286826344
-----
Iteration Number: 3468
Loss: 43.42117869390424
12 norm of gradients: 0.3459858096238331
12 norm of weights: 6.065356554028491
-----
Iteration Number: 3469
Loss: 43.415206199503174
12 norm of gradients: 0.34589867270614066
12 norm of weights: 6.065250837010275
-----
Iteration Number: 3470
Loss: 43.409236674917466
12 norm of gradients: 0.34581159240837944
12 norm of weights: 6.065145135763167
-----
Iteration Number: 3471
Loss: 43.403270117466036
12 norm of gradients: 0.34572456869128054
12 norm of weights: 6.065039450278647
-----
Iteration Number: 3472
Loss: 43.3973065245177
12 norm of gradients: 0.3456376015155739
12 norm of weights: 6.064933780548205
-----
Iteration Number: 3473
Loss: 43.39134589341764
12 norm of gradients: 0.3455506908419887
12 norm of weights: 6.064828126563336
-----
Iteration Number: 3474
Loss: 43.38538822154394
12 norm of gradients: 0.34546383663125335
12 norm of weights: 6.064722488315547
-----
Iteration Number: 3475
Loss: 43.37943350626392
12 norm of gradients: 0.3453770388440953
12 norm of weights: 6.064616865796354
-----
Iteration Number: 3476
Loss: 43.373481744931325
12 norm of gradients: 0.3452902974412417
12 norm of weights: 6.064511258997278
-----

```

Iteration Number: 3477
Loss: 43.367532934919595
l2 norm of gradients: 0.3452036123834185
l2 norm of weights: 6.064405667909855

Iteration Number: 3478
Loss: 43.36158707361247
l2 norm of gradients: 0.3451169836313516
l2 norm of weights: 6.064300092525625

Iteration Number: 3479
Loss: 43.355644158372016
l2 norm of gradients: 0.3450304111457662
l2 norm of weights: 6.064194532836136

Iteration Number: 3480
Loss: 43.349704186565006
l2 norm of gradients: 0.3449438948873872
l2 norm of weights: 6.064088988832949

Iteration Number: 3481
Loss: 43.343767155597085
l2 norm of gradients: 0.344857434816939
l2 norm of weights: 6.063983460507631

Iteration Number: 3482
Loss: 43.33783306282973
l2 norm of gradients: 0.34477103089514605
l2 norm of weights: 6.063877947851758

Iteration Number: 3483
Loss: 43.33190190566119
l2 norm of gradients: 0.34468468308273204
l2 norm of weights: 6.063772450856915

Iteration Number: 3484
Loss: 43.32597368149246
l2 norm of gradients: 0.3445983913404212
l2 norm of weights: 6.063666969514695

Iteration Number: 3485
Loss: 43.320048387684324
l2 norm of gradients: 0.3445121556289373
l2 norm of weights: 6.0635615038167

Iteration Number: 3486
Loss: 43.314126021657415
l2 norm of gradients: 0.34442597590900437

```
l2 norm of weights: 6.0634560537545426
-----
Iteration Number: 3487
Loss: 43.3082065807876
l2 norm of gradients: 0.3443398521413462
l2 norm of weights: 6.0633506193198405
-----
Iteration Number: 3488
Loss: 43.302290062493775
l2 norm of gradients: 0.3442537842866872
l2 norm of weights: 6.063245200504223
-----
Iteration Number: 3489
Loss: 43.296376464172056
l2 norm of gradients: 0.3441677723057517
l2 norm of weights: 6.0631397972993275
-----
Iteration Number: 3490
Loss: 43.2904657832318
l2 norm of gradients: 0.3440818161592643
l2 norm of weights: 6.063034409696797
-----
Iteration Number: 3491
Loss: 43.28455801705608
l2 norm of gradients: 0.34399591580795036
l2 norm of weights: 6.06292903768829
-----
Iteration Number: 3492
Loss: 43.2786531630752
l2 norm of gradients: 0.3439100712125353
l2 norm of weights: 6.062823681265467
-----
Iteration Number: 3493
Loss: 43.2727512187184
l2 norm of gradients: 0.3438242823337453
l2 norm of weights: 6.062718340419998
-----
Iteration Number: 3494
Loss: 43.2668521813533
l2 norm of gradients: 0.34373854913230706
l2 norm of weights: 6.062613015143566
-----
Iteration Number: 3495
Loss: 43.26095604846126
l2 norm of gradients: 0.3436528715689479
l2 norm of weights: 6.062507705427857
-----
Iteration Number: 3496
```

```

Loss: 43.25506281741698
12 norm of gradients: 0.3435672496043961
12 norm of weights: 6.06240241126457
-----
Iteration Number: 3497
Loss: 43.249172485666186
12 norm of gradients: 0.3434816831993806
12 norm of weights: 6.06229713264541
-----
Iteration Number: 3498
Loss: 43.24328505061615
12 norm of gradients: 0.34339617231463127
12 norm of weights: 6.062191869562092
-----
Iteration Number: 3499
Loss: 43.23740050971611
12 norm of gradients: 0.3433107169108789
12 norm of weights: 6.062086622006339
-----
Iteration Number: 3500
Loss: 43.231518860399845
12 norm of gradients: 0.34322531694885544
12 norm of weights: 6.061981389969882
-----
Iteration Number: 3501
Loss: 43.225640100077456
12 norm of gradients: 0.3431399723892939
12 norm of weights: 6.061876173444461
-----
Iteration Number: 3502
Loss: 43.2197642262092
12 norm of gradients: 0.3430546831929285
12 norm of weights: 6.061770972421825
-----
Iteration Number: 3503
Loss: 43.21389123622244
12 norm of gradients: 0.3429694493204947
12 norm of weights: 6.061665786893731
-----
Iteration Number: 3504
Loss: 43.20802112757626
12 norm of gradients: 0.3428842707327293
12 norm of weights: 6.061560616851946
-----
Iteration Number: 3505
Loss: 43.202153897693364
12 norm of gradients: 0.3427991473903705
12 norm of weights: 6.061455462288241

```

```

-----
Iteration Number: 3506
Loss: 43.19628954404167
l2 norm of gradients: 0.34271407925415825
l2 norm of weights: 6.061350323194402
-----
Iteration Number: 3507
Loss: 43.1904280640478
l2 norm of gradients: 0.34262906628483353
l2 norm of weights: 6.061245199562217
-----
Iteration Number: 3508
Loss: 43.18456945519619
l2 norm of gradients: 0.3425441084431395
l2 norm of weights: 6.061140091383489
-----
Iteration Number: 3509
Loss: 43.17871371493075
l2 norm of gradients: 0.3424592056898207
l2 norm of weights: 6.061034998650025
-----
Iteration Number: 3510
Loss: 43.172860840702285
l2 norm of gradients: 0.3423743579856237
l2 norm of weights: 6.06092992135364
-----
Iteration Number: 3511
Loss: 43.167010829986324
l2 norm of gradients: 0.3422895652912966
l2 norm of weights: 6.060824859486161
-----
Iteration Number: 3512
Loss: 43.1611636802361
l2 norm of gradients: 0.3422048275675897
l2 norm of weights: 6.0607198130394195
-----
Iteration Number: 3513
Loss: 43.155319388912545
l2 norm of gradients: 0.34212014477525526
l2 norm of weights: 6.06061478200526
-----
Iteration Number: 3514
Loss: 43.14947795350188
l2 norm of gradients: 0.3420355168750477
l2 norm of weights: 6.060509766375531
-----
Iteration Number: 3515
Loss: 43.143639371474826

```

```

12 norm of gradients: 0.34195094382772345
12 norm of weights: 6.060404766142091
-----
Iteration Number: 3516
Loss: 43.13780364028849
12 norm of gradients: 0.3418664255940413
12 norm of weights: 6.060299781296808
-----
Iteration Number: 3517
Loss: 43.13197075743168
12 norm of gradients: 0.3417819621347621
12 norm of weights: 6.0601948118315585
-----
Iteration Number: 3518
Loss: 43.126140720407335
12 norm of gradients: 0.3416975534106496
12 norm of weights: 6.060089857738225
-----
Iteration Number: 3519
Loss: 43.12031352664807
12 norm of gradients: 0.34161319938246953
12 norm of weights: 6.0599849190087
-----
Iteration Number: 3520
Loss: 43.11448917368375
12 norm of gradients: 0.34152890001099034
12 norm of weights: 6.059879995634884
-----
Iteration Number: 3521
Loss: 43.10866765897722
12 norm of gradients: 0.3414446552569831
12 norm of weights: 6.059775087608688
-----
Iteration Number: 3522
Loss: 43.10284898003827
12 norm of gradients: 0.34136046508122153
12 norm of weights: 6.059670194922028
-----
Iteration Number: 3523
Loss: 43.09703313433891
12 norm of gradients: 0.34127632944448216
12 norm of weights: 6.059565317566832
-----
Iteration Number: 3524
Loss: 43.091220119391316
12 norm of gradients: 0.3411922483075444
12 norm of weights: 6.059460455535029
-----

```

Iteration Number: 3525
Loss: 43.08540993267834
l2 norm of gradients: 0.34110822163119037
l2 norm of weights: 6.059355608818566

Iteration Number: 3526
Loss: 43.07960257173541
l2 norm of gradients: 0.3410242493762055
l2 norm of weights: 6.059250777409392

Iteration Number: 3527
Loss: 43.073798034023504
l2 norm of gradients: 0.340940331503378
l2 norm of weights: 6.059145961299468

Iteration Number: 3528
Loss: 43.06799631707282
l2 norm of gradients: 0.34085646797349944
l2 norm of weights: 6.05904116048076

Iteration Number: 3529
Loss: 43.06219741839951
l2 norm of gradients: 0.34077265874736434
l2 norm of weights: 6.058936374945242

Iteration Number: 3530
Loss: 43.05640133549083
l2 norm of gradients: 0.34068890378577077
l2 norm of weights: 6.058831604684901

Iteration Number: 3531
Loss: 43.050608065886045
l2 norm of gradients: 0.3406052030495202
l2 norm of weights: 6.058726849691727

Iteration Number: 3532
Loss: 43.044817607085804
l2 norm of gradients: 0.3405215564994172
l2 norm of weights: 6.058622109957723

Iteration Number: 3533
Loss: 43.03902995662648
l2 norm of gradients: 0.3404379640962702
l2 norm of weights: 6.058517385474896

Iteration Number: 3534
Loss: 43.03324511199571
l2 norm of gradients: 0.3403544258008911

```

12 norm of weights: 6.058412676235263
-----
Iteration Number: 3535
Loss: 43.027463070760675
12 norm of gradients: 0.34027094157409543
12 norm of weights: 6.05830798223085
-----
Iteration Number: 3536
Loss: 43.021683830427406
12 norm of gradients: 0.3401875113767026
12 norm of weights: 6.058203303453691
-----
Iteration Number: 3537
Loss: 43.015907388526095
12 norm of gradients: 0.34010413516953564
12 norm of weights: 6.058098639895825
-----
Iteration Number: 3538
Loss: 43.01013374258335
12 norm of gradients: 0.3400208129134219
12 norm of weights: 6.057993991549306
-----
Iteration Number: 3539
Loss: 43.004362890161254
12 norm of gradients: 0.339937544569192
12 norm of weights: 6.0578893584061895
-----
Iteration Number: 3540
Loss: 42.99859482876466
12 norm of gradients: 0.3398543300976813
12 norm of weights: 6.057784740458543
-----
Iteration Number: 3541
Loss: 42.99282955596153
12 norm of gradients: 0.33977116945972896
12 norm of weights: 6.0576801376984415
-----
Iteration Number: 3542
Loss: 42.987067069270466
12 norm of gradients: 0.3396880626161783
12 norm of weights: 6.057575550117965
-----
Iteration Number: 3543
Loss: 42.981307366256104
12 norm of gradients: 0.33960500952787714
12 norm of weights: 6.057470977709207
-----
Iteration Number: 3544

```


Loss: 42.975550444461504
l2 norm of gradients: 0.3395220101556774
l2 norm of weights: 6.057366420464266

Iteration Number: 3545
Loss: 42.969796301446316
l2 norm of gradients: 0.33943906446043576
l2 norm of weights: 6.057261878375249

Iteration Number: 3546
Loss: 42.96404493474564
l2 norm of gradients: 0.33935617240301286
l2 norm of weights: 6.057157351434272

Iteration Number: 3547
Loss: 42.95829634191303
l2 norm of gradients: 0.33927333394427467
l2 norm of weights: 6.057052839633457

Iteration Number: 3548
Loss: 42.952550520530224
l2 norm of gradients: 0.33919054904509116
l2 norm of weights: 6.056948342964938

Iteration Number: 3549
Loss: 42.94680746814568
l2 norm of gradients: 0.3391078176663373
l2 norm of weights: 6.056843861420852

Iteration Number: 3550
Loss: 42.9410671823175
l2 norm of gradients: 0.339025139768893
l2 norm of weights: 6.056739394993348

Iteration Number: 3551
Loss: 42.935329660615
l2 norm of gradients: 0.3389425153136428
l2 norm of weights: 6.056634943674582

Iteration Number: 3552
Loss: 42.929594900632026
l2 norm of gradients: 0.3388599442614763
l2 norm of weights: 6.056530507456718

Iteration Number: 3553
Loss: 42.92386289990176
l2 norm of gradients: 0.33877742657328824
l2 norm of weights: 6.056426086331928

```

-----
Iteration Number: 3554
Loss: 42.91813365603422
l2 norm of gradients: 0.33869496220997836
l2 norm of weights: 6.056321680292394
-----
Iteration Number: 3555
Loss: 42.912407166567064
l2 norm of gradients: 0.3386125511324516
l2 norm of weights: 6.0562172893303
-----
Iteration Number: 3556
Loss: 42.90668342911009
l2 norm of gradients: 0.3385301933016182
l2 norm of weights: 6.056112913437846
-----
Iteration Number: 3557
Loss: 42.900962441231535
l2 norm of gradients: 0.33844788867839376
l2 norm of weights: 6.056008552607236
-----
Iteration Number: 3558
Loss: 42.895244200525745
l2 norm of gradients: 0.3383656372236993
l2 norm of weights: 6.0559042068306805
-----
Iteration Number: 3559
Loss: 42.88952870455459
l2 norm of gradients: 0.3382834388984611
l2 norm of weights: 6.0557998761004015
-----
Iteration Number: 3560
Loss: 42.88381595094738
l2 norm of gradients: 0.33820129366361135
l2 norm of weights: 6.055695560408627
-----
Iteration Number: 3561
Loss: 42.87810593727922
l2 norm of gradients: 0.3381192014800877
l2 norm of weights: 6.055591259747593
-----
Iteration Number: 3562
Loss: 42.87239866113528
l2 norm of gradients: 0.3380371623088336
l2 norm of weights: 6.055486974109544
-----
Iteration Number: 3563
Loss: 42.86669412010777

```

```

12 norm of gradients: 0.3379551761107983
12 norm of weights: 6.055382703486734
-----
Iteration Number: 3564
Loss: 42.86099231181089
12 norm of gradients: 0.3378732428469367
12 norm of weights: 6.05527844787142
-----
Iteration Number: 3565
Loss: 42.85529323385006
12 norm of gradients: 0.3377913624782099
12 norm of weights: 6.055174207255874
-----
Iteration Number: 3566
Loss: 42.849596883826095
12 norm of gradients: 0.337709534965585
12 norm of weights: 6.05506998163237
-----
Iteration Number: 3567
Loss: 42.84390325932887
12 norm of gradients: 0.337627760270035
12 norm of weights: 6.054965770993193
-----
Iteration Number: 3568
Loss: 42.83821235798416
12 norm of gradients: 0.33754603835253927
12 norm of weights: 6.054861575330635
-----
Iteration Number: 3569
Loss: 42.83252417741642
12 norm of gradients: 0.33746436917408346
12 norm of weights: 6.054757394636997
-----
Iteration Number: 3570
Loss: 42.826838715228234
12 norm of gradients: 0.33738275269565926
12 norm of weights: 6.054653228904586
-----
Iteration Number: 3571
Loss: 42.82115596903514
12 norm of gradients: 0.33730118887826505
12 norm of weights: 6.054549078125719
-----
Iteration Number: 3572
Loss: 42.81547593644874
12 norm of gradients: 0.33721967768290545
12 norm of weights: 6.05444494229272
-----

```

Iteration Number: 3573
Loss: 42.80979861509839
l2 norm of gradients: 0.337138219070592
l2 norm of weights: 6.054340821397919

Iteration Number: 3574
Loss: 42.80412400262032
l2 norm of gradients: 0.33705681300234236
l2 norm of weights: 6.0542367154336585

Iteration Number: 3575
Loss: 42.7984520966389
l2 norm of gradients: 0.3369754594391813
l2 norm of weights: 6.054132624392284

Iteration Number: 3576
Loss: 42.79278289476808
l2 norm of gradients: 0.33689415834214015
l2 norm of weights: 6.0540285482661504

Iteration Number: 3577
Loss: 42.7871163946528
l2 norm of gradients: 0.33681290967225697
l2 norm of weights: 6.053924487047624

Iteration Number: 3578
Loss: 42.78145259393609
l2 norm of gradients: 0.33673171339057717
l2 norm of weights: 6.053820440729074

Iteration Number: 3579
Loss: 42.77579149023979
l2 norm of gradients: 0.33665056945815275
l2 norm of weights: 6.053716409302879

Iteration Number: 3580
Loss: 42.7701330812171
l2 norm of gradients: 0.336569477836043
l2 norm of weights: 6.053612392761428

Iteration Number: 3581
Loss: 42.764477364513844
l2 norm of gradients: 0.3364884384853142
l2 norm of weights: 6.053508391097113

Iteration Number: 3582
Loss: 42.758824337739306
l2 norm of gradients: 0.3364074513670401

```

12 norm of weights: 6.053404404302339
-----
Iteration Number: 3583
Loss: 42.753173998591606
12 norm of gradients: 0.3363265164423013
12 norm of weights: 6.0533004323695145
-----
Iteration Number: 3584
Loss: 42.74752634468265
12 norm of gradients: 0.33624563367218635
12 norm of weights: 6.05319647529106
-----
Iteration Number: 3585
Loss: 42.741881373685075
12 norm of gradients: 0.3361648030177906
12 norm of weights: 6.053092533059398
-----
Iteration Number: 3586
Loss: 42.736239083236654
12 norm of gradients: 0.3360840244402177
12 norm of weights: 6.0529886056669655
-----
Iteration Number: 3587
Loss: 42.730599471023375
12 norm of gradients: 0.3360032979005781
12 norm of weights: 6.052884693106202
-----
Iteration Number: 3588
Loss: 42.72496253468571
12 norm of gradients: 0.3359226233599903
12 norm of weights: 6.05278079536956
-----
Iteration Number: 3589
Loss: 42.71932827189028
12 norm of gradients: 0.3358420007795806
12 norm of weights: 6.052676912449492
-----
Iteration Number: 3590
Loss: 42.71369668030462
12 norm of gradients: 0.33576143012048304
12 norm of weights: 6.052573044338465
-----
Iteration Number: 3591
Loss: 42.70806775759163
12 norm of gradients: 0.3356809113438395
12 norm of weights: 6.052469191028952
-----
Iteration Number: 3592

```

```

Loss: 42.702441501425405
12 norm of gradients: 0.3356004444107997
12 norm of weights: 6.0523653525134335
-----
Iteration Number: 3593
Loss: 42.69681790947285
12 norm of gradients: 0.3355200292825218
12 norm of weights: 6.052261528784396
-----
Iteration Number: 3594
Loss: 42.69119697942268
12 norm of gradients: 0.33543966592017155
12 norm of weights: 6.052157719834336
-----
Iteration Number: 3595
Loss: 42.68557870892993
12 norm of gradients: 0.33535935428492336
12 norm of weights: 6.052053925655758
-----
Iteration Number: 3596
Loss: 42.679963095686
12 norm of gradients: 0.33527909433795944
12 norm of weights: 6.051950146241173
-----
Iteration Number: 3597
Loss: 42.6743501373839
12 norm of gradients: 0.3351988860404709
12 norm of weights: 6.051846381583097
-----
Iteration Number: 3598
Loss: 42.66873983169587
12 norm of gradients: 0.3351187293536566
12 norm of weights: 6.051742631674061
-----
Iteration Number: 3599
Loss: 42.663132176325135
12 norm of gradients: 0.3350386242387243
12 norm of weights: 6.051638896506595
-----
Iteration Number: 3600
Loss: 42.657527168945734
12 norm of gradients: 0.33495857065689016
12 norm of weights: 6.051535176073243
-----
Iteration Number: 3601
Loss: 42.6519248072344
12 norm of gradients: 0.3348785685693792
12 norm of weights: 6.051431470366555

```

```

-----
Iteration Number: 3602
Loss: 42.64632508891831
l2 norm of gradients: 0.33479861793742477
l2 norm of weights: 6.051327779379087
-----
Iteration Number: 3603
Loss: 42.64072801168583
l2 norm of gradients: 0.334718718722269
l2 norm of weights: 6.051224103103404
-----
Iteration Number: 3604
Loss: 42.635133573244715
l2 norm of gradients: 0.33463887088516325
l2 norm of weights: 6.051120441532079
-----
Iteration Number: 3605
Loss: 42.6295417712645
l2 norm of gradients: 0.3345590743873675
l2 norm of weights: 6.051016794657692
-----
Iteration Number: 3606
Loss: 42.62395260348104
l2 norm of gradients: 0.33447932919015066
l2 norm of weights: 6.050913162472829
-----
Iteration Number: 3607
Loss: 42.618366067584454
l2 norm of gradients: 0.33439963525479094
l2 norm of weights: 6.050809544970089
-----
Iteration Number: 3608
Loss: 42.61278216130611
l2 norm of gradients: 0.3343199925425754
l2 norm of weights: 6.05070594214207
-----
Iteration Number: 3609
Loss: 42.60720088232815
l2 norm of gradients: 0.3342404010148005
l2 norm of weights: 6.050602353981387
-----
Iteration Number: 3610
Loss: 42.601622228383846
l2 norm of gradients: 0.33416086063277184
l2 norm of weights: 6.050498780480655
-----
Iteration Number: 3611
Loss: 42.596046197202156

```

```

12 norm of gradients: 0.33408137135780447
12 norm of weights: 6.050395221632501
-----
Iteration Number: 3612
Loss: 42.59047278646038
12 norm of gradients: 0.3340019331512229
12 norm of weights: 6.050291677429557
-----
Iteration Number: 3613
Loss: 42.58490199391836
12 norm of gradients: 0.333922545974361
12 norm of weights: 6.050188147864466
-----
Iteration Number: 3614
Loss: 42.579333817296714
12 norm of gradients: 0.3338432097885621
12 norm of weights: 6.050084632929874
-----
Iteration Number: 3615
Loss: 42.5737682543162
12 norm of gradients: 0.33376392455517945
12 norm of weights: 6.049981132618436
-----
Iteration Number: 3616
Loss: 42.5682053026856
12 norm of gradients: 0.3336846902355759
12 norm of weights: 6.049877646922819
-----
Iteration Number: 3617
Loss: 42.562644960161954
12 norm of gradients: 0.3336055067911238
12 norm of weights: 6.04977417583569
-----
Iteration Number: 3618
Loss: 42.55708722447763
12 norm of gradients: 0.33352637418320574
12 norm of weights: 6.049670719349731
-----
Iteration Number: 3619
Loss: 42.55153209336813
12 norm of gradients: 0.3334472923732139
12 norm of weights: 6.0495672774576255
-----
Iteration Number: 3620
Loss: 42.545979564557236
12 norm of gradients: 0.33336826132255076
12 norm of weights: 6.0494638501520654
-----

```


Iteration Number: 3621
Loss: 42.54042963580476
l2 norm of gradients: 0.3332892809926287
l2 norm of weights: 6.049360437425755

Iteration Number: 3622
Loss: 42.53488230483887
l2 norm of gradients: 0.33321035134487015
l2 norm of weights: 6.0492570392714

Iteration Number: 3623
Loss: 42.529337569405314
l2 norm of gradients: 0.3331314723407078
l2 norm of weights: 6.049153655681718

Iteration Number: 3624
Loss: 42.52379542727073
l2 norm of gradients: 0.3330526439415845
l2 norm of weights: 6.049050286649432

Iteration Number: 3625
Loss: 42.51825587617724
l2 norm of gradients: 0.3329738661089537
l2 norm of weights: 6.0489469321672695

Iteration Number: 3626
Loss: 42.51271891387823
l2 norm of gradients: 0.332895138804279
l2 norm of weights: 6.048843592227972

Iteration Number: 3627
Loss: 42.507184538108554
l2 norm of gradients: 0.33281646198903475
l2 norm of weights: 6.048740266824285

Iteration Number: 3628
Loss: 42.50165274666493
l2 norm of gradients: 0.3327378356247055
l2 norm of weights: 6.048636955948959

Iteration Number: 3629
Loss: 42.49612353728147
l2 norm of gradients: 0.33265925967278664
l2 norm of weights: 6.048533659594757

Iteration Number: 3630
Loss: 42.4905969077263
l2 norm of gradients: 0.33258073409478417

```

12 norm of weights: 6.048430377754445
-----
Iteration Number: 3631
Loss: 42.48507285576466
12 norm of gradients: 0.33250225885221496
12 norm of weights: 6.048327110420799
-----
Iteration Number: 3632
Loss: 42.47955137916378
12 norm of gradients: 0.33242383390660646
12 norm of weights: 6.048223857586602
-----
Iteration Number: 3633
Loss: 42.47403247570268
12 norm of gradients: 0.3323454592194973
12 norm of weights: 6.048120619244643
-----
Iteration Number: 3634
Loss: 42.468516143135304
12 norm of gradients: 0.3322671347524369
12 norm of weights: 6.048017395387719
-----
Iteration Number: 3635
Loss: 42.463002379258135
12 norm of gradients: 0.3321888604669857
12 norm of weights: 6.047914186008636
-----
Iteration Number: 3636
Loss: 42.45749118181963
12 norm of gradients: 0.3321106363247154
12 norm of weights: 6.047810991100204
-----
Iteration Number: 3637
Loss: 42.45198254863759
12 norm of gradients: 0.3320324622872085
12 norm of weights: 6.047707810655245
-----
Iteration Number: 3638
Loss: 42.446476477467854
12 norm of gradients: 0.33195433831605936
12 norm of weights: 6.047604644666584
-----
Iteration Number: 3639
Loss: 42.44097296609847
12 norm of gradients: 0.33187626437287293
12 norm of weights: 6.047501493127054
-----
Iteration Number: 3640

```

Loss: 42.43547201230685
l2 norm of gradients: 0.33179824041926614
l2 norm of weights: 6.047398356029499

Iteration Number: 3641
Loss: 42.429973613915514
l2 norm of gradients: 0.33172026641686697
l2 norm of weights: 6.047295233366765

Iteration Number: 3642
Loss: 42.42447776868291
l2 norm of gradients: 0.33164234232731504
l2 norm of weights: 6.047192125131711

Iteration Number: 3643
Loss: 42.41898447442046
l2 norm of gradients: 0.3315644681122615
l2 norm of weights: 6.047089031317197

Iteration Number: 3644
Loss: 42.41349372890855
l2 norm of gradients: 0.3314866437333693
l2 norm of weights: 6.0469859519160964

Iteration Number: 3645
Loss: 42.40800552995076
l2 norm of gradients: 0.3314088691523129
l2 norm of weights: 6.046882886921285

Iteration Number: 3646
Loss: 42.40251987536763
l2 norm of gradients: 0.3313311443307785
l2 norm of weights: 6.046779836325648

Iteration Number: 3647
Loss: 42.39703676292688
l2 norm of gradients: 0.33125346923046445
l2 norm of weights: 6.046676800122079

Iteration Number: 3648
Loss: 42.39155619047967
l2 norm of gradients: 0.33117584381308074
l2 norm of weights: 6.046573778303476

Iteration Number: 3649
Loss: 42.3860781557851
l2 norm of gradients: 0.3310982680403495
l2 norm of weights: 6.046470770862747

```

-----
Iteration Number: 3650
Loss: 42.38060265669858
l2 norm of gradients: 0.33102074187400465
l2 norm of weights: 6.046367777792806
-----
Iteration Number: 3651
Loss: 42.37512969099855
l2 norm of gradients: 0.3309432652757925
l2 norm of weights: 6.046264799086574
-----
Iteration Number: 3652
Loss: 42.36965925650994
l2 norm of gradients: 0.33086583820747145
l2 norm of weights: 6.046161834736979
-----
Iteration Number: 3653
Loss: 42.36419135105899
l2 norm of gradients: 0.3307884606308122
l2 norm of weights: 6.046058884736957
-----
Iteration Number: 3654
Loss: 42.35872597245989
l2 norm of gradients: 0.3307111325075975
l2 norm of weights: 6.045955949079452
-----
Iteration Number: 3655
Loss: 42.35326311852455
l2 norm of gradients: 0.3306338537996228
l2 norm of weights: 6.0458530277574125
-----
Iteration Number: 3656
Loss: 42.3478027870984
l2 norm of gradients: 0.33055662446869566
l2 norm of weights: 6.045750120763796
-----
Iteration Number: 3657
Loss: 42.34234497599449
l2 norm of gradients: 0.33047944447663646
l2 norm of weights: 6.0456472280915685
-----
Iteration Number: 3658
Loss: 42.33688968304304
l2 norm of gradients: 0.3304023137852781
l2 norm of weights: 6.0455443497337
-----
Iteration Number: 3659
Loss: 42.33143690608177

```

```

12 norm of gradients: 0.33032523235646594
12 norm of weights: 6.04544148568317
-----
Iteration Number: 3660
Loss: 42.32598664291841
12 norm of gradients: 0.33024820015205814
12 norm of weights: 6.045338635932964
-----
Iteration Number: 3661
Loss: 42.320538891428455
12 norm of gradients: 0.3301712171339256
12 norm of weights: 6.045235800476076
-----
Iteration Number: 3662
Loss: 42.31509364942652
12 norm of gradients: 0.33009428326395224
12 norm of weights: 6.045132979305505
-----
Iteration Number: 3663
Loss: 42.30965091475579
12 norm of gradients: 0.3300173985040344
12 norm of weights: 6.045030172414258
-----
Iteration Number: 3664
Loss: 42.30421068526169
12 norm of gradients: 0.32994056281608186
12 norm of weights: 6.044927379795352
-----
Iteration Number: 3665
Loss: 42.29877295879789
12 norm of gradients: 0.32986377616201723
12 norm of weights: 6.044824601441806
-----
Iteration Number: 3666
Loss: 42.29333773318988
12 norm of gradients: 0.32978703850377633
12 norm of weights: 6.044721837346649
-----
Iteration Number: 3667
Loss: 42.28790500631531
12 norm of gradients: 0.3297103498033079
12 norm of weights: 6.044619087502918
-----
Iteration Number: 3668
Loss: 42.2824747760011
12 norm of gradients: 0.3296337100225739
12 norm of weights: 6.044516351903654
-----

```

Iteration Number: 3669
Loss: 42.27704704011746
l2 norm of gradients: 0.3295571191235499
l2 norm of weights: 6.044413630541909

Iteration Number: 3670
Loss: 42.27162179651955
l2 norm of gradients: 0.32948057706822453
l2 norm of weights: 6.044310923410738

Iteration Number: 3671
Loss: 42.26619904305675
l2 norm of gradients: 0.3294040838185999
l2 norm of weights: 6.044208230503207

Iteration Number: 3672
Loss: 42.26077877760444
l2 norm of gradients: 0.32932763933669146
l2 norm of weights: 6.044105551812385

Iteration Number: 3673
Loss: 42.25536099802946
l2 norm of gradients: 0.3292512435845284
l2 norm of weights: 6.044002887331351

Iteration Number: 3674
Loss: 42.249945702172056
l2 norm of gradients: 0.32917489652415344
l2 norm of weights: 6.043900237053191

Iteration Number: 3675
Loss: 42.24453288792715
l2 norm of gradients: 0.3290985981176227
l2 norm of weights: 6.043797600970996

Iteration Number: 3676
Loss: 42.23912255314812
l2 norm of gradients: 0.32902234832700644
l2 norm of weights: 6.043694979077865

Iteration Number: 3677
Loss: 42.23371469572132
l2 norm of gradients: 0.32894614711438847
l2 norm of weights: 6.0435923713669055

Iteration Number: 3678
Loss: 42.2283093135059
l2 norm of gradients: 0.32886999444186643

```

12 norm of weights: 6.043489777831229
-----
Iteration Number: 3679
Loss: 42.22290640440455
12 norm of gradients: 0.32879389027155187
12 norm of weights: 6.043387198463957
-----
Iteration Number: 3680
Loss: 42.217505966262685
12 norm of gradients: 0.3287178345655703
12 norm of weights: 6.043284633258216
-----
Iteration Number: 3681
Loss: 42.21210799700045
12 norm of gradients: 0.32864182728606134
12 norm of weights: 6.0431820822071405
-----
Iteration Number: 3682
Loss: 42.20671249447208
12 norm of gradients: 0.32856586839517854
12 norm of weights: 6.043079545303872
-----
Iteration Number: 3683
Loss: 42.20131945658031
12 norm of gradients: 0.3284899578550898
12 norm of weights: 6.042977022541557
-----
Iteration Number: 3684
Loss: 42.19592888119779
12 norm of gradients: 0.328414095627977
12 norm of weights: 6.042874513913351
-----
Iteration Number: 3685
Loss: 42.19054076623136
12 norm of gradients: 0.32833828167603646
12 norm of weights: 6.042772019412417
-----
Iteration Number: 3686
Loss: 42.18515510958533
12 norm of gradients: 0.3282625159614788
12 norm of weights: 6.042669539031923
-----
Iteration Number: 3687
Loss: 42.17977190912865
12 norm of gradients: 0.3281867984465289
12 norm of weights: 6.042567072765045
-----
Iteration Number: 3688

```

```

Loss: 42.17439116277349
l2 norm of gradients: 0.3281111290934263
l2 norm of weights: 6.0424646206049655
-----
Iteration Number: 3689
Loss: 42.16901286841094
l2 norm of gradients: 0.3280355078644248
l2 norm of weights: 6.042362182544874
-----
Iteration Number: 3690
Loss: 42.163637023944936
l2 norm of gradients: 0.32795993472179297
l2 norm of weights: 6.042259758577967
-----
Iteration Number: 3691
Loss: 42.158263627304066
l2 norm of gradients: 0.327884409627814
l2 norm of weights: 6.042157348697449
-----
Iteration Number: 3692
Loss: 42.15289267636452
l2 norm of gradients: 0.3278089325447856
l2 norm of weights: 6.04205495289653
-----
Iteration Number: 3693
Loss: 42.14752416904624
l2 norm of gradients: 0.32773350343502033
l2 norm of weights: 6.041952571168426
-----
Iteration Number: 3694
Loss: 42.14215810326893
l2 norm of gradients: 0.3276581222608455
l2 norm of weights: 6.041850203506361
-----
Iteration Number: 3695
Loss: 42.13679447693522
l2 norm of gradients: 0.32758278898460347
l2 norm of weights: 6.041747849903567
-----
Iteration Number: 3696
Loss: 42.13143328797385
l2 norm of gradients: 0.32750750356865127
l2 norm of weights: 6.041645510353281
-----
Iteration Number: 3697
Loss: 42.126074534297075
l2 norm of gradients: 0.32743226597536107
l2 norm of weights: 6.041543184848748

```



```

-----
Iteration Number: 3698
Loss: 42.120718213809496
l2 norm of gradients: 0.32735707616711995
l2 norm of weights: 6.041440873383219
-----
Iteration Number: 3699
Loss: 42.115364324455214
l2 norm of gradients: 0.3272819341063304
l2 norm of weights: 6.041338575949952
-----
Iteration Number: 3700
Loss: 42.11001286415501
l2 norm of gradients: 0.3272068397554094
l2 norm of weights: 6.041236292542213
-----
Iteration Number: 3701
Loss: 42.10466383082938
l2 norm of gradients: 0.32713179307678997
l2 norm of weights: 6.0411340231532735
-----
Iteration Number: 3702
Loss: 42.09931722240485
l2 norm of gradients: 0.3270567940329198
l2 norm of weights: 6.041031767776412
-----
Iteration Number: 3703
Loss: 42.093973036827975
l2 norm of gradients: 0.32698184258626206
l2 norm of weights: 6.040929526404912
-----
Iteration Number: 3704
Loss: 42.08863127202484
l2 norm of gradients: 0.3269069386992955
l2 norm of weights: 6.0408272990320695
-----
Iteration Number: 3705
Loss: 42.08329192593874
l2 norm of gradients: 0.32683208233451405
l2 norm of weights: 6.040725085651181
-----
Iteration Number: 3706
Loss: 42.077954996477835
l2 norm of gradients: 0.3267572734544272
l2 norm of weights: 6.040622886255552
-----
Iteration Number: 3707
Loss: 42.07262048163239

```

```

12 norm of gradients: 0.32668251202156
12 norm of weights: 6.040520700838498
-----
Iteration Number: 3708
Loss: 42.067288379302035
12 norm of gradients: 0.3266077979984532
12 norm of weights: 6.040418529393334
-----
Iteration Number: 3709
Loss: 42.06195868744958
12 norm of gradients: 0.32653313134766326
12 norm of weights: 6.0403163719133905
-----
Iteration Number: 3710
Loss: 42.056631404027016
12 norm of gradients: 0.32645851203176196
12 norm of weights: 6.040214228391997
-----
Iteration Number: 3711
Loss: 42.051306526970656
12 norm of gradients: 0.32638394001333737
12 norm of weights: 6.040112098822492
-----
Iteration Number: 3712
Loss: 42.04598405424494
12 norm of gradients: 0.3263094152549932
12 norm of weights: 6.0400099831982255
-----
Iteration Number: 3713
Loss: 42.040663983789806
12 norm of gradients: 0.32623493771934875
12 norm of weights: 6.039907881512549
-----
Iteration Number: 3714
Loss: 42.03534631356415
12 norm of gradients: 0.32616050736903973
12 norm of weights: 6.03980579375882
-----
Iteration Number: 3715
Loss: 42.030031041546984
12 norm of gradients: 0.3260861241667175
12 norm of weights: 6.039703719930407
-----
Iteration Number: 3716
Loss: 42.02471816567613
12 norm of gradients: 0.32601178807504977
12 norm of weights: 6.039601660020683
-----

```

Iteration Number: 3717
Loss: 42.01940768393211
l2 norm of gradients: 0.32593749905672004
l2 norm of weights: 6.039499614023026

Iteration Number: 3718
Loss: 42.014099594263435
l2 norm of gradients: 0.32586325707442815
l2 norm of weights: 6.039397581930824

Iteration Number: 3719
Loss: 42.00879389464167
l2 norm of gradients: 0.32578906209089015
l2 norm of weights: 6.039295563737469

Iteration Number: 3720
Loss: 42.00349058303951
l2 norm of gradients: 0.32571491406883835
l2 norm of weights: 6.039193559436363

Iteration Number: 3721
Loss: 41.99818965741629
l2 norm of gradients: 0.3256408129710212
l2 norm of weights: 6.039091569020909

Iteration Number: 3722
Loss: 41.992891115774285
l2 norm of gradients: 0.32556675876020397
l2 norm of weights: 6.038989592484522

Iteration Number: 3723
Loss: 41.98759495605461
l2 norm of gradients: 0.3254927513991677
l2 norm of weights: 6.03888762982062

Iteration Number: 3724
Loss: 41.982301176265686
l2 norm of gradients: 0.32541879085071057
l2 norm of weights: 6.038785681022631

Iteration Number: 3725
Loss: 41.97700977436333
l2 norm of gradients: 0.32534487707764703
l2 norm of weights: 6.038683746083986

Iteration Number: 3726
Loss: 41.97172074833837
l2 norm of gradients: 0.32527101004280773

```

12 norm of weights: 6.0385818249981265
-----
Iteration Number: 3727
Loss: 41.96643409618176
12 norm of gradients: 0.32519718970904066
12 norm of weights: 6.0384799177584965
-----
Iteration Number: 3728
Loss: 41.96114981588423
12 norm of gradients: 0.32512341603921
12 norm of weights: 6.0383780243585505
-----
Iteration Number: 3729
Loss: 41.9558679054258
12 norm of gradients: 0.3250496889961969
12 norm of weights: 6.038276144791746
-----
Iteration Number: 3730
Loss: 41.95058836279384
12 norm of gradients: 0.3249760085428992
12 norm of weights: 6.03817427905155
-----
Iteration Number: 3731
Loss: 41.94531118598661
12 norm of gradients: 0.3249023746422316
12 norm of weights: 6.038072427131434
-----
Iteration Number: 3732
Loss: 41.940036372991464
12 norm of gradients: 0.32482878725712566
12 norm of weights: 6.037970589024877
-----
Iteration Number: 3733
Loss: 41.93476392181922
12 norm of gradients: 0.32475524635053005
12 norm of weights: 6.037868764725365
-----
Iteration Number: 3734
Loss: 41.92949383047293
12 norm of gradients: 0.3246817518854102
12 norm of weights: 6.03776695422639
-----
Iteration Number: 3735
Loss: 41.92422609694899
12 norm of gradients: 0.3246083038247489
12 norm of weights: 6.037665157521449
-----
Iteration Number: 3736

```

```

Loss: 41.918960719227414
l2 norm of gradients: 0.3245349021315456
l2 norm of weights: 6.037563374604049
-----
Iteration Number: 3737
Loss: 41.91369769535413
l2 norm of gradients: 0.3244615467688175
l2 norm of weights: 6.0374616054677
-----
Iteration Number: 3738
Loss: 41.90843702331086
l2 norm of gradients: 0.32438823769959835
l2 norm of weights: 6.037359850105921
-----
Iteration Number: 3739
Loss: 41.90317870109951
l2 norm of gradients: 0.32431497488693967
l2 norm of weights: 6.037258108512238
-----
Iteration Number: 3740
Loss: 41.89792272676994
l2 norm of gradients: 0.32424175829391
l2 norm of weights: 6.0371563806801785
-----
Iteration Number: 3741
Loss: 41.89266909831471
l2 norm of gradients: 0.3241685878835953
l2 norm of weights: 6.037054666603283
-----
Iteration Number: 3742
Loss: 41.88741781375408
l2 norm of gradients: 0.32409546361909886
l2 norm of weights: 6.036952966275094
-----
Iteration Number: 3743
Loss: 41.88216887110794
l2 norm of gradients: 0.3240223854635418
l2 norm of weights: 6.036851279689163
-----
Iteration Number: 3744
Loss: 41.87692226838964
l2 norm of gradients: 0.32394935338006203
l2 norm of weights: 6.036749606839046
-----
Iteration Number: 3745
Loss: 41.87167800362236
l2 norm of gradients: 0.32387636733181563
l2 norm of weights: 6.036647947718307

```

```

-----
Iteration Number: 3746
Loss: 41.866436074848224
l2 norm of gradients: 0.32380342728197603
l2 norm of weights: 6.036546302320517
-----
Iteration Number: 3747
Loss: 41.861196480072664
l2 norm of gradients: 0.3237305331937343
l2 norm of weights: 6.0364446706392485
-----
Iteration Number: 3748
Loss: 41.85595921734895
l2 norm of gradients: 0.3236576850302993
l2 norm of weights: 6.036343052668088
-----
Iteration Number: 3749
Loss: 41.850724284693825
l2 norm of gradients: 0.32358488275489755
l2 norm of weights: 6.036241448400625
-----
Iteration Number: 3750
Loss: 41.84549168013685
l2 norm of gradients: 0.3235121263307732
l2 norm of weights: 6.036139857830452
-----
Iteration Number: 3751
Loss: 41.84026140173161
l2 norm of gradients: 0.32343941572118867
l2 norm of weights: 6.036038280951174
-----
Iteration Number: 3752
Loss: 41.83503344750776
l2 norm of gradients: 0.3233667508894238
l2 norm of weights: 6.035936717756397
-----
Iteration Number: 3753
Loss: 41.829807815504054
l2 norm of gradients: 0.3232941317987767
l2 norm of weights: 6.035835168239736
-----
Iteration Number: 3754
Loss: 41.82458450375291
l2 norm of gradients: 0.32322155841256317
l2 norm of weights: 6.035733632394814
-----
Iteration Number: 3755
Loss: 41.81936351032121

```

```

12 norm of gradients: 0.3231490306941173
12 norm of weights: 6.035632110215258
-----
Iteration Number: 3756
Loss: 41.81414483323783
12 norm of gradients: 0.323076548606791
12 norm of weights: 6.035530601694701
-----
Iteration Number: 3757
Loss: 41.808928470565775
12 norm of gradients: 0.3230041121139547
12 norm of weights: 6.035429106826784
-----
Iteration Number: 3758
Loss: 41.80371442033805
12 norm of gradients: 0.3229317211789964
12 norm of weights: 6.035327625605152
-----
Iteration Number: 3759
Loss: 41.79850268062996
12 norm of gradients: 0.32285937576532275
12 norm of weights: 6.035226158023461
-----
Iteration Number: 3760
Loss: 41.793293249467325
12 norm of gradients: 0.3227870758363586
12 norm of weights: 6.0351247040753675
-----
Iteration Number: 3761
Loss: 41.78808612495107
12 norm of gradients: 0.3227148213555468
12 norm of weights: 6.035023263754539
-----
Iteration Number: 3762
Loss: 41.782881305091095
12 norm of gradients: 0.32264261228634916
12 norm of weights: 6.034921837054647
-----
Iteration Number: 3763
Loss: 41.777678787989984
12 norm of gradients: 0.3225704485922452
12 norm of weights: 6.034820423969368
-----
Iteration Number: 3764
Loss: 41.77247857168076
12 norm of gradients: 0.32249833023673324
12 norm of weights: 6.0347190244923885
-----

```

Iteration Number: 3765
Loss: 41.767280654260645
l2 norm of gradients: 0.32242625718333
l2 norm of weights: 6.034617638617401

Iteration Number: 3766
Loss: 41.76208503376009
l2 norm of gradients: 0.32235422939557073
l2 norm of weights: 6.034516266338098

Iteration Number: 3767
Loss: 41.756891708272406
l2 norm of gradients: 0.3222822468370094
l2 norm of weights: 6.034414907648188

Iteration Number: 3768
Loss: 41.751700675867035
l2 norm of gradients: 0.3222103094712182
l2 norm of weights: 6.034313562541377

Iteration Number: 3769
Loss: 41.74651193461335
l2 norm of gradients: 0.32213841726178843
l2 norm of weights: 6.034212231011383

Iteration Number: 3770
Loss: 41.74132548259764
l2 norm of gradients: 0.3220665701723297
l2 norm of weights: 6.034110913051927

Iteration Number: 3771
Loss: 41.73614131789446
l2 norm of gradients: 0.3219947681664707
l2 norm of weights: 6.034009608656739

Iteration Number: 3772
Loss: 41.73095943857362
l2 norm of gradients: 0.3219230112078587
l2 norm of weights: 6.033908317819553

Iteration Number: 3773
Loss: 41.725779842722346
l2 norm of gradients: 0.32185129926015965
l2 norm of weights: 6.03380704053411

Iteration Number: 3774
Loss: 41.72060252842681
l2 norm of gradients: 0.32177963228705897


```
12 norm of weights: 6.033705776794157
-----
Iteration Number: 3775
Loss: 41.71542749377859
12 norm of gradients: 0.32170801025226026
12 norm of weights: 6.033604526593448
-----
Iteration Number: 3776
Loss: 41.710254736847034
12 norm of gradients: 0.3216364331194866
12 norm of weights: 6.0335032899257435
-----
Iteration Number: 3777
Loss: 41.7050842557453
12 norm of gradients: 0.32156490085247974
12 norm of weights: 6.033402066784807
-----
Iteration Number: 3778
Loss: 41.69991604856947
12 norm of gradients: 0.32149341341500076
12 norm of weights: 6.033300857164412
-----
Iteration Number: 3779
Loss: 41.69475011338826
12 norm of gradients: 0.32142197077082957
12 norm of weights: 6.033199661058337
-----
Iteration Number: 3780
Loss: 41.68958644832678
12 norm of gradients: 0.32135057288376534
12 norm of weights: 6.033098478460365
-----
Iteration Number: 3781
Loss: 41.68442505146076
12 norm of gradients: 0.3212792197176265
12 norm of weights: 6.03299730936429
-----
Iteration Number: 3782
Loss: 41.679265920891666
12 norm of gradients: 0.32120791123625037
12 norm of weights: 6.032896153763906
-----
Iteration Number: 3783
Loss: 41.67410905475329
12 norm of gradients: 0.32113664740349374
12 norm of weights: 6.032795011653016
-----
Iteration Number: 3784
```

```

Loss: 41.66895445111472
l2 norm of gradients: 0.32106542818323286
l2 norm of weights: 6.03269388302543
-----
Iteration Number: 3785
Loss: 41.663802108095915
l2 norm of gradients: 0.32099425353936284
l2 norm of weights: 6.032592767874962
-----
Iteration Number: 3786
Loss: 41.658652023808806
l2 norm of gradients: 0.32092312343579876
l2 norm of weights: 6.032491666195435
-----
Iteration Number: 3787
Loss: 41.653504196373525
l2 norm of gradients: 0.3208520378364744
l2 norm of weights: 6.032390577980677
-----
Iteration Number: 3788
Loss: 41.648358623881904
l2 norm of gradients: 0.32078099670534377
l2 norm of weights: 6.032289503224519
-----
Iteration Number: 3789
Loss: 41.64321530445593
l2 norm of gradients: 0.3207100000063796
l2 norm of weights: 6.0321884419208045
-----
Iteration Number: 3790
Loss: 41.63807423622969
l2 norm of gradients: 0.3206390477035748
l2 norm of weights: 6.032087394063376
-----
Iteration Number: 3791
Loss: 41.632935417301375
l2 norm of gradients: 0.32056813976094145
l2 norm of weights: 6.031986359646087
-----
Iteration Number: 3792
Loss: 41.627798845803625
l2 norm of gradients: 0.3204972761425115
l2 norm of weights: 6.031885338662796
-----
Iteration Number: 3793
Loss: 41.62266451986426
l2 norm of gradients: 0.32042645681233617
l2 norm of weights: 6.0317843311073664

```

```

-----
Iteration Number: 3794
Loss: 41.61753243759503
l2 norm of gradients: 0.3203556817344867
l2 norm of weights: 6.031683336973669
-----
Iteration Number: 3795
Loss: 41.61240259713296
l2 norm of gradients: 0.3202849508730541
l2 norm of weights: 6.031582356255579
-----
Iteration Number: 3796
Loss: 41.60727499659955
l2 norm of gradients: 0.32021426419214877
l2 norm of weights: 6.03148138894698
-----
Iteration Number: 3797
Loss: 41.602149634123855
l2 norm of gradients: 0.3201436216559013
l2 norm of weights: 6.03138043504176
-----
Iteration Number: 3798
Loss: 41.59702650786945
l2 norm of gradients: 0.32007302322846193
l2 norm of weights: 6.031279494533814
-----
Iteration Number: 3799
Loss: 41.591905615927246
l2 norm of gradients: 0.32000246887400075
l2 norm of weights: 6.031178567417042
-----
Iteration Number: 3800
Loss: 41.58678695646698
l2 norm of gradients: 0.3199319585567078
l2 norm of weights: 6.031077653685351
-----
Iteration Number: 3801
Loss: 41.58167052761896
l2 norm of gradients: 0.31986149224079313
l2 norm of weights: 6.030976753332653
-----
Iteration Number: 3802
Loss: 41.5765563275257
l2 norm of gradients: 0.31979106989048683
l2 norm of weights: 6.030875866352867
-----
Iteration Number: 3803
Loss: 41.57144435432243

```

```

12 norm of gradients: 0.3197206914700388
12 norm of weights: 6.030774992739918
-----
Iteration Number: 3804
Loss: 41.566334606165626
12 norm of gradients: 0.31965035694371907
12 norm of weights: 6.030674132487735
-----
Iteration Number: 3805
Loss: 41.56122708121317
12 norm of gradients: 0.319580066275818
12 norm of weights: 6.030573285590257
-----
Iteration Number: 3806
Loss: 41.55612177758366
12 norm of gradients: 0.31950981943064577
12 norm of weights: 6.030472452041425
-----
Iteration Number: 3807
Loss: 41.5510186934452
12 norm of gradients: 0.3194396163725329
12 norm of weights: 6.030371631835188
-----
Iteration Number: 3808
Loss: 41.54591782695883
12 norm of gradients: 0.31936945706583003
12 norm of weights: 6.030270824965502
-----
Iteration Number: 3809
Loss: 41.54081917626402
12 norm of gradients: 0.3192993414749082
12 norm of weights: 6.030170031426326
-----
Iteration Number: 3810
Loss: 41.53572273953652
12 norm of gradients: 0.31922926956415854
12 norm of weights: 6.030069251211626
-----
Iteration Number: 3811
Loss: 41.53062851492285
12 norm of gradients: 0.31915924129799256
12 norm of weights: 6.029968484315375
-----
Iteration Number: 3812
Loss: 41.52553650059298
12 norm of gradients: 0.31908925664084214
12 norm of weights: 6.029867730731552
-----

```

Iteration Number: 3813
Loss: 41.52044669469465
l2 norm of gradients: 0.3190193155571596
l2 norm of weights: 6.029766990454141

Iteration Number: 3814
Loss: 41.5153590954243
l2 norm of gradients: 0.3189494180114175
l2 norm of weights: 6.029666263477132

Iteration Number: 3815
Loss: 41.510273700908435
l2 norm of gradients: 0.318879563968109
l2 norm of weights: 6.029565549794523

Iteration Number: 3816
Loss: 41.50519050935402
l2 norm of gradients: 0.31880975339174783
l2 norm of weights: 6.029464849400313

Iteration Number: 3817
Loss: 41.50010951891448
l2 norm of gradients: 0.31873998624686783
l2 norm of weights: 6.0293641622885135

Iteration Number: 3818
Loss: 41.495030727765794
l2 norm of gradients: 0.31867026249802394
l2 norm of weights: 6.029263488453135

Iteration Number: 3819
Loss: 41.489954134077045
l2 norm of gradients: 0.31860058210979136
l2 norm of weights: 6.029162827888199

Iteration Number: 3820
Loss: 41.48487973603776
l2 norm of gradients: 0.3185309450467659
l2 norm of weights: 6.029062180587732

Iteration Number: 3821
Loss: 41.479807531818935
l2 norm of gradients: 0.3184613512735641
l2 norm of weights: 6.028961546545765

Iteration Number: 3822
Loss: 41.47473751960306
l2 norm of gradients: 0.31839180075482326

```

12 norm of weights: 6.028860925756335
-----
Iteration Number: 3823
Loss: 41.46966969757725
12 norm of gradients: 0.3183222934552014
12 norm of weights: 6.028760318213485
-----
Iteration Number: 3824
Loss: 41.46460406392903
12 norm of gradients: 0.31825282933937704
12 norm of weights: 6.028659723911265
-----
Iteration Number: 3825
Loss: 41.4595406168493
12 norm of gradients: 0.3181834083720497
12 norm of weights: 6.02855914284373
-----
Iteration Number: 3826
Loss: 41.45447935449935
12 norm of gradients: 0.31811403051793985
12 norm of weights: 6.028458575004941
-----
Iteration Number: 3827
Loss: 41.4494202751026
12 norm of gradients: 0.31804469574178845
12 norm of weights: 6.028358020388965
-----
Iteration Number: 3828
Loss: 41.44436337683838
12 norm of gradients: 0.3179754040083576
12 norm of weights: 6.0282574789898735
-----
Iteration Number: 3829
Loss: 41.4393086578938
12 norm of gradients: 0.3179061552824302
12 norm of weights: 6.028156950801745
-----
Iteration Number: 3830
Loss: 41.434256116482906
12 norm of gradients: 0.31783694952881025
12 norm of weights: 6.028056435818664
-----
Iteration Number: 3831
Loss: 41.42920575080747
12 norm of gradients: 0.31776778671232253
12 norm of weights: 6.027955934034722
-----
Iteration Number: 3832

```

Loss: 41.4241575590608
l2 norm of gradients: 0.31769866679781295
l2 norm of weights: 6.027855445444012

Iteration Number: 3833
Loss: 41.41911153943707
l2 norm of gradients: 0.3176295897501484
l2 norm of weights: 6.027754970040638

Iteration Number: 3834
Loss: 41.414067690136385
l2 norm of gradients: 0.31756055553421686
l2 norm of weights: 6.027654507818705

Iteration Number: 3835
Loss: 41.40902600937965
l2 norm of gradients: 0.3174915641149274
l2 norm of weights: 6.027554058772329

Iteration Number: 3836
Loss: 41.40398649538359
l2 norm of gradients: 0.3174226154572102
l2 norm of weights: 6.027453622895626

Iteration Number: 3837
Loss: 41.39894914634143
l2 norm of gradients: 0.3173537095260167
l2 norm of weights: 6.027353200182723

Iteration Number: 3838
Loss: 41.39391396047404
l2 norm of gradients: 0.3172848462863194
l2 norm of weights: 6.02725279062775

Iteration Number: 3839
Loss: 41.38888093599723
l2 norm of gradients: 0.3172160257031121
l2 norm of weights: 6.027152394224842

Iteration Number: 3840
Loss: 41.3838500711137
l2 norm of gradients: 0.31714724774140973
l2 norm of weights: 6.027052010968142

Iteration Number: 3841
Loss: 41.378821364058744
l2 norm of gradients: 0.31707851236624873
l2 norm of weights: 6.026951640851798

```
-----  
Iteration Number: 3842  
Loss: 41.373794813053685  
l2 norm of gradients: 0.31700981954268675  
l2 norm of weights: 6.026851283869963  
-----  
Iteration Number: 3843  
Loss: 41.36877041630293  
l2 norm of gradients: 0.3169411692358025  
l2 norm of weights: 6.026750940016795  
-----  
Iteration Number: 3844  
Loss: 41.363748172053384  
l2 norm of gradients: 0.3168725614106965  
l2 norm of weights: 6.0266506092864605  
-----  
Iteration Number: 3845  
Loss: 41.35872807851361  
l2 norm of gradients: 0.31680399603249026  
l2 norm of weights: 6.026550291673129  
-----  
Iteration Number: 3846  
Loss: 41.353710133916785  
l2 norm of gradients: 0.3167354730663271  
l2 norm of weights: 6.026449987170977  
-----  
Iteration Number: 3847  
Loss: 41.348694336492244  
l2 norm of gradients: 0.31666699247737146  
l2 norm of weights: 6.026349695774188  
-----  
Iteration Number: 3848  
Loss: 41.343680684470854  
l2 norm of gradients: 0.3165985542308094  
l2 norm of weights: 6.026249417476946  
-----  
Iteration Number: 3849  
Loss: 41.33866917609146  
l2 norm of gradients: 0.31653015829184844  
l2 norm of weights: 6.026149152273448  
-----  
Iteration Number: 3850  
Loss: 41.333659809584894  
l2 norm of gradients: 0.31646180462571794  
l2 norm of weights: 6.0260489001578925  
-----  
Iteration Number: 3851  
Loss: 41.32865258320145
```



```

12 norm of gradients: 0.3163934931976681
12 norm of weights: 6.025948661124482
-----
Iteration Number: 3852
Loss: 41.323647495162156
12 norm of gradients: 0.3163252239729714
12 norm of weights: 6.025848435167428
-----
Iteration Number: 3853
Loss: 41.31864454370971
12 norm of gradients: 0.31625699691692183
12 norm of weights: 6.025748222280947
-----
Iteration Number: 3854
Loss: 41.31364372710976
12 norm of gradients: 0.31618881199483456
12 norm of weights: 6.02564802245926
-----
Iteration Number: 3855
Loss: 41.30864504357857
12 norm of gradients: 0.316120669172047
12 norm of weights: 6.025547835696593
-----
Iteration Number: 3856
Loss: 41.30364849137906
12 norm of gradients: 0.31605256841391793
12 norm of weights: 6.025447661987182
-----
Iteration Number: 3857
Loss: 41.29865406875539
12 norm of gradients: 0.31598450968582786
12 norm of weights: 6.025347501325263
-----
Iteration Number: 3858
Loss: 41.29366177396621
12 norm of gradients: 0.31591649295317914
12 norm of weights: 6.0252473537050815
-----
Iteration Number: 3859
Loss: 41.288671605246954
12 norm of gradients: 0.315848518181396
12 norm of weights: 6.0251472191208855
-----
Iteration Number: 3860
Loss: 41.283683560876455
12 norm of gradients: 0.3157805853359241
12 norm of weights: 6.025047097566932
-----

```

Iteration Number: 3861
Loss: 41.27869763908798
l2 norm of gradients: 0.3157126943822314
l2 norm of weights: 6.024946989037481

Iteration Number: 3862
Loss: 41.273713838157434
l2 norm of gradients: 0.3156448452858073
l2 norm of weights: 6.024846893526799

Iteration Number: 3863
Loss: 41.26873215633972
l2 norm of gradients: 0.31557703801216314
l2 norm of weights: 6.02474681102916

Iteration Number: 3864
Loss: 41.263752591886075
l2 norm of gradients: 0.31550927252683253
l2 norm of weights: 6.024646741538838

Iteration Number: 3865
Loss: 41.258775143079795
l2 norm of gradients: 0.3154415487953707
l2 norm of weights: 6.02454668505012

Iteration Number: 3866
Loss: 41.25379980817707
l2 norm of gradients: 0.31537386678335466
l2 norm of weights: 6.024446641557293

Iteration Number: 3867
Loss: 41.24882658543526
l2 norm of gradients: 0.3153062264563837
l2 norm of weights: 6.024346611054651

Iteration Number: 3868
Loss: 41.24385547314597
l2 norm of gradients: 0.315238627780079
l2 norm of weights: 6.0242465935364935

Iteration Number: 3869
Loss: 41.238886469562736
l2 norm of gradients: 0.3151710707200838
l2 norm of weights: 6.024146588997128

Iteration Number: 3870
Loss: 41.233919572968354
l2 norm of gradients: 0.31510355524206335

```
l2 norm of weights: 6.024046597430863
-----
Iteration Number: 3871
Loss: 41.228954781639196
l2 norm of gradients: 0.315036081311705
l2 norm of weights: 6.023946618832018
-----
Iteration Number: 3872
Loss: 41.223992093839776
l2 norm of gradients: 0.31496864889471804
l2 norm of weights: 6.023846653194912
-----
Iteration Number: 3873
Loss: 41.21903150786382
l2 norm of gradients: 0.31490125795683416
l2 norm of weights: 6.023746700513875
-----
Iteration Number: 3874
Loss: 41.21407302198651
l2 norm of gradients: 0.31483390846380693
l2 norm of weights: 6.0236467607832385
-----
Iteration Number: 3875
Loss: 41.20911663448658
l2 norm of gradients: 0.31476660038141224
l2 norm of weights: 6.023546833997341
-----
Iteration Number: 3876
Loss: 41.20416234366256
l2 norm of gradients: 0.3146993336754479
l2 norm of weights: 6.023446920150527
-----
Iteration Number: 3877
Loss: 41.19921014777415
l2 norm of gradients: 0.31463210831173444
l2 norm of weights: 6.023347019237146
-----
Iteration Number: 3878
Loss: 41.19426004513554
l2 norm of gradients: 0.31456492425611415
l2 norm of weights: 6.023247131251553
-----
Iteration Number: 3879
Loss: 41.18931203402448
l2 norm of gradients: 0.31449778147445173
l2 norm of weights: 6.023147256188108
-----
Iteration Number: 3880
```

Loss: 41.1843661127338
l2 norm of gradients: 0.31443067993263424
l2 norm of weights: 6.023047394041178

Iteration Number: 3881
Loss: 41.179422279562694
l2 norm of gradients: 0.3143636195965709
l2 norm of weights: 6.022947544805133

Iteration Number: 3882
Loss: 41.17448053280465
l2 norm of gradients: 0.3142966004321932
l2 norm of weights: 6.022847708474351

Iteration Number: 3883
Loss: 41.16954087075315
l2 norm of gradients: 0.31422962240545527
l2 norm of weights: 6.022747885043214

Iteration Number: 3884
Loss: 41.164603291704104
l2 norm of gradients: 0.3141626854823334
l2 norm of weights: 6.022648074506108

Iteration Number: 3885
Loss: 41.15966779397839
l2 norm of gradients: 0.31409578962882595
l2 norm of weights: 6.022548276857429

Iteration Number: 3886
Loss: 41.154734375875
l2 norm of gradients: 0.31402893481095434
l2 norm of weights: 6.022448492091574

Iteration Number: 3887
Loss: 41.149803035670146
l2 norm of gradients: 0.313962120994762
l2 norm of weights: 6.022348720202948

Iteration Number: 3888
Loss: 41.14487377168993
l2 norm of gradients: 0.3138953481463148
l2 norm of weights: 6.022248961185959

Iteration Number: 3889
Loss: 41.13994658224517
l2 norm of gradients: 0.3138286162317011
l2 norm of weights: 6.022149215035022

```
-----
Iteration Number: 3890
Loss: 41.135021465658824
l2 norm of gradients: 0.31376192521703195
l2 norm of weights: 6.022049481744559
-----
Iteration Number: 3891
Loss: 41.130098420211446
l2 norm of gradients: 0.31369527506844064
l2 norm of weights: 6.021949761308994
-----
Iteration Number: 3892
Loss: 41.12517744424101
l2 norm of gradients: 0.31362866575208315
l2 norm of weights: 6.0218500537227575
-----
Iteration Number: 3893
Loss: 41.120258536048695
l2 norm of gradients: 0.3135620972341379
l2 norm of weights: 6.021750358980286
-----
Iteration Number: 3894
Loss: 41.11534169396466
l2 norm of gradients: 0.3134955694808059
l2 norm of weights: 6.021650677076024
-----
Iteration Number: 3895
Loss: 41.11042691631054
l2 norm of gradients: 0.3134290824583109
l2 norm of weights: 6.021551008004415
-----
Iteration Number: 3896
Loss: 41.105514201394826
l2 norm of gradients: 0.31336263613289894
l2 norm of weights: 6.021451351759913
-----
Iteration Number: 3897
Loss: 41.10060354753045
l2 norm of gradients: 0.313296230470839
l2 norm of weights: 6.021351708336975
-----
Iteration Number: 3898
Loss: 41.09569495306772
l2 norm of gradients: 0.3132298654384225
l2 norm of weights: 6.021252077730065
-----
Iteration Number: 3899
Loss: 41.090788416329495
```

```

12 norm of gradients: 0.3131635410019636
12 norm of weights: 6.021152459933652
-----
Iteration Number: 3900
Loss: 41.085883935639835
12 norm of gradients: 0.31309725712779907
12 norm of weights: 6.021052854942209
-----
Iteration Number: 3901
Loss: 41.08098150932116
12 norm of gradients: 0.3130310137822886
12 norm of weights: 6.020953262750215
-----
Iteration Number: 3902
Loss: 41.076081135720266
12 norm of gradients: 0.31296481093181416
12 norm of weights: 6.020853683352154
-----
Iteration Number: 3903
Loss: 41.07118281315656
12 norm of gradients: 0.31289864854278093
12 norm of weights: 6.020754116742517
-----
Iteration Number: 3904
Loss: 41.06628653997491
12 norm of gradients: 0.3128325265816165
12 norm of weights: 6.020654562915799
-----
Iteration Number: 3905
Loss: 41.06139231450507
12 norm of gradients: 0.3127664450147713
12 norm of weights: 6.020555021866499
-----
Iteration Number: 3906
Loss: 41.056500135079084
12 norm of gradients: 0.3127004038087188
12 norm of weights: 6.0204554935891235
-----
Iteration Number: 3907
Loss: 41.05161000006434
12 norm of gradients: 0.31263440292995487
12 norm of weights: 6.020355978078184
-----
Iteration Number: 3908
Loss: 41.04672190779222
12 norm of gradients: 0.3125684423449985
12 norm of weights: 6.020256475328196
-----

```

Iteration Number: 3909
Loss: 41.0418358566021
l2 norm of gradients: 0.31250252202039147
l2 norm of weights: 6.020156985333681

Iteration Number: 3910
Loss: 41.03695184483684
l2 norm of gradients: 0.31243664192269827
l2 norm of weights: 6.020057508089167

Iteration Number: 3911
Loss: 41.0320698708565
l2 norm of gradients: 0.31237080201850637
l2 norm of weights: 6.019958043589184

Iteration Number: 3912
Loss: 41.02718993300343
l2 norm of gradients: 0.3123050022744261
l2 norm of weights: 6.01985859182827

Iteration Number: 3913
Loss: 41.02231202962989
l2 norm of gradients: 0.3122392426570908
l2 norm of weights: 6.019759152800968

Iteration Number: 3914
Loss: 41.01743615909161
l2 norm of gradients: 0.3121735231331565
l2 norm of weights: 6.019659726501825

Iteration Number: 3915
Loss: 41.0125623197429
l2 norm of gradients: 0.3121078436693024
l2 norm of weights: 6.019560312925394

Iteration Number: 3916
Loss: 41.00769050993558
l2 norm of gradients: 0.3120422042322305
l2 norm of weights: 6.019460912066235

Iteration Number: 3917
Loss: 41.00282072804395
l2 norm of gradients: 0.31197660478866585
l2 norm of weights: 6.019361523918908

Iteration Number: 3918
Loss: 40.997952972415284
l2 norm of gradients: 0.3119110453053565

```

12 norm of weights: 6.019262148477986
-----
Iteration Number: 3919
Loss: 40.99308724141774
12 norm of gradients: 0.31184552574907337
12 norm of weights: 6.01916278573804
-----
Iteration Number: 3920
Loss: 40.98822353340987
12 norm of gradients: 0.31178004608661053
12 norm of weights: 6.019063435693648
-----
Iteration Number: 3921
Loss: 40.983361846757596
12 norm of gradients: 0.311714606284785
12 norm of weights: 6.0189640983393975
-----
Iteration Number: 3922
Loss: 40.978502179832475
12 norm of gradients: 0.3116492063104368
12 norm of weights: 6.018864773669875
-----
Iteration Number: 3923
Loss: 40.97364453100687
12 norm of gradients: 0.3115838461304292
12 norm of weights: 6.018765461679677
-----
Iteration Number: 3924
Loss: 40.96878889864445
12 norm of gradients: 0.3115185257116485
12 norm of weights: 6.018666162363402
-----
Iteration Number: 3925
Loss: 40.96393528113843
12 norm of gradients: 0.31145324502100374
12 norm of weights: 6.018566875715656
-----
Iteration Number: 3926
Loss: 40.95908367683172
12 norm of gradients: 0.31138800402542766
12 norm of weights: 6.018467601731048
-----
Iteration Number: 3927
Loss: 40.9542340841116
12 norm of gradients: 0.31132280269187546
12 norm of weights: 6.018368340404194
-----
Iteration Number: 3928

```



```

Loss: 40.94938650137436
l2 norm of gradients: 0.3112576409873259
l2 norm of weights: 6.0182690917297155
-----
Iteration Number: 3929
Loss: 40.94454092698359
l2 norm of gradients: 0.31119251887878097
l2 norm of weights: 6.018169855702236
-----
Iteration Number: 3930
Loss: 40.93969735931675
l2 norm of gradients: 0.3111274363332654
l2 norm of weights: 6.018070632316387
-----
Iteration Number: 3931
Loss: 40.93485579678206
l2 norm of gradients: 0.3110623933178274
l2 norm of weights: 6.017971421566804
-----
Iteration Number: 3932
Loss: 40.930016237730136
l2 norm of gradients: 0.3109973897995384
l2 norm of weights: 6.017872223448128
-----
Iteration Number: 3933
Loss: 40.92517868056398
l2 norm of gradients: 0.3109324257454927
l2 norm of weights: 6.017773037955006
-----
Iteration Number: 3934
Loss: 40.92034312368621
l2 norm of gradients: 0.31086750112280814
l2 norm of weights: 6.017673865082088
-----
Iteration Number: 3935
Loss: 40.91550956546506
l2 norm of gradients: 0.3108026158986258
l2 norm of weights: 6.017574704824031
-----
Iteration Number: 3936
Loss: 40.91067800431254
l2 norm of gradients: 0.3107377700401095
l2 norm of weights: 6.017475557175496
-----
Iteration Number: 3937
Loss: 40.905848438609624
l2 norm of gradients: 0.31067296351444706
l2 norm of weights: 6.017376422131149

```

```
-----  
Iteration Number: 3938  
Loss: 40.90102086674978  
l2 norm of gradients: 0.31060819628884895  
l2 norm of weights: 6.017277299685663  
-----  
Iteration Number: 3939  
Loss: 40.89619528714166  
l2 norm of gradients: 0.3105434683305492  
l2 norm of weights: 6.017178189833713  
-----  
Iteration Number: 3940  
Loss: 40.89137169816864  
l2 norm of gradients: 0.3104787796068051  
l2 norm of weights: 6.017079092569982  
-----  
Iteration Number: 3941  
Loss: 40.88655009824184  
l2 norm of gradients: 0.31041413008489704  
l2 norm of weights: 6.016980007889154  
-----  
Iteration Number: 3942  
Loss: 40.8817304857769  
l2 norm of gradients: 0.3103495197321291  
l2 norm of weights: 6.016880935785925  
-----  
Iteration Number: 3943  
Loss: 40.87691285916031  
l2 norm of gradients: 0.3102849485158283  
l2 norm of weights: 6.0167818762549885  
-----  
Iteration Number: 3944  
Loss: 40.872097216782585  
l2 norm of gradients: 0.3102204164033452  
l2 norm of weights: 6.016682829291049  
-----  
Iteration Number: 3945  
Loss: 40.86728355708238  
l2 norm of gradients: 0.3101559233620536  
l2 norm of weights: 6.0165837948888115  
-----  
Iteration Number: 3946  
Loss: 40.86247187846651  
l2 norm of gradients: 0.31009146935935084  
l2 norm of weights: 6.016484773042988  
-----  
Iteration Number: 3947  
Loss: 40.8576621793247
```

```

12 norm of gradients: 0.31002705436265726
12 norm of weights: 6.016385763748297
-----
Iteration Number: 3948
Loss: 40.85285445808903
12 norm of gradients: 0.30996267833941704
12 norm of weights: 6.0162867669994595
-----
Iteration Number: 3949
Loss: 40.84804871316382
12 norm of gradients: 0.30989834125709753
12 norm of weights: 6.016187782791201
-----
Iteration Number: 3950
Loss: 40.843244942967964
12 norm of gradients: 0.30983404308318924
12 norm of weights: 6.016088811118257
-----
Iteration Number: 3951
Loss: 40.83844314591395
12 norm of gradients: 0.30976978378520653
12 norm of weights: 6.0159898519753625
-----
Iteration Number: 3952
Loss: 40.83364332043872
12 norm of gradients: 0.3097055633306869
12 norm of weights: 6.015890905357259
-----
Iteration Number: 3953
Loss: 40.8288454649429
12 norm of gradients: 0.30964138168719124
12 norm of weights: 6.015791971258695
-----
Iteration Number: 3954
Loss: 40.82404957787009
12 norm of gradients: 0.30957723882230415
12 norm of weights: 6.015693049674422
-----
Iteration Number: 3955
Loss: 40.81925565762287
12 norm of gradients: 0.30951313470363345
12 norm of weights: 6.015594140599196
-----
Iteration Number: 3956
Loss: 40.8144637026473
12 norm of gradients: 0.3094490692988105
12 norm of weights: 6.015495244027782
-----

```

Iteration Number: 3957
Loss: 40.80967371134945
l2 norm of gradients: 0.30938504257549004
l2 norm of weights: 6.015396359954942

Iteration Number: 3958
Loss: 40.804885682173754
l2 norm of gradients: 0.30932105450135033
l2 norm of weights: 6.015297488375453

Iteration Number: 3959
Loss: 40.800099613562544
l2 norm of gradients: 0.3092571050440933
l2 norm of weights: 6.015198629284089

Iteration Number: 3960
Loss: 40.79531550393385
l2 norm of gradients: 0.30919319417144414
l2 norm of weights: 6.015099782675631

Iteration Number: 3961
Loss: 40.79053335172827
l2 norm of gradients: 0.3091293218511516
l2 norm of weights: 6.015000948544869

Iteration Number: 3962
Loss: 40.78575315536264
l2 norm of gradients: 0.30906548805098794
l2 norm of weights: 6.014902126886591

Iteration Number: 3963
Loss: 40.780974913307624
l2 norm of gradients: 0.3090016927387491
l2 norm of weights: 6.014803317695597

Iteration Number: 3964
Loss: 40.77619862398447
l2 norm of gradients: 0.30893793588225427
l2 norm of weights: 6.014704520966686

Iteration Number: 3965
Loss: 40.7714242858423
l2 norm of gradients: 0.3088742174493465
l2 norm of weights: 6.014605736694665

Iteration Number: 3966
Loss: 40.76665189731319
l2 norm of gradients: 0.30881053740789216

```

12 norm of weights: 6.0145069648743466
-----
Iteration Number: 3967
Loss: 40.76188145683555
12 norm of gradients: 0.3087468957257811
12 norm of weights: 6.014408205500546
-----
Iteration Number: 3968
Loss: 40.75711296288747
12 norm of gradients: 0.3086832923709271
12 norm of weights: 6.014309458568085
-----
Iteration Number: 3969
Loss: 40.75234641390762
12 norm of gradients: 0.3086197273112671
12 norm of weights: 6.014210724071789
-----
Iteration Number: 3970
Loss: 40.74758180832384
12 norm of gradients: 0.30855620051476196
12 norm of weights: 6.014112002006486
-----
Iteration Number: 3971
Loss: 40.74281914460201
12 norm of gradients: 0.30849271194939576
12 norm of weights: 6.014013292367019
-----
Iteration Number: 3972
Loss: 40.73805842118984
12 norm of gradients: 0.30842926158317663
12 norm of weights: 6.013914595148223
-----
Iteration Number: 3973
Loss: 40.733299636556396
12 norm of gradients: 0.308365849384136
12 norm of weights: 6.013815910344945
-----
Iteration Number: 3974
Loss: 40.72854278914778
12 norm of gradients: 0.30830247532032895
12 norm of weights: 6.013717237952037
-----
Iteration Number: 3975
Loss: 40.72378787741812
12 norm of gradients: 0.3082391393598341
12 norm of weights: 6.013618577964352
-----
Iteration Number: 3976

```

Loss: 40.71903489982973
l2 norm of gradients: 0.30817584147075394
l2 norm of weights: 6.013519930376751

Iteration Number: 3977
Loss: 40.714283854857705
l2 norm of gradients: 0.3081125816212144
l2 norm of weights: 6.013421295184099

Iteration Number: 3978
Loss: 40.70953474094932
l2 norm of gradients: 0.3080493597793651
l2 norm of weights: 6.013322672381265

Iteration Number: 3979
Loss: 40.70478755657555
l2 norm of gradients: 0.3079861759133793
l2 norm of weights: 6.013224061963125

Iteration Number: 3980
Loss: 40.70004230020832
l2 norm of gradients: 0.3079230299914538
l2 norm of weights: 6.013125463924558

Iteration Number: 3981
Loss: 40.6952989702937
l2 norm of gradients: 0.3078599219818093
l2 norm of weights: 6.013026878260448

Iteration Number: 3982
Loss: 40.69055756533328
l2 norm of gradients: 0.30779685185268996
l2 norm of weights: 6.012928304965684

Iteration Number: 3983
Loss: 40.68581808378147
l2 norm of gradients: 0.3077338195723638
l2 norm of weights: 6.0128297440351615

Iteration Number: 3984
Loss: 40.681080524102946
l2 norm of gradients: 0.30767082510912225
l2 norm of weights: 6.012731195463776

Iteration Number: 3985
Loss: 40.67634488478922
l2 norm of gradients: 0.3076078684312806
l2 norm of weights: 6.012632659246434

```
-----
Iteration Number: 3986
Loss: 40.67161116431109
l2 norm of gradients: 0.30754494950717776
l2 norm of weights: 6.012534135378042
-----
Iteration Number: 3987
Loss: 40.66687936113901
l2 norm of gradients: 0.30748206830517655
l2 norm of weights: 6.012435623853515
-----
Iteration Number: 3988
Loss: 40.66214947375791
l2 norm of gradients: 0.307419224793663
l2 norm of weights: 6.012337124667768
-----
Iteration Number: 3989
Loss: 40.65742150064554
l2 norm of gradients: 0.3073564189410474
l2 norm of weights: 6.0122386378157255
-----
Iteration Number: 3990
Loss: 40.652695440293186
l2 norm of gradients: 0.30729365071576337
l2 norm of weights: 6.012140163292315
-----
Iteration Number: 3991
Loss: 40.64797129119912
l2 norm of gradients: 0.30723092008626834
l2 norm of weights: 6.012041701092468
-----
Iteration Number: 3992
Loss: 40.6432490518103
l2 norm of gradients: 0.30716822702104357
l2 norm of weights: 6.011943251211122
-----
Iteration Number: 3993
Loss: 40.63852872063602
l2 norm of gradients: 0.30710557148859385
l2 norm of weights: 6.011844813643218
-----
Iteration Number: 3994
Loss: 40.63381029617018
l2 norm of gradients: 0.30704295345744786
l2 norm of weights: 6.011746388383702
-----
Iteration Number: 3995
Loss: 40.629093776901406
```

```

12 norm of gradients: 0.3069803728961579
12 norm of weights: 6.011647975427527
-----
Iteration Number: 3996
Loss: 40.624379161323226
12 norm of gradients: 0.30691782977330023
12 norm of weights: 6.011549574769647
-----
Iteration Number: 3997
Loss: 40.61966644792262
12 norm of gradients: 0.3068553240574746
12 norm of weights: 6.011451186405022
-----
Iteration Number: 3998
Loss: 40.614955635190285
12 norm of gradients: 0.30679285571730447
12 norm of weights: 6.011352810328621
-----
Iteration Number: 3999
Loss: 40.610246721644145
12 norm of gradients: 0.30673042472143724
12 norm of weights: 6.0112544465354105
-----
Iteration Number: 4000
Loss: 40.60553970577379
12 norm of gradients: 0.30666803103854395
12 norm of weights: 6.0111560950203655
-----
Iteration Number: 4001
Loss: 40.60083458607004
12 norm of gradients: 0.30660567463731947
12 norm of weights: 6.011057755778467
-----
Iteration Number: 4002
Loss: 40.59613136103647
12 norm of gradients: 0.3065433554864824
12 norm of weights: 6.010959428804698
-----
Iteration Number: 4003
Loss: 40.59143002919977
12 norm of gradients: 0.3064810735547751
12 norm of weights: 6.010861114094048
-----
Iteration Number: 4004
Loss: 40.58673058903574
12 norm of gradients: 0.3064188288109636
12 norm of weights: 6.01076281164151
-----

```


Iteration Number: 4005
Loss: 40.58203303907664
l2 norm of gradients: 0.30635662122383794
l2 norm of weights: 6.010664521442083

Iteration Number: 4006
Loss: 40.577337377811254
l2 norm of gradients: 0.3062944507622116
l2 norm of weights: 6.010566243490768

Iteration Number: 4007
Loss: 40.57264360376528
l2 norm of gradients: 0.3062323173949222
l2 norm of weights: 6.010467977782574

Iteration Number: 4008
Loss: 40.567951715442625
l2 norm of gradients: 0.3061702210908309
l2 norm of weights: 6.010369724312513

Iteration Number: 4009
Loss: 40.5632617113579
l2 norm of gradients: 0.30610816181882267
l2 norm of weights: 6.010271483075601

Iteration Number: 4010
Loss: 40.55857359002775
l2 norm of gradients: 0.30604613954780624
l2 norm of weights: 6.010173254066861

Iteration Number: 4011
Loss: 40.55388734996107
l2 norm of gradients: 0.3059841542467143
l2 norm of weights: 6.010075037281318

Iteration Number: 4012
Loss: 40.54920298968007
l2 norm of gradients: 0.3059222058845032
l2 norm of weights: 6.009976832714003

Iteration Number: 4013
Loss: 40.544520507706935
l2 norm of gradients: 0.30586029443015306
l2 norm of weights: 6.009878640359951

Iteration Number: 4014
Loss: 40.539839902575814
l2 norm of gradients: 0.3057984198526678

```

12 norm of weights: 6.009780460214203
-----
Iteration Number: 4015
Loss: 40.53516117278639
12 norm of gradients: 0.3057365821210753
12 norm of weights: 6.009682292271803
-----
Iteration Number: 4016
Loss: 40.530484316871046
12 norm of gradients: 0.3056747812044269
12 norm of weights: 6.0095841365278
-----
Iteration Number: 4017
Loss: 40.52580933335675
12 norm of gradients: 0.30561301707179817
12 norm of weights: 6.009485992977249
-----
Iteration Number: 4018
Loss: 40.52113622078315
12 norm of gradients: 0.3055512896922881
12 norm of weights: 6.009387861615206
-----
Iteration Number: 4019
Loss: 40.51646497766414
12 norm of gradients: 0.3054895990350197
12 norm of weights: 6.009289742436737
-----
Iteration Number: 4020
Loss: 40.511795602534406
12 norm of gradients: 0.30542794506913973
12 norm of weights: 6.009191635436908
-----
Iteration Number: 4021
Loss: 40.50712809394389
12 norm of gradients: 0.30536632776381867
12 norm of weights: 6.009093540610791
-----
Iteration Number: 4022
Loss: 40.50246245039124
12 norm of gradients: 0.3053047470882511
12 norm of weights: 6.008995457953464
-----
Iteration Number: 4023
Loss: 40.49779867043718
12 norm of gradients: 0.30524320301165503
12 norm of weights: 6.008897387460007
-----
Iteration Number: 4024

```

```

Loss: 40.49313675261447
l2 norm of gradients: 0.30518169550327245
l2 norm of weights: 6.008799329125508
-----
Iteration Number: 4025
Loss: 40.48847669546392
l2 norm of gradients: 0.30512022453236937
l2 norm of weights: 6.008701282945055
-----
Iteration Number: 4026
Loss: 40.48381849751324
l2 norm of gradients: 0.3050587900682351
l2 norm of weights: 6.008603248913745
-----
Iteration Number: 4027
Loss: 40.47916215731402
l2 norm of gradients: 0.30499739208018334
l2 norm of weights: 6.008505227026677
-----
Iteration Number: 4028
Loss: 40.47450767342025
l2 norm of gradients: 0.30493603053755114
l2 norm of weights: 6.008407217278955
-----
Iteration Number: 4029
Loss: 40.469855044362475
l2 norm of gradients: 0.30487470540969974
l2 norm of weights: 6.008309219665688
-----
Iteration Number: 4030
Loss: 40.46520426868686
l2 norm of gradients: 0.3048134166660139
l2 norm of weights: 6.0082112341819895
-----
Iteration Number: 4031
Loss: 40.460555344949505
l2 norm of gradients: 0.30475216427590235
l2 norm of weights: 6.008113260822976
-----
Iteration Number: 4032
Loss: 40.455908271690866
l2 norm of gradients: 0.3046909482087976
l2 norm of weights: 6.0080152995837715
-----
Iteration Number: 4033
Loss: 40.45126304746669
l2 norm of gradients: 0.304629768434156
l2 norm of weights: 6.007917350459501

```

```
-----  
Iteration Number: 4034  
Loss: 40.446619670838864  
l2 norm of gradients: 0.30456862492145764  
l2 norm of weights: 6.007819413445299  
-----  
Iteration Number: 4035  
Loss: 40.44197814035526  
l2 norm of gradients: 0.3045075176402065  
l2 norm of weights: 6.007721488536298  
-----  
Iteration Number: 4036  
Loss: 40.43733845455866  
l2 norm of gradients: 0.3044464465599304  
l2 norm of weights: 6.007623575727639  
-----  
Iteration Number: 4037  
Loss: 40.43270061200904  
l2 norm of gradients: 0.30438541165018085  
l2 norm of weights: 6.0075256750144685  
-----  
Iteration Number: 4038  
Loss: 40.42806461129411  
l2 norm of gradients: 0.30432441288053325  
l2 norm of weights: 6.0074277863919345  
-----  
Iteration Number: 4039  
Loss: 40.42343045094217  
l2 norm of gradients: 0.30426345022058693  
l2 norm of weights: 6.007329909855191  
-----  
Iteration Number: 4040  
Loss: 40.418798129531226  
l2 norm of gradients: 0.30420252363996475  
l2 norm of weights: 6.007232045399396  
-----  
Iteration Number: 4041  
Loss: 40.41416764562601  
l2 norm of gradients: 0.30414163310831377  
l2 norm of weights: 6.007134193019714  
-----  
Iteration Number: 4042  
Loss: 40.40953899777008  
l2 norm of gradients: 0.30408077859530447  
l2 norm of weights: 6.00703635271131  
-----  
Iteration Number: 4043  
Loss: 40.40491218455224
```

```
12 norm of gradients: 0.3040199600706314
12 norm of weights: 6.006938524469357
-----
Iteration Number: 4044
Loss: 40.40028720453888
12 norm of gradients: 0.30395917750401275
12 norm of weights: 6.00684070828903
-----
Iteration Number: 4045
Loss: 40.39566405628904
12 norm of gradients: 0.3038984308651908
12 norm of weights: 6.006742904165511
-----
Iteration Number: 4046
Loss: 40.391042738390695
12 norm of gradients: 0.3038377201239313
12 norm of weights: 6.0066451120939846
-----
Iteration Number: 4047
Loss: 40.38642324939595
12 norm of gradients: 0.303777045250024
12 norm of weights: 6.0065473320696405
-----
Iteration Number: 4048
Loss: 40.3818055878992
12 norm of gradients: 0.3037164062132824
12 norm of weights: 6.006449564087671
-----
Iteration Number: 4049
Loss: 40.37718975245188
12 norm of gradients: 0.3036558029835439
12 norm of weights: 6.006351808143277
-----
Iteration Number: 4050
Loss: 40.37257574164179
12 norm of gradients: 0.30359523553066947
12 norm of weights: 6.00625406423166
-----
Iteration Number: 4051
Loss: 40.367963554066165
12 norm of gradients: 0.30353470382454417
12 norm of weights: 6.006156332348027
-----
Iteration Number: 4052
Loss: 40.36335318828963
12 norm of gradients: 0.30347420783507667
12 norm of weights: 6.006058612487591
-----
```

Iteration Number: 4053
Loss: 40.35874464288248
l2 norm of gradients: 0.30341374753219946
l2 norm of weights: 6.0059609046455655

Iteration Number: 4054
Loss: 40.354137916457276
l2 norm of gradients: 0.30335332288586886
l2 norm of weights: 6.005863208817173

Iteration Number: 4055
Loss: 40.34953300756881
l2 norm of gradients: 0.30329293386606504
l2 norm of weights: 6.0057655249976385

Iteration Number: 4056
Loss: 40.34492991481511
l2 norm of gradients: 0.30323258044279183
l2 norm of weights: 6.00566785318219

Iteration Number: 4057
Loss: 40.340328636796954
l2 norm of gradients: 0.303172262586077
l2 norm of weights: 6.0055701933660615

Iteration Number: 4058
Loss: 40.335729172074856
l2 norm of gradients: 0.303111980265972
l2 norm of weights: 6.005472545544491

Iteration Number: 4059
Loss: 40.331131519267316
l2 norm of gradients: 0.30305173345255204
l2 norm of weights: 6.005374909712721

Iteration Number: 4060
Loss: 40.326535676955636
l2 norm of gradients: 0.30299152211591623
l2 norm of weights: 6.005277285865998

Iteration Number: 4061
Loss: 40.32194164372983
l2 norm of gradients: 0.3029313462261873
l2 norm of weights: 6.005179673999574

Iteration Number: 4062
Loss: 40.31734941818937
l2 norm of gradients: 0.30287120575351195

```
l2 norm of weights: 6.005082074108703
-----
Iteration Number: 4063
Loss: 40.3127589989447
l2 norm of gradients: 0.30281110066806055
l2 norm of weights: 6.004984486188646
-----
Iteration Number: 4064
Loss: 40.30817038456958
l2 norm of gradients: 0.3027510309400271
l2 norm of weights: 6.004886910234667
-----
Iteration Number: 4065
Loss: 40.30358357367362
l2 norm of gradients: 0.3026909965396297
l2 norm of weights: 6.004789346242034
-----
Iteration Number: 4066
Loss: 40.29899856486363
l2 norm of gradients: 0.3026309974371099
l2 norm of weights: 6.00469179420602
-----
Iteration Number: 4067
Loss: 40.2944153567349
l2 norm of gradients: 0.3025710336027333
l2 norm of weights: 6.004594254121902
-----
Iteration Number: 4068
Loss: 40.28983394790582
l2 norm of gradients: 0.3025111050067889
l2 norm of weights: 6.004496725984963
-----
Iteration Number: 4069
Loss: 40.2852543369658
l2 norm of gradients: 0.30245121161958977
l2 norm of weights: 6.004399209790487
-----
Iteration Number: 4070
Loss: 40.28067652252033
l2 norm of gradients: 0.3023913534114726
l2 norm of weights: 6.004301705533765
-----
Iteration Number: 4071
Loss: 40.27610050319791
l2 norm of gradients: 0.30233153035279775
l2 norm of weights: 6.004204213210091
-----
Iteration Number: 4072
```

```
Loss: 40.271526277590155
12 norm of gradients: 0.3022717424139496
12 norm of weights: 6.004106732814764
-----
Iteration Number: 4073
Loss: 40.26695384431948
12 norm of gradients: 0.302211989565336
12 norm of weights: 6.004009264343087
-----
Iteration Number: 4074
Loss: 40.26238320198917
12 norm of gradients: 0.30215227177738857
12 norm of weights: 6.0039118077903675
-----
Iteration Number: 4075
Loss: 40.25781434921996
12 norm of gradients: 0.3020925890205628
12 norm of weights: 6.003814363151917
-----
Iteration Number: 4076
Loss: 40.25324728463076
12 norm of gradients: 0.30203294126533775
12 norm of weights: 6.003716930423051
-----
Iteration Number: 4077
Loss: 40.248682006834585
12 norm of gradients: 0.3019733284822164
12 norm of weights: 6.0036195095990905
-----
Iteration Number: 4078
Loss: 40.24411851444886
12 norm of gradients: 0.30191375064172526
12 norm of weights: 6.003522100675361
-----
Iteration Number: 4079
Loss: 40.23955680610686
12 norm of gradients: 0.30185420771441457
12 norm of weights: 6.003424703647189
-----
Iteration Number: 4080
Loss: 40.23499688041199
12 norm of gradients: 0.3017946996708585
12 norm of weights: 6.003327318509909
-----
Iteration Number: 4081
Loss: 40.23043873599793
12 norm of gradients: 0.30173522648165463
12 norm of weights: 6.003229945258856
```



```

-----
Iteration Number: 4082
Loss: 40.225882371490975
l2 norm of gradients: 0.30167578811742446
l2 norm of weights: 6.003132583889376
-----
Iteration Number: 4083
Loss: 40.22132778551949
l2 norm of gradients: 0.3016163845488132
l2 norm of weights: 6.00303523439681
-----
Iteration Number: 4084
Loss: 40.21677497670345
l2 norm of gradients: 0.3015570157464895
l2 norm of weights: 6.002937896776511
-----
Iteration Number: 4085
Loss: 40.21222394366937
l2 norm of gradients: 0.301497681681146
l2 norm of weights: 6.002840571023832
-----
Iteration Number: 4086
Loss: 40.207674685061356
l2 norm of gradients: 0.3014383823234988
l2 norm of weights: 6.0027432571341315
-----
Iteration Number: 4087
Loss: 40.20312719949699
l2 norm of gradients: 0.30137911764428793
l2 norm of weights: 6.002645955102774
-----
Iteration Number: 4088
Loss: 40.19858148563098
l2 norm of gradients: 0.30131988761427675
l2 norm of weights: 6.002548664925124
-----
Iteration Number: 4089
Loss: 40.19403754207707
l2 norm of gradients: 0.3012606922042527
l2 norm of weights: 6.002451386596553
-----
Iteration Number: 4090
Loss: 40.18949536748596
l2 norm of gradients: 0.3012015313850265
l2 norm of weights: 6.002354120112439
-----
Iteration Number: 4091
Loss: 40.184954960483154

```

```

12 norm of gradients: 0.3011424051274328
12 norm of weights: 6.0022568654681585
-----
Iteration Number: 4092
Loss: 40.180416319707035
12 norm of gradients: 0.30108331340232974
12 norm of weights: 6.002159622659096
-----
Iteration Number: 4093
Loss: 40.1758794438232
12 norm of gradients: 0.30102425618059914
12 norm of weights: 6.00206239168064
-----
Iteration Number: 4094
Loss: 40.171344331444935
12 norm of gradients: 0.30096523343314663
12 norm of weights: 6.001965172528183
-----
Iteration Number: 4095
Loss: 40.16681098123723
12 norm of gradients: 0.30090624513090125
12 norm of weights: 6.00186796519712
-----
Iteration Number: 4096
Loss: 40.16227939182081
12 norm of gradients: 0.30084729124481585
12 norm of weights: 6.001770769682853
-----
Iteration Number: 4097
Loss: 40.15774956187641
12 norm of gradients: 0.30078837174586676
12 norm of weights: 6.001673585980787
-----
Iteration Number: 4098
Loss: 40.15322149002008
12 norm of gradients: 0.300729486605054
12 norm of weights: 6.001576414086329
-----
Iteration Number: 4099
Loss: 40.14869517491996
12 norm of gradients: 0.30067063579340136
12 norm of weights: 6.0014792539948925
-----
Iteration Number: 4100
Loss: 40.14417061521616
12 norm of gradients: 0.3006118192819558
12 norm of weights: 6.001382105701897
-----

```

Iteration Number: 4101
Loss: 40.139647809579245
l2 norm of gradients: 0.3005530370417884
l2 norm of weights: 6.00128496920276

Iteration Number: 4102
Loss: 40.13512675664047
l2 norm of gradients: 0.30049428904399356
l2 norm of weights: 6.00118784449291

Iteration Number: 4103
Loss: 40.13060745507176
l2 norm of gradients: 0.3004355752596894
l2 norm of weights: 6.001090731567776

Iteration Number: 4104
Loss: 40.12608990351342
l2 norm of gradients: 0.30037689566001735
l2 norm of weights: 6.000993630422791

Iteration Number: 4105
Loss: 40.1215741006427
l2 norm of gradients: 0.30031825021614267
l2 norm of weights: 6.000896541053394

Iteration Number: 4106
Loss: 40.11706004510426
l2 norm of gradients: 0.30025963889925433
l2 norm of weights: 6.000799463455026

Iteration Number: 4107
Loss: 40.11254773556265
l2 norm of gradients: 0.3002010616805645
l2 norm of weights: 6.000702397623134

Iteration Number: 4108
Loss: 40.10803717069875
l2 norm of gradients: 0.30014251853130913
l2 norm of weights: 6.000605343553168

Iteration Number: 4109
Loss: 40.10352834914841
l2 norm of gradients: 0.3000840094227477
l2 norm of weights: 6.000508301240582

Iteration Number: 4110
Loss: 40.09902126959091
l2 norm of gradients: 0.30002553432616325

```
l2 norm of weights: 6.000411270680836
-----
Iteration Number: 4111
Loss: 40.094515930682746
l2 norm of gradients: 0.29996709321286225
l2 norm of weights: 6.000314251869391
-----
Iteration Number: 4112
Loss: 40.09001233109787
l2 norm of gradients: 0.2999086860541749
l2 norm of weights: 6.000217244801713
-----
Iteration Number: 4113
Loss: 40.085510469522056
l2 norm of gradients: 0.2998503128214547
l2 norm of weights: 6.000120249473275
-----
Iteration Number: 4114
Loss: 40.08101034460724
l2 norm of gradients: 0.2997919734860789
l2 norm of weights: 6.000023265879551
-----
Iteration Number: 4115
Loss: 40.076511955026845
l2 norm of gradients: 0.299733668019448
l2 norm of weights: 5.999926294016021
-----
Iteration Number: 4116
Loss: 40.072015299455046
l2 norm of gradients: 0.2996753963929863
l2 norm of weights: 5.999829333878165
-----
Iteration Number: 4117
Loss: 40.067520376586096
l2 norm of gradients: 0.29961715857814136
l2 norm of weights: 5.999732385461473
-----
Iteration Number: 4118
Loss: 40.06302718506017
l2 norm of gradients: 0.29955895454638437
l2 norm of weights: 5.9996354487614365
-----
Iteration Number: 4119
Loss: 40.05853572358671
l2 norm of gradients: 0.2995007842692099
l2 norm of weights: 5.999538523773549
-----
Iteration Number: 4120
```

Loss: 40.05404599082779
l2 norm of gradients: 0.2994426477181361
l2 norm of weights: 5.999441610493309

Iteration Number: 4121
Loss: 40.049557985475346
l2 norm of gradients: 0.2993845448647046
l2 norm of weights: 5.999344708916223

Iteration Number: 4122
Loss: 40.045071706205526
l2 norm of gradients: 0.2993264756804804
l2 norm of weights: 5.999247819037797

Iteration Number: 4123
Loss: 40.04058715169464
l2 norm of gradients: 0.299268440137052
l2 norm of weights: 5.999150940853542

Iteration Number: 4124
Loss: 40.03610432064319
l2 norm of gradients: 0.2992104382060313
l2 norm of weights: 5.999054074358973

Iteration Number: 4125
Loss: 40.03162321172604
l2 norm of gradients: 0.29915246985905375
l2 norm of weights: 5.998957219549611

Iteration Number: 4126
Loss: 40.02714382363045
l2 norm of gradients: 0.2990945350677781
l2 norm of weights: 5.99886037642098

Iteration Number: 4127
Loss: 40.022666155051624
l2 norm of gradients: 0.29903663380388673
l2 norm of weights: 5.998763544968605

Iteration Number: 4128
Loss: 40.01819020468298
l2 norm of gradients: 0.2989787660390851
l2 norm of weights: 5.99866672518802

Iteration Number: 4129
Loss: 40.01371597120506
l2 norm of gradients: 0.2989209317451025
l2 norm of weights: 5.99856991707476

```

-----
Iteration Number: 4130
Loss: 40.00924345331774
l2 norm of gradients: 0.29886313089369115
l2 norm of weights: 5.9984731206243636
-----
Iteration Number: 4131
Loss: 40.004772649714724
l2 norm of gradients: 0.2988053634566271
l2 norm of weights: 5.9983763358323765
-----
Iteration Number: 4132
Loss: 40.000303559096686
l2 norm of gradients: 0.2987476294057096
l2 norm of weights: 5.998279562694345
-----
Iteration Number: 4133
Loss: 39.9958361801497
l2 norm of gradients: 0.2986899287127612
l2 norm of weights: 5.99818280120582
-----
Iteration Number: 4134
Loss: 39.99137051158693
l2 norm of gradients: 0.2986322613496278
l2 norm of weights: 5.9980860513623595
-----
Iteration Number: 4135
Loss: 39.98690655209754
l2 norm of gradients: 0.29857462728817896
l2 norm of weights: 5.997989313159521
-----
Iteration Number: 4136
Loss: 39.98244430038423
l2 norm of gradients: 0.29851702650030726
l2 norm of weights: 5.997892586592869
-----
Iteration Number: 4137
Loss: 39.9779837551473
l2 norm of gradients: 0.2984594589579288
l2 norm of weights: 5.99779587165797
-----
Iteration Number: 4138
Loss: 39.97352491509943
l2 norm of gradients: 0.29840192463298293
l2 norm of weights: 5.9976991683503975
-----
Iteration Number: 4139
Loss: 39.96906777894097

```

```

12 norm of gradients: 0.2983444234974324
12 norm of weights: 5.997602476665726
-----
Iteration Number: 4140
Loss: 39.964612345389035
12 norm of gradients: 0.2982869555232633
12 norm of weights: 5.997505796599535
-----
Iteration Number: 4141
Loss: 39.96015861313902
12 norm of gradients: 0.29822952068248476
12 norm of weights: 5.997409128147408
-----
Iteration Number: 4142
Loss: 39.95570658090566
12 norm of gradients: 0.2981721189471296
12 norm of weights: 5.997312471304932
-----
Iteration Number: 4143
Loss: 39.95125624738436
12 norm of gradients: 0.29811475028925377
12 norm of weights: 5.9972158260677
-----
Iteration Number: 4144
Loss: 39.94680761132203
12 norm of gradients: 0.2980574146809364
12 norm of weights: 5.997119192431304
-----
Iteration Number: 4145
Loss: 39.94236067138916
12 norm of gradients: 0.29800011209428007
12 norm of weights: 5.997022570391346
-----
Iteration Number: 4146
Loss: 39.9379154263387
12 norm of gradients: 0.29794284250141034
12 norm of weights: 5.996925959943428
-----
Iteration Number: 4147
Loss: 39.93347187487967
12 norm of gradients: 0.2978856058744765
12 norm of weights: 5.996829361083159
-----
Iteration Number: 4148
Loss: 39.929030015709266
12 norm of gradients: 0.29782840218565065
12 norm of weights: 5.996732773806146
-----

```

Iteration Number: 4149
Loss: 39.9245898475693
l2 norm of gradients: 0.29777123140712836
l2 norm of weights: 5.996636198108008

Iteration Number: 4150
Loss: 39.920151369164316
l2 norm of gradients: 0.29771409351112826
l2 norm of weights: 5.996539633984361

Iteration Number: 4151
Loss: 39.915714579229814
l2 norm of gradients: 0.2976569884698925
l2 norm of weights: 5.996443081430829

Iteration Number: 4152
Loss: 39.91127947647547
l2 norm of gradients: 0.297599916255686
l2 norm of weights: 5.996346540443039

Iteration Number: 4153
Loss: 39.906846059626616
l2 norm of gradients: 0.29754287684079733
l2 norm of weights: 5.99625001101662

Iteration Number: 4154
Loss: 39.90241432743092
l2 norm of gradients: 0.29748587019753797
l2 norm of weights: 5.996153493147209

Iteration Number: 4155
Loss: 39.897984278591004
l2 norm of gradients: 0.29742889629824254
l2 norm of weights: 5.9960569868304425

Iteration Number: 4156
Loss: 39.89355591184208
l2 norm of gradients: 0.29737195511526926
l2 norm of weights: 5.995960492061962

Iteration Number: 4157
Loss: 39.889129225919575
l2 norm of gradients: 0.297315046620999
l2 norm of weights: 5.995864008837414

Iteration Number: 4158
Loss: 39.88470421954627
l2 norm of gradients: 0.2972581707878359


```

12 norm of weights: 5.99576753715245
-----
Iteration Number: 4159
Loss: 39.88028089146922
12 norm of gradients: 0.29720132758820755
12 norm of weights: 5.995671077002723
-----
Iteration Number: 4160
Loss: 39.8758592404106
12 norm of gradients: 0.2971445169945644
12 norm of weights: 5.995574628383891
-----
Iteration Number: 4161
Loss: 39.87143926510688
12 norm of gradients: 0.2970877389793802
12 norm of weights: 5.995478191291614
-----
Iteration Number: 4162
Loss: 39.86702096428161
12 norm of gradients: 0.2970309935151514
12 norm of weights: 5.995381765721559
-----
Iteration Number: 4163
Loss: 39.86260433671263
12 norm of gradients: 0.2969742805743981
12 norm of weights: 5.995285351669394
-----
Iteration Number: 4164
Loss: 39.85818938108891
12 norm of gradients: 0.29691760012966323
12 norm of weights: 5.995188949130793
-----
Iteration Number: 4165
Loss: 39.85377609618839
12 norm of gradients: 0.2968609521535129
12 norm of weights: 5.995092558101433
-----
Iteration Number: 4166
Loss: 39.849364480738735
12 norm of gradients: 0.29680433661853606
12 norm of weights: 5.9949961785769945
-----
Iteration Number: 4167
Loss: 39.844954533492334
12 norm of gradients: 0.29674775349734506
12 norm of weights: 5.994899810553162
-----
Iteration Number: 4168

```

Loss: 39.84054625317343
l2 norm of gradients: 0.296691202762575
l2 norm of weights: 5.994803454025624

Iteration Number: 4169
Loss: 39.836139638544815
l2 norm of gradients: 0.2966346843868844
l2 norm of weights: 5.994707108990073

Iteration Number: 4170
Loss: 39.83173468834008
l2 norm of gradients: 0.29657819834295435
l2 norm of weights: 5.994610775442205

Iteration Number: 4171
Loss: 39.82733140133555
l2 norm of gradients: 0.2965217446034894
l2 norm of weights: 5.994514453377721

Iteration Number: 4172
Loss: 39.82292977625201
l2 norm of gradients: 0.29646532314121676
l2 norm of weights: 5.9944181427923215

Iteration Number: 4173
Loss: 39.818529811849025
l2 norm of gradients: 0.29640893392888695
l2 norm of weights: 5.994321843681717

Iteration Number: 4174
Loss: 39.814131506883406
l2 norm of gradients: 0.2963525769392733
l2 norm of weights: 5.994225556041618

Iteration Number: 4175
Loss: 39.80973486011319
l2 norm of gradients: 0.296296252145172
l2 norm of weights: 5.994129279867739

Iteration Number: 4176
Loss: 39.80533987027029
l2 norm of gradients: 0.2962399595194027
l2 norm of weights: 5.9940330151558

Iteration Number: 4177
Loss: 39.800946536150825
l2 norm of gradients: 0.29618369903480746
l2 norm of weights: 5.993936761901522

```

-----
Iteration Number: 4178
Loss: 39.79655485647044
l2 norm of gradients: 0.2961274706642516
l2 norm of weights: 5.993840520100634
-----
Iteration Number: 4179
Loss: 39.79216483001789
l2 norm of gradients: 0.29607127438062325
l2 norm of weights: 5.9937442897488635
-----
Iteration Number: 4180
Loss: 39.787776455531706
l2 norm of gradients: 0.2960151101568337
l2 norm of weights: 5.993648070841946
-----
Iteration Number: 4181
Loss: 39.78338973179478
l2 norm of gradients: 0.2959589779658167
l2 norm of weights: 5.993551863375619
-----
Iteration Number: 4182
Loss: 39.77900465755927
l2 norm of gradients: 0.29590287778052954
l2 norm of weights: 5.993455667345625
-----
Iteration Number: 4183
Loss: 39.77462123157871
l2 norm of gradients: 0.29584680957395193
l2 norm of weights: 5.9933594827477075
-----
Iteration Number: 4184
Loss: 39.770239452638826
l2 norm of gradients: 0.29579077331908654
l2 norm of weights: 5.993263309577617
-----
Iteration Number: 4185
Loss: 39.765859319490104
l2 norm of gradients: 0.2957347689889591
l2 norm of weights: 5.993167147831105
-----
Iteration Number: 4186
Loss: 39.76148083091552
l2 norm of gradients: 0.29567879655661805
l2 norm of weights: 5.99307099750393
-----
Iteration Number: 4187
Loss: 39.75710398566589

```

```

12 norm of gradients: 0.2956228559951348
12 norm of weights: 5.992974858591849
-----
Iteration Number: 4188
Loss: 39.75272878253705
12 norm of gradients: 0.2955669472776034
12 norm of weights: 5.99287873109063
-----
Iteration Number: 4189
Loss: 39.74835522027297
12 norm of gradients: 0.295511070377141
12 norm of weights: 5.992782614996039
-----
Iteration Number: 4190
Loss: 39.743983297664016
12 norm of gradients: 0.2954552252668875
12 norm of weights: 5.992686510303846
-----
Iteration Number: 4191
Loss: 39.73961301348467
12 norm of gradients: 0.29539941192000546
12 norm of weights: 5.992590417009828
-----
Iteration Number: 4192
Loss: 39.73524436650358
12 norm of gradients: 0.2953436303096803
12 norm of weights: 5.992494335109765
-----
Iteration Number: 4193
Loss: 39.73087735551514
12 norm of gradients: 0.2952878804091204
12 norm of weights: 5.992398264599436
-----
Iteration Number: 4194
Loss: 39.72651197926486
12 norm of gradients: 0.2952321621915566
12 norm of weights: 5.9923022054746315
-----
Iteration Number: 4195
Loss: 39.722148236562276
12 norm of gradients: 0.29517647563024296
12 norm of weights: 5.992206157731138
-----
Iteration Number: 4196
Loss: 39.71778612616595
12 norm of gradients: 0.2951208206984559
12 norm of weights: 5.9921101213647505
-----

```

Iteration Number: 4197
Loss: 39.7134256468823
l2 norm of gradients: 0.29506519736949466
l2 norm of weights: 5.992014096371267

Iteration Number: 4198
Loss: 39.709066797476694
l2 norm of gradients: 0.2950096056166814
l2 norm of weights: 5.991918082746489

Iteration Number: 4199
Loss: 39.70470957673857
l2 norm of gradients: 0.29495404541336073
l2 norm of weights: 5.9918220804862194

Iteration Number: 4200
Loss: 39.70035398345197
l2 norm of gradients: 0.29489851673290024
l2 norm of weights: 5.99172608958627

Iteration Number: 4201
Loss: 39.6960000164149
l2 norm of gradients: 0.29484301954869
l2 norm of weights: 5.991630110042449

Iteration Number: 4202
Loss: 39.69164767441113
l2 norm of gradients: 0.294787553834143
l2 norm of weights: 5.9915341418505745

Iteration Number: 4203
Loss: 39.68729695621811
l2 norm of gradients: 0.29473211956269457
l2 norm of weights: 5.991438185006467

Iteration Number: 4204
Loss: 39.68294786063605
l2 norm of gradients: 0.2946767167078031
l2 norm of weights: 5.991342239505947

Iteration Number: 4205
Loss: 39.67860038645752
l2 norm of gradients: 0.29462134524294936
l2 norm of weights: 5.991246305344842

Iteration Number: 4206
Loss: 39.67425453248938
l2 norm of gradients: 0.2945660051416368

```

12 norm of weights: 5.991150382518986
-----
Iteration Number: 4207
Loss: 39.669910297493836
12 norm of gradients: 0.2945106963773915
12 norm of weights: 5.991054471024208
-----
Iteration Number: 4208
Loss: 39.66556768029357
12 norm of gradients: 0.29445541892376226
12 norm of weights: 5.990958570856349
-----
Iteration Number: 4209
Loss: 39.661226679684184
12 norm of gradients: 0.29440017275432045
12 norm of weights: 5.990862682011248
-----
Iteration Number: 4210
Loss: 39.656887294457675
12 norm of gradients: 0.2943449578426599
12 norm of weights: 5.990766804484753
-----
Iteration Number: 4211
Loss: 39.652549523421314
12 norm of gradients: 0.2942897741623972
12 norm of weights: 5.990670938272712
-----
Iteration Number: 4212
Loss: 39.648213365352774
12 norm of gradients: 0.2942346216871714
12 norm of weights: 5.9905750833709766
-----
Iteration Number: 4213
Loss: 39.64387881908733
12 norm of gradients: 0.2941795003906442
12 norm of weights: 5.990479239775402
-----
Iteration Number: 4214
Loss: 39.63954588340726
12 norm of gradients: 0.29412441024649966
12 norm of weights: 5.990383407481849
-----
Iteration Number: 4215
Loss: 39.63521455712184
12 norm of gradients: 0.29406935122844463
12 norm of weights: 5.99028758648618
-----
Iteration Number: 4216

```

```

Loss: 39.63088483904833
l2 norm of gradients: 0.2940143233102083
l2 norm of weights: 5.990191776784262
-----
Iteration Number: 4217
Loss: 39.6265567279724
l2 norm of gradients: 0.29395932646554246
l2 norm of weights: 5.990095978371967
-----
Iteration Number: 4218
Loss: 39.622230222721655
l2 norm of gradients: 0.29390436066822123
l2 norm of weights: 5.990000191245166
-----
Iteration Number: 4219
Loss: 39.61790532210023
l2 norm of gradients: 0.29384942589204144
l2 norm of weights: 5.989904415399739
-----
Iteration Number: 4220
Loss: 39.61358202491971
l2 norm of gradients: 0.2937945221108223
l2 norm of weights: 5.989808650831566
-----
Iteration Number: 4221
Loss: 39.60926032999402
l2 norm of gradients: 0.2937396492984055
l2 norm of weights: 5.989712897536533
-----
Iteration Number: 4222
Loss: 39.60494023613044
l2 norm of gradients: 0.29368480742865516
l2 norm of weights: 5.989617155510527
-----
Iteration Number: 4223
Loss: 39.600621742145506
l2 norm of gradients: 0.29362999647545773
l2 norm of weights: 5.989521424749441
-----
Iteration Number: 4224
Loss: 39.5963048468595
l2 norm of gradients: 0.29357521641272233
l2 norm of weights: 5.98942570524917
-----
Iteration Number: 4225
Loss: 39.59198954909282
l2 norm of gradients: 0.2935204672143802
l2 norm of weights: 5.989329997005613

```

```

-----
Iteration Number: 4226
Loss: 39.58767584764586
l2 norm of gradients: 0.2934657488543853
l2 norm of weights: 5.9892343000146715
-----
Iteration Number: 4227
Loss: 39.58336374136344
l2 norm of gradients: 0.2934110613067136
l2 norm of weights: 5.989138614272255
-----
Iteration Number: 4228
Loss: 39.57905322905406
l2 norm of gradients: 0.2933564045453639
l2 norm of weights: 5.98904293977427
-----
Iteration Number: 4229
Loss: 39.57474430954011
l2 norm of gradients: 0.293301778544357
l2 norm of weights: 5.988947276516632
-----
Iteration Number: 4230
Loss: 39.57043698164317
l2 norm of gradients: 0.29324718327773613
l2 norm of weights: 5.988851624495257
-----
Iteration Number: 4231
Loss: 39.566131244190665
l2 norm of gradients: 0.29319261871956687
l2 norm of weights: 5.988755983706065
-----
Iteration Number: 4232
Loss: 39.56182709600715
l2 norm of gradients: 0.2931380848439372
l2 norm of weights: 5.988660354144981
-----
Iteration Number: 4233
Loss: 39.55752453593414
l2 norm of gradients: 0.2930835816249574
l2 norm of weights: 5.988564735807933
-----
Iteration Number: 4234
Loss: 39.55322356276979
l2 norm of gradients: 0.29302910903675994
l2 norm of weights: 5.988469128690849
-----
Iteration Number: 4235
Loss: 39.54892417537576

```



```

12 norm of gradients: 0.29297466705349956
12 norm of weights: 5.988373532789668
-----
Iteration Number: 4236
Loss: 39.544626372563556
12 norm of gradients: 0.29292025564935364
12 norm of weights: 5.9882779481003245
-----
Iteration Number: 4237
Loss: 39.540330153165165
12 norm of gradients: 0.29286587479852116
12 norm of weights: 5.988182374618762
-----
Iteration Number: 4238
Loss: 39.53603551602132
12 norm of gradients: 0.29281152447522407
12 norm of weights: 5.988086812340924
-----
Iteration Number: 4239
Loss: 39.53174245996592
12 norm of gradients: 0.292757204653706
12 norm of weights: 5.987991261262761
-----
Iteration Number: 4240
Loss: 39.5274509838295
12 norm of gradients: 0.29270291530823306
12 norm of weights: 5.987895721380224
-----
Iteration Number: 4241
Loss: 39.52316108645724
12 norm of gradients: 0.2926486564130937
12 norm of weights: 5.98780019268927
-----
Iteration Number: 4242
Loss: 39.51887276667702
12 norm of gradients: 0.2925944279425981
12 norm of weights: 5.987704675185856
-----
Iteration Number: 4243
Loss: 39.51458602332434
12 norm of gradients: 0.2925402298710791
12 norm of weights: 5.987609168865946
-----
Iteration Number: 4244
Loss: 39.51030085527918
12 norm of gradients: 0.2924860621728915
12 norm of weights: 5.987513673725506
-----

```

Iteration Number: 4245
Loss: 39.5060172613297
l2 norm of gradients: 0.2924319248224123
l2 norm of weights: 5.987418189760506

Iteration Number: 4246
Loss: 39.5017352403531
l2 norm of gradients: 0.2923778177940406
l2 norm of weights: 5.9873227169669185

Iteration Number: 4247
Loss: 39.49745479117843
l2 norm of gradients: 0.29232374106219766
l2 norm of weights: 5.987227255340721

Iteration Number: 4248
Loss: 39.493175912647345
l2 norm of gradients: 0.2922696946013269
l2 norm of weights: 5.987131804877892

Iteration Number: 4249
Loss: 39.48889860362282
l2 norm of gradients: 0.2922156783858938
l2 norm of weights: 5.987036365574417

Iteration Number: 4250
Loss: 39.48462286293746
l2 norm of gradients: 0.2921616923903859
l2 norm of weights: 5.986940937426281

Iteration Number: 4251
Loss: 39.48034868945034
l2 norm of gradients: 0.29210773658931305
l2 norm of weights: 5.986845520429476

Iteration Number: 4252
Loss: 39.47607608202186
l2 norm of gradients: 0.2920538109572067
l2 norm of weights: 5.986750114579995

Iteration Number: 4253
Loss: 39.471805039469736
l2 norm of gradients: 0.29199991546862086
l2 norm of weights: 5.986654719873836

Iteration Number: 4254
Loss: 39.467535560676424
l2 norm of gradients: 0.2919460500981313

```

12 norm of weights: 5.986559336307
-----
Iteration Number: 4255
Loss: 39.46326764448977
12 norm of gradients: 0.29189221482033584
12 norm of weights: 5.98646396387549
-----
Iteration Number: 4256
Loss: 39.45900128974377
12 norm of gradients: 0.29183840960985435
12 norm of weights: 5.986368602575315
-----
Iteration Number: 4257
Loss: 39.45473649533507
12 norm of gradients: 0.2917846344413287
12 norm of weights: 5.986273252402485
-----
Iteration Number: 4258
Loss: 39.45047326007713
12 norm of gradients: 0.2917308892894228
12 norm of weights: 5.986177913353016
-----
Iteration Number: 4259
Loss: 39.44621158285916
12 norm of gradients: 0.2916771741288225
12 norm of weights: 5.986082585422925
-----
Iteration Number: 4260
Loss: 39.44195146252869
12 norm of gradients: 0.29162348893423545
12 norm of weights: 5.985987268608234
-----
Iteration Number: 4261
Loss: 39.43769289795153
12 norm of gradients: 0.29156983368039163
12 norm of weights: 5.985891962904967
-----
Iteration Number: 4262
Loss: 39.433435887985716
12 norm of gradients: 0.2915162083420426
12 norm of weights: 5.9857966683091535
-----
Iteration Number: 4263
Loss: 39.429180431497564
12 norm of gradients: 0.29146261289396197
12 norm of weights: 5.985701384816824
-----
Iteration Number: 4264

```

```

Loss: 39.42492652734127
l2 norm of gradients: 0.2914090473109452
l2 norm of weights: 5.985606112424015
-----
Iteration Number: 4265
Loss: 39.420674174392914
l2 norm of gradients: 0.2913555115678097
l2 norm of weights: 5.985510851126762
-----
Iteration Number: 4266
Loss: 39.416423371521155
l2 norm of gradients: 0.29130200563939485
l2 norm of weights: 5.98541560092111
-----
Iteration Number: 4267
Loss: 39.41217411759276
l2 norm of gradients: 0.2912485295005616
l2 norm of weights: 5.985320361803103
-----
Iteration Number: 4268
Loss: 39.40792641146887
l2 norm of gradients: 0.2911950831261933
l2 norm of weights: 5.9852251337687905
-----
Iteration Number: 4269
Loss: 39.4036802520274
l2 norm of gradients: 0.2911416664911945
l2 norm of weights: 5.985129916814223
-----
Iteration Number: 4270
Loss: 39.39943563813688
l2 norm of gradients: 0.29108827957049194
l2 norm of weights: 5.985034710935458
-----
Iteration Number: 4271
Loss: 39.39519256868119
l2 norm of gradients: 0.29103492233903416
l2 norm of weights: 5.984939516128552
-----
Iteration Number: 4272
Loss: 39.39095104250608
l2 norm of gradients: 0.2909815947717914
l2 norm of weights: 5.98484433238957
-----
Iteration Number: 4273
Loss: 39.386711058506535
l2 norm of gradients: 0.2909282968437559
l2 norm of weights: 5.984749159714575

```

```

-----
Iteration Number: 4274
Loss: 39.3824726155607
l2 norm of gradients: 0.29087502852994124
l2 norm of weights: 5.984653998099637
-----
Iteration Number: 4275
Loss: 39.37823571253381
l2 norm of gradients: 0.29082178980538326
l2 norm of weights: 5.9845588475408285
-----
Iteration Number: 4276
Loss: 39.37400034832458
l2 norm of gradients: 0.2907685806451391
l2 norm of weights: 5.984463708034224
-----
Iteration Number: 4277
Loss: 39.36976652179687
l2 norm of gradients: 0.29071540102428806
l2 norm of weights: 5.984368579575904
-----
Iteration Number: 4278
Loss: 39.36553423182798
l2 norm of gradients: 0.2906622509179309
l2 norm of weights: 5.9842734621619496
-----
Iteration Number: 4279
Loss: 39.361303477307224
l2 norm of gradients: 0.29060913030119
l2 norm of weights: 5.9841783557884485
-----
Iteration Number: 4280
Loss: 39.35707425711761
l2 norm of gradients: 0.29055603914920963
l2 norm of weights: 5.984083260451487
-----
Iteration Number: 4281
Loss: 39.352846570143
l2 norm of gradients: 0.29050297743715575
l2 norm of weights: 5.983988176147159
-----
Iteration Number: 4282
Loss: 39.34862041527289
l2 norm of gradients: 0.2904499451402157
l2 norm of weights: 5.98389310287156
-----
Iteration Number: 4283
Loss: 39.34439579138006

```

```

12 norm of gradients: 0.29039694223359885
12 norm of weights: 5.98379804062079
-----
Iteration Number: 4284
Loss: 39.340172697369255
12 norm of gradients: 0.290343968692536
12 norm of weights: 5.983702989390951
-----
Iteration Number: 4285
Loss: 39.335951132120954
12 norm of gradients: 0.29029102449227956
12 norm of weights: 5.983607949178147
-----
Iteration Number: 4286
Loss: 39.33173109452099
12 norm of gradients: 0.29023810960810353
12 norm of weights: 5.983512919978488
-----
Iteration Number: 4287
Loss: 39.32751258347011
12 norm of gradients: 0.2901852240153036
12 norm of weights: 5.983417901788087
-----
Iteration Number: 4288
Loss: 39.323295597854134
12 norm of gradients: 0.29013236768919715
12 norm of weights: 5.983322894603059
-----
Iteration Number: 4289
Loss: 39.319080136569184
12 norm of gradients: 0.29007954060512275
12 norm of weights: 5.983227898419524
-----
Iteration Number: 4290
Loss: 39.314866198512796
12 norm of gradients: 0.2900267427384408
12 norm of weights: 5.983132913233602
-----
Iteration Number: 4291
Loss: 39.31065378257272
12 norm of gradients: 0.2899739740645333
12 norm of weights: 5.983037939041421
-----
Iteration Number: 4292
Loss: 39.30644288765192
12 norm of gradients: 0.28992123455880353
12 norm of weights: 5.982942975839109
-----

```

Iteration Number: 4293
Loss: 39.302233512641415
l2 norm of gradients: 0.2898685241966765
l2 norm of weights: 5.982848023622797

Iteration Number: 4294
Loss: 39.298025656453746
l2 norm of gradients: 0.28981584295359863
l2 norm of weights: 5.982753082388622

Iteration Number: 4295
Loss: 39.29381931797436
l2 norm of gradients: 0.2897631908050378
l2 norm of weights: 5.982658152132722

Iteration Number: 4296
Loss: 39.2896144961227
l2 norm of gradients: 0.28971056772648335
l2 norm of weights: 5.9825632328512395

Iteration Number: 4297
Loss: 39.28541118977467
l2 norm of gradients: 0.2896579736934462
l2 norm of weights: 5.982468324540318

Iteration Number: 4298
Loss: 39.281209397861026
l2 norm of gradients: 0.28960540868145845
l2 norm of weights: 5.982373427196109

Iteration Number: 4299
Loss: 39.27700911926698
l2 norm of gradients: 0.289552872666074
l2 norm of weights: 5.982278540814761

Iteration Number: 4300
Loss: 39.27281035291101
l2 norm of gradients: 0.28950036562286774
l2 norm of weights: 5.982183665392431

Iteration Number: 4301
Loss: 39.26861309769621
l2 norm of gradients: 0.28944788752743633
l2 norm of weights: 5.982088800925277

Iteration Number: 4302
Loss: 39.26441735253346
l2 norm of gradients: 0.28939543835539766

```
12 norm of weights: 5.981993947409461
-----
Iteration Number: 4303
Loss: 39.26022311631851
12 norm of gradients: 0.2893430180823909
12 norm of weights: 5.981899104841148
-----
Iteration Number: 4304
Loss: 39.25603038798537
12 norm of gradients: 0.2892906266840767
12 norm of weights: 5.981804273216505
-----
Iteration Number: 4305
Loss: 39.251839166427935
12 norm of gradients: 0.28923826413613707
12 norm of weights: 5.981709452531703
-----
Iteration Number: 4306
Loss: 39.24764945056683
12 norm of gradients: 0.28918593041427526
12 norm of weights: 5.981614642782918
-----
Iteration Number: 4307
Loss: 39.24346123930412
12 norm of gradients: 0.28913362549421584
12 norm of weights: 5.9815198439663275
-----
Iteration Number: 4308
Loss: 39.23927453156141
12 norm of gradients: 0.28908134935170465
12 norm of weights: 5.981425056078112
-----
Iteration Number: 4309
Loss: 39.235089326265474
12 norm of gradients: 0.28902910196250914
12 norm of weights: 5.981330279114455
-----
Iteration Number: 4310
Loss: 39.23090562232756
12 norm of gradients: 0.28897688330241755
12 norm of weights: 5.981235513071547
-----
Iteration Number: 4311
Loss: 39.226723418650124
12 norm of gradients: 0.2889246933472396
12 norm of weights: 5.981140757945575
-----
Iteration Number: 4312
```


Loss: 39.22254271417681
l2 norm of gradients: 0.2888725320728063
l2 norm of weights: 5.981046013732736

Iteration Number: 4313
Loss: 39.218363507817635
l2 norm of gradients: 0.28882039945496973
l2 norm of weights: 5.980951280429224

Iteration Number: 4314
Loss: 39.21418579849619
l2 norm of gradients: 0.2887682954696034
l2 norm of weights: 5.980856558031241

Iteration Number: 4315
Loss: 39.21000958511916
l2 norm of gradients: 0.288716220092602
l2 norm of weights: 5.980761846534992

Iteration Number: 4316
Loss: 39.20583486663326
l2 norm of gradients: 0.2886641732998811
l2 norm of weights: 5.980667145936682

Iteration Number: 4317
Loss: 39.20166164196079
l2 norm of gradients: 0.2886121550673778
l2 norm of weights: 5.98057245623252

Iteration Number: 4318
Loss: 39.19748991001525
l2 norm of gradients: 0.2885601653710502
l2 norm of weights: 5.980477777418721

Iteration Number: 4319
Loss: 39.19331966972088
l2 norm of gradients: 0.2885082041868775
l2 norm of weights: 5.980383109491501

Iteration Number: 4320
Loss: 39.18915092003302
l2 norm of gradients: 0.2884562714908601
l2 norm of weights: 5.980288452447079

Iteration Number: 4321
Loss: 39.18498365985564
l2 norm of gradients: 0.2884043672590195
l2 norm of weights: 5.980193806281678

```
-----  
Iteration Number: 4322  
Loss: 39.18081788811952  
l2 norm of gradients: 0.28835249146739833  
l2 norm of weights: 5.980099170991523  
-----  
Iteration Number: 4323  
Loss: 39.17665360377598  
l2 norm of gradients: 0.2883006440920603  
l2 norm of weights: 5.980004546572845  
-----  
Iteration Number: 4324  
Loss: 39.172490805734554  
l2 norm of gradients: 0.2882488251090901  
l2 norm of weights: 5.979909933021874  
-----  
Iteration Number: 4325  
Loss: 39.168329492949105  
l2 norm of gradients: 0.2881970344945936  
l2 norm of weights: 5.979815330334847  
-----  
Iteration Number: 4326  
Loss: 39.16416966433796  
l2 norm of gradients: 0.28814527222469766  
l2 norm of weights: 5.979720738508002  
-----  
Iteration Number: 4327  
Loss: 39.16001131884972  
l2 norm of gradients: 0.28809353827555007  
l2 norm of weights: 5.979626157537581  
-----  
Iteration Number: 4328  
Loss: 39.155854455411486  
l2 norm of gradients: 0.28804183262331984  
l2 norm of weights: 5.979531587419827  
-----  
Iteration Number: 4329  
Loss: 39.151699072963815  
l2 norm of gradients: 0.28799015524419685  
l2 norm of weights: 5.979437028150991  
-----  
Iteration Number: 4330  
Loss: 39.14754517046263  
l2 norm of gradients: 0.2879385061143919  
l2 norm of weights: 5.979342479727322  
-----  
Iteration Number: 4331  
Loss: 39.14339274682554
```

```

12 norm of gradients: 0.28788688521013694
12 norm of weights: 5.979247942145076
-----
Iteration Number: 4332
Loss: 39.139241801000466
12 norm of gradients: 0.28783529250768475
12 norm of weights: 5.979153415400508
-----
Iteration Number: 4333
Loss: 39.135092331926415
12 norm of gradients: 0.28778372798330903
12 norm of weights: 5.979058899489881
-----
Iteration Number: 4334
Loss: 39.13094433855933
12 norm of gradients: 0.2877321916133045
12 norm of weights: 5.9789643944094575
-----
Iteration Number: 4335
Loss: 39.126797819836746
12 norm of gradients: 0.2876806833739867
12 norm of weights: 5.978869900155504
-----
Iteration Number: 4336
Loss: 39.12265277470718
12 norm of gradients: 0.2876292032416922
12 norm of weights: 5.978775416724292
-----
Iteration Number: 4337
Loss: 39.118509202113906
12 norm of gradients: 0.2875777511927784
12 norm of weights: 5.978680944112093
-----
Iteration Number: 4338
Loss: 39.114367101012135
12 norm of gradients: 0.2875263272036234
12 norm of weights: 5.978586482315184
-----
Iteration Number: 4339
Loss: 39.11022647033143
12 norm of gradients: 0.28747493125062645
12 norm of weights: 5.9784920313298455
-----
Iteration Number: 4340
Loss: 39.106087309042756
12 norm of gradients: 0.2874235633102073
12 norm of weights: 5.978397591152357
-----

```

Iteration Number: 4341
Loss: 39.10194961609161
l2 norm of gradients: 0.28737222335880686
l2 norm of weights: 5.978303161779007

Iteration Number: 4342
Loss: 39.097813390429586
l2 norm of gradients: 0.2873209113728867
l2 norm of weights: 5.978208743206081

Iteration Number: 4343
Loss: 39.09367863101018
l2 norm of gradients: 0.2872696273289291
l2 norm of weights: 5.978114335429874

Iteration Number: 4344
Loss: 39.0895453367791
l2 norm of gradients: 0.2872183712034373
l2 norm of weights: 5.978019938446679

Iteration Number: 4345
Loss: 39.08541350670992
l2 norm of gradients: 0.2871671429729351
l2 norm of weights: 5.9779255522527945

Iteration Number: 4346
Loss: 39.081283139740876
l2 norm of gradients: 0.2871159426139673
l2 norm of weights: 5.977831176844521

Iteration Number: 4347
Loss: 39.077154234836144
l2 norm of gradients: 0.2870647701030992
l2 norm of weights: 5.977736812218164

Iteration Number: 4348
Loss: 39.07302679096314
l2 norm of gradients: 0.28701362541691705
l2 norm of weights: 5.977642458370029

Iteration Number: 4349
Loss: 39.0689008070777
l2 norm of gradients: 0.2869625085320276
l2 norm of weights: 5.977548115296427

Iteration Number: 4350
Loss: 39.06477628212281
l2 norm of gradients: 0.2869114194250584

```

12 norm of weights: 5.9774537829936705
-----
Iteration Number: 4351
Loss: 39.0606532150816
12 norm of gradients: 0.2868603580726577
12 norm of weights: 5.977359461458078
-----
Iteration Number: 4352
Loss: 39.05653160491522
12 norm of gradients: 0.2868093244514944
12 norm of weights: 5.977265150685967
-----
Iteration Number: 4353
Loss: 39.05241145057449
12 norm of gradients: 0.28675831853825795
12 norm of weights: 5.9771708506736605
-----
Iteration Number: 4354
Loss: 39.048292751043014
12 norm of gradients: 0.28670734030965855
12 norm of weights: 5.977076561417485
-----
Iteration Number: 4355
Loss: 39.04417550526603
12 norm of gradients: 0.286656389742427
12 norm of weights: 5.976982282913768
-----
Iteration Number: 4356
Loss: 39.04005971223065
12 norm of gradients: 0.28660546681331467
12 norm of weights: 5.976888015158843
-----
Iteration Number: 4357
Loss: 39.03594537088175
12 norm of gradients: 0.2865545714990936
12 norm of weights: 5.976793758149042
-----
Iteration Number: 4358
Loss: 39.031832480223464
12 norm of gradients: 0.2865037037765564
12 norm of weights: 5.976699511880704
-----
Iteration Number: 4359
Loss: 39.02772103919531
12 norm of gradients: 0.286452863622516
12 norm of weights: 5.9766052763501705
-----
Iteration Number: 4360

```

```
Loss: 39.02361104677152
12 norm of gradients: 0.28640205101380634
12 norm of weights: 5.9765110515537865
-----
Iteration Number: 4361
Loss: 39.019502501940096
12 norm of gradients: 0.2863512659272814
12 norm of weights: 5.976416837487895
-----
Iteration Number: 4362
Loss: 39.0153954036676
12 norm of gradients: 0.286300508339816
12 norm of weights: 5.976322634148849
-----
Iteration Number: 4363
Loss: 39.01128975092077
12 norm of gradients: 0.2862497782283054
12 norm of weights: 5.976228441533002
-----
Iteration Number: 4364
Loss: 39.00718554268656
12 norm of gradients: 0.2861990755696653
12 norm of weights: 5.976134259636708
-----
Iteration Number: 4365
Loss: 39.00308277793341
12 norm of gradients: 0.2861484003408319
12 norm of weights: 5.976040088456327
-----
Iteration Number: 4366
Loss: 38.99898145564474
12 norm of gradients: 0.2860977525187618
12 norm of weights: 5.975945927988221
-----
Iteration Number: 4367
Loss: 38.99488157479693
12 norm of gradients: 0.2860471320804322
12 norm of weights: 5.975851778228754
-----
Iteration Number: 4368
Loss: 38.99078313437057
12 norm of gradients: 0.2859965390028406
12 norm of weights: 5.975757639174296
-----
Iteration Number: 4369
Loss: 38.98668613334502
12 norm of gradients: 0.2859459732630049
12 norm of weights: 5.975663510821216
```

```

-----
Iteration Number: 4370
Loss: 38.982590570707366
l2 norm of gradients: 0.28589543483796354
l2 norm of weights: 5.97556939316589
-----
Iteration Number: 4371
Loss: 38.978496445428064
l2 norm of gradients: 0.2858449237047752
l2 norm of weights: 5.975475286204694
-----
Iteration Number: 4372
Loss: 38.974403756504785
l2 norm of gradients: 0.2857944398405191
l2 norm of weights: 5.975381189934008
-----
Iteration Number: 4373
Loss: 38.97031250292427
l2 norm of gradients: 0.28574398322229455
l2 norm of weights: 5.975287104350215
-----
Iteration Number: 4374
Loss: 38.96622268365353
l2 norm of gradients: 0.28569355382722145
l2 norm of weights: 5.975193029449701
-----
Iteration Number: 4375
Loss: 38.96213429769009
l2 norm of gradients: 0.2856431516324398
l2 norm of weights: 5.975098965228856
-----
Iteration Number: 4376
Loss: 38.9580473440309
l2 norm of gradients: 0.2855927766151103
l2 norm of weights: 5.975004911684071
-----
Iteration Number: 4377
Loss: 38.953961821650466
l2 norm of gradients: 0.28554242875241354
l2 norm of weights: 5.974910868811741
-----
Iteration Number: 4378
Loss: 38.949877729542685
l2 norm of gradients: 0.2854921080215505
l2 norm of weights: 5.974816836608264
-----
Iteration Number: 4379
Loss: 38.94579506671336

```

```

12 norm of gradients: 0.2854418143997427
12 norm of weights: 5.9747228150700415
-----
Iteration Number: 4380
Loss: 38.94171383214376
12 norm of gradients: 0.28539154786423143
12 norm of weights: 5.974628804193476
-----
Iteration Number: 4381
Loss: 38.93763402481685
12 norm of gradients: 0.2853413083922788
12 norm of weights: 5.974534803974976
-----
Iteration Number: 4382
Loss: 38.933555643748036
12 norm of gradients: 0.2852910959611664
12 norm of weights: 5.9744408144109515
-----
Iteration Number: 4383
Loss: 38.92947868791326
12 norm of gradients: 0.2852409105481969
12 norm of weights: 5.974346835497814
-----
Iteration Number: 4384
Loss: 38.925403156316996
12 norm of gradients: 0.28519075213069245
12 norm of weights: 5.974252867231979
-----
Iteration Number: 4385
Loss: 38.921329047959205
12 norm of gradients: 0.28514062068599577
12 norm of weights: 5.974158909609865
-----
Iteration Number: 4386
Loss: 38.91725636183274
12 norm of gradients: 0.28509051619146963
12 norm of weights: 5.974064962627895
-----
Iteration Number: 4387
Loss: 38.91318509694728
12 norm of gradients: 0.2850404386244969
12 norm of weights: 5.973971026282494
-----
Iteration Number: 4388
Loss: 38.909115252277125
12 norm of gradients: 0.28499038796248066
12 norm of weights: 5.973877100570087
-----

```


Iteration Number: 4389
Loss: 38.905046826856164
l2 norm of gradients: 0.2849403641828442
l2 norm of weights: 5.973783185487107

Iteration Number: 4390
Loss: 38.90097981967385
l2 norm of gradients: 0.28489036726303063
l2 norm of weights: 5.973689281029984

Iteration Number: 4391
Loss: 38.89691422973446
l2 norm of gradients: 0.28484039718050347
l2 norm of weights: 5.973595387195158

Iteration Number: 4392
Loss: 38.89285005604134
l2 norm of gradients: 0.28479045391274604
l2 norm of weights: 5.973501503979066

Iteration Number: 4393
Loss: 38.88878729759363
l2 norm of gradients: 0.284740537437262
l2 norm of weights: 5.9734076313781515

Iteration Number: 4394
Loss: 38.88472595340546
l2 norm of gradients: 0.2846906477315748
l2 norm of weights: 5.973313769388858

Iteration Number: 4395
Loss: 38.880666022475964
l2 norm of gradients: 0.28464078477322813
l2 norm of weights: 5.973219918007635

Iteration Number: 4396
Loss: 38.876607503832794
l2 norm of gradients: 0.2845909485397856
l2 norm of weights: 5.973126077230932

Iteration Number: 4397
Loss: 38.8725503964668
l2 norm of gradients: 0.2845411390088309
l2 norm of weights: 5.973032247055202

Iteration Number: 4398
Loss: 38.8684946993888
l2 norm of gradients: 0.28449135615796733

```

12 norm of weights: 5.9729384274769055
-----
Iteration Number: 4399
Loss: 38.86444041161904
12 norm of gradients: 0.2844415999648188
12 norm of weights: 5.9728446184924975
-----
Iteration Number: 4400
Loss: 38.86038753217419
12 norm of gradients: 0.2843918704070288
12 norm of weights: 5.972750820098444
-----
Iteration Number: 4401
Loss: 38.85633606005405
12 norm of gradients: 0.2843421674622607
12 norm of weights: 5.972657032291209
-----
Iteration Number: 4402
Loss: 38.852285994278574
12 norm of gradients: 0.284292491108198
12 norm of weights: 5.97256325506726
-----
Iteration Number: 4403
Loss: 38.8482373338702
12 norm of gradients: 0.284242841322544
12 norm of weights: 5.9724694884230685
-----
Iteration Number: 4404
Loss: 38.844190077834156
12 norm of gradients: 0.28419321808302195
12 norm of weights: 5.972375732355109
-----
Iteration Number: 4405
Loss: 38.84014422519652
12 norm of gradients: 0.284143621367375
12 norm of weights: 5.972281986859859
-----
Iteration Number: 4406
Loss: 38.836099774976894
12 norm of gradients: 0.2840940511533661
12 norm of weights: 5.9721882519337965
-----
Iteration Number: 4407
Loss: 38.83205672618984
12 norm of gradients: 0.28404450741877807
12 norm of weights: 5.972094527573406
-----
Iteration Number: 4408

```

Loss: 38.82801507784591
l2 norm of gradients: 0.28399499014141366
l2 norm of weights: 5.972000813775171

Iteration Number: 4409
Loss: 38.82397482898889
l2 norm of gradients: 0.28394549929909535
l2 norm of weights: 5.971907110535582

Iteration Number: 4410
Loss: 38.81993597862654
l2 norm of gradients: 0.2838960348696654
l2 norm of weights: 5.97181341785113

Iteration Number: 4411
Loss: 38.81589852578515
l2 norm of gradients: 0.28384659683098595
l2 norm of weights: 5.971719735718307

Iteration Number: 4412
Loss: 38.8118624694912
l2 norm of gradients: 0.283797185160939
l2 norm of weights: 5.971626064133613

Iteration Number: 4413
Loss: 38.807827808768764
l2 norm of gradients: 0.28374779983742604
l2 norm of weights: 5.971532403093544

Iteration Number: 4414
Loss: 38.8037945426498
l2 norm of gradients: 0.2836984408383687
l2 norm of weights: 5.971438752594606

Iteration Number: 4415
Loss: 38.79976267014238
l2 norm of gradients: 0.28364910814170796
l2 norm of weights: 5.971345112633305

Iteration Number: 4416
Loss: 38.79573219030569
l2 norm of gradients: 0.2835998017254047
l2 norm of weights: 5.971251483206146

Iteration Number: 4417
Loss: 38.791703102137525
l2 norm of gradients: 0.28355052156743965
l2 norm of weights: 5.971157864309642

```

-----
Iteration Number: 4418
Loss: 38.787675404684606
l2 norm of gradients: 0.28350126764581307
l2 norm of weights: 5.971064255940307
-----
Iteration Number: 4419
Loss: 38.78364909699283
l2 norm of gradients: 0.28345203993854473
l2 norm of weights: 5.970970658094658
-----
Iteration Number: 4420
Loss: 38.779624178066605
l2 norm of gradients: 0.2834028384236745
l2 norm of weights: 5.970877070769216
-----
Iteration Number: 4421
Loss: 38.775600646944866
l2 norm of gradients: 0.2833536630792614
l2 norm of weights: 5.9707834939605
-----
Iteration Number: 4422
Loss: 38.77157850267686
l2 norm of gradients: 0.28330451388338446
l2 norm of weights: 5.970689927665038
-----
Iteration Number: 4423
Loss: 38.767557744285384
l2 norm of gradients: 0.2832553908141423
l2 norm of weights: 5.970596371879358
-----
Iteration Number: 4424
Loss: 38.763538370815176
l2 norm of gradients: 0.2832062938496529
l2 norm of weights: 5.970502826599989
-----
Iteration Number: 4425
Loss: 38.759520381299076
l2 norm of gradients: 0.283157222968054
l2 norm of weights: 5.970409291823468
-----
Iteration Number: 4426
Loss: 38.755503774774766
l2 norm of gradients: 0.28310817814750294
l2 norm of weights: 5.97031576754633
-----
Iteration Number: 4427
Loss: 38.75148855027411

```

```

12 norm of gradients: 0.28305915936617654
12 norm of weights: 5.9702222537651135
-----
Iteration Number: 4428
Loss: 38.74747470685132
12 norm of gradients: 0.2830101666022712
12 norm of weights: 5.970128750476361
-----
Iteration Number: 4429
Loss: 38.74346224354091
12 norm of gradients: 0.28296119983400286
12 norm of weights: 5.970035257676619
-----
Iteration Number: 4430
Loss: 38.73945115939326
12 norm of gradients: 0.2829122590396069
12 norm of weights: 5.969941775362433
-----
Iteration Number: 4431
Loss: 38.73544145343141
12 norm of gradients: 0.28286334419733833
12 norm of weights: 5.969848303530356
-----
Iteration Number: 4432
Loss: 38.73143312472093
12 norm of gradients: 0.28281445528547156
12 norm of weights: 5.96975484217694
-----
Iteration Number: 4433
Loss: 38.72742617229013
12 norm of gradients: 0.2827655922823005
12 norm of weights: 5.969661391298741
-----
Iteration Number: 4434
Loss: 38.72342059519454
12 norm of gradients: 0.2827167551661385
12 norm of weights: 5.969567950892319
-----
Iteration Number: 4435
Loss: 38.7194163924806
12 norm of gradients: 0.28266794391531835
12 norm of weights: 5.969474520954235
-----
Iteration Number: 4436
Loss: 38.71541356319388
12 norm of gradients: 0.28261915850819236
12 norm of weights: 5.969381101481053
-----

```

Iteration Number: 4437
Loss: 38.71141210639055
l2 norm of gradients: 0.28257039892313207
l2 norm of weights: 5.969287692469341

Iteration Number: 4438
Loss: 38.70741202111072
l2 norm of gradients: 0.2825216651385285
l2 norm of weights: 5.9691942939156695

Iteration Number: 4439
Loss: 38.703413306405025
l2 norm of gradients: 0.2824729571327922
l2 norm of weights: 5.96910090581661

Iteration Number: 4440
Loss: 38.69941596133054
l2 norm of gradients: 0.2824242748843528
l2 norm of weights: 5.969007528168739

Iteration Number: 4441
Loss: 38.695419984938134
l2 norm of gradients: 0.28237561837165953
l2 norm of weights: 5.968914160968636

Iteration Number: 4442
Loss: 38.69142537628182
l2 norm of gradients: 0.28232698757318087
l2 norm of weights: 5.9688208042128785

Iteration Number: 4443
Loss: 38.68743213441701
l2 norm of gradients: 0.2822783824674046
l2 norm of weights: 5.968727457898054

Iteration Number: 4444
Loss: 38.68344025839253
l2 norm of gradients: 0.28222980303283784
l2 norm of weights: 5.968634122020747

Iteration Number: 4445
Loss: 38.679449747274774
l2 norm of gradients: 0.282181249248007
l2 norm of weights: 5.968540796577548

Iteration Number: 4446
Loss: 38.67546060010836
l2 norm of gradients: 0.2821327210914578

```
l2 norm of weights: 5.968447481565048
-----
Iteration Number: 4447
Loss: 38.67147281597441
l2 norm of gradients: 0.28208421854175514
l2 norm of weights: 5.968354176979843
-----
Iteration Number: 4448
Loss: 38.66748639391075
l2 norm of gradients: 0.2820357415774833
l2 norm of weights: 5.968260882818528
-----
Iteration Number: 4449
Loss: 38.66350133298183
l2 norm of gradients: 0.28198729017724566
l2 norm of weights: 5.968167599077708
-----
Iteration Number: 4450
Loss: 38.659517632255
l2 norm of gradients: 0.28193886431966486
l2 norm of weights: 5.968074325753982
-----
Iteration Number: 4451
Loss: 38.655535290795136
l2 norm of gradients: 0.2818904639833828
l2 norm of weights: 5.967981062843957
-----
Iteration Number: 4452
Loss: 38.65155430765074
l2 norm of gradients: 0.2818420891470605
l2 norm of weights: 5.967887810344242
-----
Iteration Number: 4453
Loss: 38.64757468190038
l2 norm of gradients: 0.28179373978937833
l2 norm of weights: 5.967794568251446
-----
Iteration Number: 4454
Loss: 38.643596412606435
l2 norm of gradients: 0.2817454158890355
l2 norm of weights: 5.967701336562187
-----
Iteration Number: 4455
Loss: 38.639619498828786
l2 norm of gradients: 0.2816971174247507
l2 norm of weights: 5.9676081152730776
-----
Iteration Number: 4456
```

Loss: 38.63564393963313
l2 norm of gradients: 0.28164884437526155
l2 norm of weights: 5.967514904380739

Iteration Number: 4457
Loss: 38.63166973410435
l2 norm of gradients: 0.28160059671932475
l2 norm of weights: 5.967421703881794

Iteration Number: 4458
Loss: 38.627696881288415
l2 norm of gradients: 0.28155237443571635
l2 norm of weights: 5.967328513772865

Iteration Number: 4459
Loss: 38.62372538026868
l2 norm of gradients: 0.2815041775032313
l2 norm of weights: 5.9672353340505815

Iteration Number: 4460
Loss: 38.619755230112524
l2 norm of gradients: 0.28145600590068376
l2 norm of weights: 5.967142164711572

Iteration Number: 4461
Loss: 38.61578642988932
l2 norm of gradients: 0.28140785960690656
l2 norm of weights: 5.9670490057524725

Iteration Number: 4462
Loss: 38.61181897867609
l2 norm of gradients: 0.2813597386007521
l2 norm of weights: 5.966955857169915

Iteration Number: 4463
Loss: 38.60785287555053
l2 norm of gradients: 0.28131164286109145
l2 norm of weights: 5.966862718960539

Iteration Number: 4464
Loss: 38.603888119578826
l2 norm of gradients: 0.28126357236681493
l2 norm of weights: 5.966769591120986

Iteration Number: 4465
Loss: 38.59992470983163
l2 norm of gradients: 0.28121552709683173
l2 norm of weights: 5.966676473647898


```

-----
Iteration Number: 4466
Loss: 38.59596264539185
l2 norm of gradients: 0.28116750703007
l2 norm of weights: 5.9665833665379235
-----
Iteration Number: 4467
Loss: 38.59200192534213
l2 norm of gradients: 0.281119512145477
l2 norm of weights: 5.966490269787711
-----
Iteration Number: 4468
Loss: 38.58804254875199
l2 norm of gradients: 0.2810715424220187
l2 norm of weights: 5.9663971833939105
-----
Iteration Number: 4469
Loss: 38.584084514711115
l2 norm of gradients: 0.2810235978386803
l2 norm of weights: 5.966304107353177
-----
Iteration Number: 4470
Loss: 38.58012782228157
l2 norm of gradients: 0.28097567837446574
l2 norm of weights: 5.966211041662167
-----
Iteration Number: 4471
Loss: 38.57617247056002
l2 norm of gradients: 0.2809277840083979
l2 norm of weights: 5.966117986317543
-----
Iteration Number: 4472
Loss: 38.57221845862811
l2 norm of gradients: 0.2808799147195187
l2 norm of weights: 5.966024941315964
-----
Iteration Number: 4473
Loss: 38.56826578554945
l2 norm of gradients: 0.2808320704868887
l2 norm of weights: 5.965931906654096
-----
Iteration Number: 4474
Loss: 38.5643144504362
l2 norm of gradients: 0.28078425128958745
l2 norm of weights: 5.965838882328606
-----
Iteration Number: 4475
Loss: 38.56036445235165

```

```

12 norm of gradients: 0.2807364571067135
12 norm of weights: 5.965745868336167
-----
Iteration Number: 4476
Loss: 38.556415790388876
12 norm of gradients: 0.280688687917384
12 norm of weights: 5.965652864673448
-----
Iteration Number: 4477
Loss: 38.55246846363819
12 norm of gradients: 0.280640943700735
12 norm of weights: 5.965559871337128
-----
Iteration Number: 4478
Loss: 38.54852247117626
12 norm of gradients: 0.2805932244359215
12 norm of weights: 5.965466888323882
-----
Iteration Number: 4479
Loss: 38.54457781210168
12 norm of gradients: 0.2805455301021171
12 norm of weights: 5.965373915630393
-----
Iteration Number: 4480
Loss: 38.54063448550137
12 norm of gradients: 0.2804978606785143
12 norm of weights: 5.965280953253345
-----
Iteration Number: 4481
Loss: 38.53669249045229
12 norm of gradients: 0.2804502161443243
12 norm of weights: 5.965188001189421
-----
Iteration Number: 4482
Loss: 38.532751826063574
12 norm of gradients: 0.2804025964787771
12 norm of weights: 5.965095059435313
-----
Iteration Number: 4483
Loss: 38.52881249142139
12 norm of gradients: 0.2803550016611213
12 norm of weights: 5.965002127987711
-----
Iteration Number: 4484
Loss: 38.524874485614774
12 norm of gradients: 0.28030743167062455
12 norm of weights: 5.964909206843309
-----

```

Iteration Number: 4485
Loss: 38.52093780774676
l2 norm of gradients: 0.2802598864865729
l2 norm of weights: 5.964816295998803

Iteration Number: 4486
Loss: 38.517002456904656
l2 norm of gradients: 0.2802123660882713
l2 norm of weights: 5.964723395450894

Iteration Number: 4487
Loss: 38.51306843217169
l2 norm of gradients: 0.280164870455043
l2 norm of weights: 5.964630505196282

Iteration Number: 4488
Loss: 38.509135732666905
l2 norm of gradients: 0.2801173995662305
l2 norm of weights: 5.964537625231673

Iteration Number: 4489
Loss: 38.5052043574779
l2 norm of gradients: 0.2800699534011944
l2 norm of weights: 5.964444755553773

Iteration Number: 4490
Loss: 38.501274305696654
l2 norm of gradients: 0.2800225319393143
l2 norm of weights: 5.9643518961592905

Iteration Number: 4491
Loss: 38.497345576438164
l2 norm of gradients: 0.27997513515998834
l2 norm of weights: 5.964259047044941

Iteration Number: 4492
Loss: 38.493418168785574
l2 norm of gradients: 0.2799277630426331
l2 norm of weights: 5.964166208207435

Iteration Number: 4493
Loss: 38.48949208185728
l2 norm of gradients: 0.27988041556668397
l2 norm of weights: 5.964073379643494

Iteration Number: 4494
Loss: 38.48556731473836
l2 norm of gradients: 0.27983309271159473

```

12 norm of weights: 5.963980561349837
-----
Iteration Number: 4495
Loss: 38.4816438665319
12 norm of gradients: 0.27978579445683793
12 norm of weights: 5.963887753323185
-----
Iteration Number: 4496
Loss: 38.47772173634996
12 norm of gradients: 0.2797385207819044
12 norm of weights: 5.963794955560265
-----
Iteration Number: 4497
Loss: 38.473800923300495
12 norm of gradients: 0.27969127166630386
12 norm of weights: 5.963702168057804
-----
Iteration Number: 4498
Loss: 38.46988142648216
12 norm of gradients: 0.2796440470895642
12 norm of weights: 5.9636093908125325
-----
Iteration Number: 4499
Loss: 38.46596324500538
12 norm of gradients: 0.27959684703123183
12 norm of weights: 5.9635166238211825
-----
Iteration Number: 4500
Loss: 38.46204637797839
12 norm of gradients: 0.2795496714708722
12 norm of weights: 5.9634238670804915
-----
Iteration Number: 4501
Loss: 38.458130824498994
12 norm of gradients: 0.2795025203880684
12 norm of weights: 5.963331120587196
-----
Iteration Number: 4502
Loss: 38.454216583677216
12 norm of gradients: 0.2794553937624226
12 norm of weights: 5.963238384338038
-----
Iteration Number: 4503
Loss: 38.45030365463333
12 norm of gradients: 0.2794082915735551
12 norm of weights: 5.9631456583297595
-----
Iteration Number: 4504

```

Loss: 38.446392036471856
l2 norm of gradients: 0.2793612138011049
l2 norm of weights: 5.9630529425591074

Iteration Number: 4505
Loss: 38.44248172830573
l2 norm of gradients: 0.2793141604247292
l2 norm of weights: 5.962960237022829

Iteration Number: 4506
Loss: 38.438572729251746
l2 norm of gradients: 0.27926713142410364
l2 norm of weights: 5.962867541717675

Iteration Number: 4507
Loss: 38.434665038423006
l2 norm of gradients: 0.2792201267789222
l2 norm of weights: 5.962774856640401

Iteration Number: 4508
Loss: 38.430758654918755
l2 norm of gradients: 0.27917314646889746
l2 norm of weights: 5.962682181787761

Iteration Number: 4509
Loss: 38.42685357787381
l2 norm of gradients: 0.27912619047376014
l2 norm of weights: 5.962589517156514

Iteration Number: 4510
Loss: 38.42294980638912
l2 norm of gradients: 0.27907925877325945
l2 norm of weights: 5.962496862743423

Iteration Number: 4511
Loss: 38.419047339598684
l2 norm of gradients: 0.27903235134716275
l2 norm of weights: 5.962404218545248

Iteration Number: 4512
Loss: 38.41514617661108
l2 norm of gradients: 0.278985468175256
l2 norm of weights: 5.9623115845587575

Iteration Number: 4513
Loss: 38.41124631653195
l2 norm of gradients: 0.278938609237343
l2 norm of weights: 5.962218960780721

```

-----
Iteration Number: 4514
Loss: 38.40734775850048
l2 norm of gradients: 0.2788917745132463
l2 norm of weights: 5.962126347207908
-----
Iteration Number: 4515
Loss: 38.40345050162671
l2 norm of gradients: 0.2788449639828067
l2 norm of weights: 5.962033743837093
-----
Iteration Number: 4516
Loss: 38.39955454503837
l2 norm of gradients: 0.2787981776258829
l2 norm of weights: 5.961941150665052
-----
Iteration Number: 4517
Loss: 38.39565988785246
l2 norm of gradients: 0.27875141542235204
l2 norm of weights: 5.961848567688565
-----
Iteration Number: 4518
Loss: 38.39176652920132
l2 norm of gradients: 0.2787046773521097
l2 norm of weights: 5.961755994904412
-----
Iteration Number: 4519
Loss: 38.38787446818833
l2 norm of gradients: 0.2786579633950693
l2 norm of weights: 5.961663432309378
-----
Iteration Number: 4520
Loss: 38.383983703962926
l2 norm of gradients: 0.27861127353116283
l2 norm of weights: 5.961570879900249
-----
Iteration Number: 4521
Loss: 38.38009423563613
l2 norm of gradients: 0.27856460774034
l2 norm of weights: 5.961478337673813
-----
Iteration Number: 4522
Loss: 38.376206062340735
l2 norm of gradients: 0.2785179660025692
l2 norm of weights: 5.961385805626863
-----
Iteration Number: 4523
Loss: 38.37231918319908

```

```

12 norm of gradients: 0.2784713482978367
12 norm of weights: 5.961293283756191
-----
Iteration Number: 4524
Loss: 38.3684335973453
12 norm of gradients: 0.2784247546061469
12 norm of weights: 5.961200772058596
-----
Iteration Number: 4525
Loss: 38.36454930390151
12 norm of gradients: 0.27837818490752253
12 norm of weights: 5.961108270530874
-----
Iteration Number: 4526
Loss: 38.36066630199906
12 norm of gradients: 0.27833163918200426
12 norm of weights: 5.961015779169829
-----
Iteration Number: 4527
Loss: 38.35678459077245
12 norm of gradients: 0.27828511740965084
12 norm of weights: 5.960923297972265
-----
Iteration Number: 4528
Loss: 38.352904169363526
12 norm of gradients: 0.2782386195705392
12 norm of weights: 5.960830826934986
-----
Iteration Number: 4529
Loss: 38.349025036885976
12 norm of gradients: 0.27819214564476424
12 norm of weights: 5.960738366054802
-----
Iteration Number: 4530
Loss: 38.34514719247834
12 norm of gradients: 0.2781456956124391
12 norm of weights: 5.9606459153285245
-----
Iteration Number: 4531
Loss: 38.341270635285426
12 norm of gradients: 0.27809926945369495
12 norm of weights: 5.9605534747529685
-----
Iteration Number: 4532
Loss: 38.3373953644289
12 norm of gradients: 0.2780528671486807
12 norm of weights: 5.960461044324949
-----

```

Iteration Number: 4533
Loss: 38.33352137905385
l2 norm of gradients: 0.2780064886775636
l2 norm of weights: 5.960368624041284

Iteration Number: 4534
Loss: 38.32964867828908
l2 norm of gradients: 0.2779601340205287
l2 norm of weights: 5.960276213898798

Iteration Number: 4535
Loss: 38.3257772612819
l2 norm of gradients: 0.2779138031577792
l2 norm of weights: 5.960183813894311

Iteration Number: 4536
Loss: 38.32190712717103
l2 norm of gradients: 0.2778674960695361
l2 norm of weights: 5.960091424024652

Iteration Number: 4537
Loss: 38.318038275091304
l2 norm of gradients: 0.27782121273603844
l2 norm of weights: 5.959999044286649

Iteration Number: 4538
Loss: 38.31417070417482
l2 norm of gradients: 0.27777495313754325
l2 norm of weights: 5.959906674677133

Iteration Number: 4539
Loss: 38.31030441358344
l2 norm of gradients: 0.2777287172543254
l2 norm of weights: 5.959814315192938

Iteration Number: 4540
Loss: 38.30643940244297
l2 norm of gradients: 0.27768250506667785
l2 norm of weights: 5.959721965830898

Iteration Number: 4541
Loss: 38.302575669900996
l2 norm of gradients: 0.27763631655491117
l2 norm of weights: 5.959629626587855

Iteration Number: 4542
Loss: 38.29871321510034
l2 norm of gradients: 0.2775901516993541


```

12 norm of weights: 5.959537297460648
-----
Iteration Number: 4543
Loss: 38.29485203718671
12 norm of gradients: 0.27754401048035293
12 norm of weights: 5.95944497844612
-----
Iteration Number: 4544
Loss: 38.29099213530463
12 norm of gradients: 0.27749789287827237
12 norm of weights: 5.959352669541118
-----
Iteration Number: 4545
Loss: 38.287133508599396
12 norm of gradients: 0.2774517988734943
12 norm of weights: 5.95926037074249
-----
Iteration Number: 4546
Loss: 38.28327615622253
12 norm of gradients: 0.277405728446419
12 norm of weights: 5.9591680820470865
-----
Iteration Number: 4547
Loss: 38.27942007732477
12 norm of gradients: 0.2773596815774641
12 norm of weights: 5.959075803451761
-----
Iteration Number: 4548
Loss: 38.27556527104494
12 norm of gradients: 0.27731365824706533
12 norm of weights: 5.9589835349533695
-----
Iteration Number: 4549
Loss: 38.27171173654494
12 norm of gradients: 0.2772676584356761
12 norm of weights: 5.958891276548769
-----
Iteration Number: 4550
Loss: 38.26785947295551
12 norm of gradients: 0.2772216821237678
12 norm of weights: 5.958799028234822
-----
Iteration Number: 4551
Loss: 38.26400847945018
12 norm of gradients: 0.2771757292918291
12 norm of weights: 5.958706790008389
-----
Iteration Number: 4552

```

```
Loss: 38.26015875516376
12 norm of gradients: 0.27712979992036696
12 norm of weights: 5.958614561866337
-----
Iteration Number: 4553
Loss: 38.256310299258175
12 norm of gradients: 0.2770838939899057
12 norm of weights: 5.958522343805534
-----
Iteration Number: 4554
Loss: 38.25246311089302
12 norm of gradients: 0.2770380114809875
12 norm of weights: 5.958430135822848
-----
Iteration Number: 4555
Loss: 38.24861718921569
12 norm of gradients: 0.27699215237417224
12 norm of weights: 5.958337937915155
-----
Iteration Number: 4556
Loss: 38.244772533383504
12 norm of gradients: 0.27694631665003744
12 norm of weights: 5.9582457500793256
-----
Iteration Number: 4557
Loss: 38.24092914254987
12 norm of gradients: 0.2769005042891785
12 norm of weights: 5.958153572312241
-----
Iteration Number: 4558
Loss: 38.23708701587362
12 norm of gradients: 0.2768547152722081
12 norm of weights: 5.95806140461078
-----
Iteration Number: 4559
Loss: 38.23324615250811
12 norm of gradients: 0.2768089495797569
12 norm of weights: 5.9579692469718255
-----
Iteration Number: 4560
Loss: 38.22940655162318
12 norm of gradients: 0.2767632071924729
12 norm of weights: 5.9578770993922605
-----
Iteration Number: 4561
Loss: 38.22556821237321
12 norm of gradients: 0.2767174880910222
12 norm of weights: 5.957784961868973
```

```

-----
Iteration Number: 4562
Loss: 38.221731133917636
l2 norm of gradients: 0.2766717922560879
l2 norm of weights: 5.957692834398852
-----
Iteration Number: 4563
Loss: 38.21789531542461
l2 norm of gradients: 0.2766261196683712
l2 norm of weights: 5.9576007169787895
-----
Iteration Number: 4564
Loss: 38.21406075603882
l2 norm of gradients: 0.2765804703085905
l2 norm of weights: 5.95750860960568
-----
Iteration Number: 4565
Loss: 38.21022745494175
l2 norm of gradients: 0.27653484415748203
l2 norm of weights: 5.957416512276421
-----
Iteration Number: 4566
Loss: 38.20639541129253
l2 norm of gradients: 0.2764892411957994
l2 norm of weights: 5.95732442498791
-----
Iteration Number: 4567
Loss: 38.20256462425346
l2 norm of gradients: 0.2764436614043138
l2 norm of weights: 5.9572323477370475
-----
Iteration Number: 4568
Loss: 38.198735092984876
l2 norm of gradients: 0.27639810476381405
l2 norm of weights: 5.95714028052074
-----
Iteration Number: 4569
Loss: 38.194906816662055
l2 norm of gradients: 0.27635257125510626
l2 norm of weights: 5.9570482233358915
-----
Iteration Number: 4570
Loss: 38.191079794456535
l2 norm of gradients: 0.27630706085901424
l2 norm of weights: 5.956956176179411
-----
Iteration Number: 4571
Loss: 38.18725402552528

```

```

12 norm of gradients: 0.27626157355637915
12 norm of weights: 5.95686413904821
-----
Iteration Number: 4572
Loss: 38.183429509030496
12 norm of gradients: 0.2762161093280597
12 norm of weights: 5.956772111939202
-----
Iteration Number: 4573
Loss: 38.179606244168035
12 norm of gradients: 0.27617066815493196
12 norm of weights: 5.956680094849301
-----
Iteration Number: 4574
Loss: 38.17578423008155
12 norm of gradients: 0.2761252500178895
12 norm of weights: 5.956588087775426
-----
Iteration Number: 4575
Loss: 38.17196346595465
12 norm of gradients: 0.27607985489784337
12 norm of weights: 5.956496090714497
-----
Iteration Number: 4576
Loss: 38.1681439509605
12 norm of gradients: 0.2760344827757218
12 norm of weights: 5.956404103663438
-----
Iteration Number: 4577
Loss: 38.164325684270516
12 norm of gradients: 0.2759891336324707
12 norm of weights: 5.956312126619172
-----
Iteration Number: 4578
Loss: 38.16050866505903
12 norm of gradients: 0.27594380744905317
12 norm of weights: 5.956220159578627
-----
Iteration Number: 4579
Loss: 38.15669289249385
12 norm of gradients: 0.2758985042064497
12 norm of weights: 5.956128202538733
-----
Iteration Number: 4580
Loss: 38.15287836575542
12 norm of gradients: 0.27585322388565825
12 norm of weights: 5.956036255496422
-----

```

Iteration Number: 4581
Loss: 38.149065084025246
l2 norm of gradients: 0.27580796646769395
l2 norm of weights: 5.955944318448629

Iteration Number: 4582
Loss: 38.145253046477315
l2 norm of gradients: 0.2757627319335894
l2 norm of weights: 5.95585239139229

Iteration Number: 4583
Loss: 38.14144225228248
l2 norm of gradients: 0.2757175202643945
l2 norm of weights: 5.955760474324345

Iteration Number: 4584
Loss: 38.137632700619534
l2 norm of gradients: 0.27567233144117625
l2 norm of weights: 5.955668567241734

Iteration Number: 4585
Loss: 38.133824390681426
l2 norm of gradients: 0.27562716544501914
l2 norm of weights: 5.955576670141403

Iteration Number: 4586
Loss: 38.130017321639926
l2 norm of gradients: 0.2755820222570249
l2 norm of weights: 5.955484783020295

Iteration Number: 4587
Loss: 38.126211492673335
l2 norm of gradients: 0.2755369018583125
l2 norm of weights: 5.95539290587536

Iteration Number: 4588
Loss: 38.12240690296094
l2 norm of gradients: 0.27549180423001807
l2 norm of weights: 5.95530103870355

Iteration Number: 4589
Loss: 38.118603551695415
l2 norm of gradients: 0.2754467293532951
l2 norm of weights: 5.955209181501815

Iteration Number: 4590
Loss: 38.11480143806102
l2 norm of gradients: 0.27540167720931424

```
12 norm of weights: 5.955117334267114
-----
Iteration Number: 4591
Loss: 38.111000561223065
12 norm of gradients: 0.2753566477792632
12 norm of weights: 5.955025496996402
-----
Iteration Number: 4592
Loss: 38.10720092038501
12 norm of gradients: 0.27531164104434713
12 norm of weights: 5.954933669686641
-----
Iteration Number: 4593
Loss: 38.103402514729
12 norm of gradients: 0.27526665698578817
12 norm of weights: 5.954841852334791
-----
Iteration Number: 4594
Loss: 38.0996053434455
12 norm of gradients: 0.2752216955848256
12 norm of weights: 5.9547500449378195
-----
Iteration Number: 4595
Loss: 38.09580940570862
12 norm of gradients: 0.27517675682271603
12 norm of weights: 5.95465824749269
-----
Iteration Number: 4596
Loss: 38.09201470072234
12 norm of gradients: 0.2751318406807331
12 norm of weights: 5.954566459996374
-----
Iteration Number: 4597
Loss: 38.088221227661116
12 norm of gradients: 0.27508694714016746
12 norm of weights: 5.954474682445842
-----
Iteration Number: 4598
Loss: 38.084428985723996
12 norm of gradients: 0.27504207618232696
12 norm of weights: 5.95438291483807
-----
Iteration Number: 4599
Loss: 38.080637974096454
12 norm of gradients: 0.27499722778853664
12 norm of weights: 5.954291157170032
-----
Iteration Number: 4600
```

Loss: 38.07684819198145
l2 norm of gradients: 0.27495240194013837
l2 norm of weights: 5.954199409438706

Iteration Number: 4601
Loss: 38.07305963855926
l2 norm of gradients: 0.2749075986184913
l2 norm of weights: 5.954107671641074

Iteration Number: 4602
Loss: 38.06927231301709
l2 norm of gradients: 0.2748628178049715
l2 norm of weights: 5.954015943774119

Iteration Number: 4603
Loss: 38.0654862145665
l2 norm of gradients: 0.27481805948097227
l2 norm of weights: 5.953924225834827

Iteration Number: 4604
Loss: 38.06170134239263
l2 norm of gradients: 0.27477332362790363
l2 norm of weights: 5.953832517820183

Iteration Number: 4605
Loss: 38.05791769569162
l2 norm of gradients: 0.2747286102271927
l2 norm of weights: 5.953740819727179

Iteration Number: 4606
Loss: 38.05413527366264
l2 norm of gradients: 0.27468391926028385
l2 norm of weights: 5.9536491315528055

Iteration Number: 4607
Loss: 38.05035407549681
l2 norm of gradients: 0.27463925070863804
l2 norm of weights: 5.953557453294059

Iteration Number: 4608
Loss: 38.04657410039873
l2 norm of gradients: 0.2745946045537336
l2 norm of weights: 5.953465784947935

Iteration Number: 4609
Loss: 38.04279534755849
l2 norm of gradients: 0.2745499807770654
l2 norm of weights: 5.95337412651143

```
-----  
Iteration Number: 4610  
Loss: 38.03901781617505  
l2 norm of gradients: 0.2745053793601455  
l2 norm of weights: 5.953282477981549  
-----  
Iteration Number: 4611  
Loss: 38.03524150546942  
l2 norm of gradients: 0.27446080028450287  
l2 norm of weights: 5.953190839355296  
-----  
Iteration Number: 4612  
Loss: 38.03146641461189  
l2 norm of gradients: 0.2744162435316834  
l2 norm of weights: 5.953099210629671  
-----  
Iteration Number: 4613  
Loss: 38.027692542824596  
l2 norm of gradients: 0.27437170908324976  
l2 norm of weights: 5.953007591801689  
-----  
Iteration Number: 4614  
Loss: 38.0239198893087  
l2 norm of gradients: 0.2743271969207815  
l2 norm of weights: 5.952915982868355  
-----  
Iteration Number: 4615  
Loss: 38.02014845325876  
l2 norm of gradients: 0.27428270702587526  
l2 norm of weights: 5.952824383826684  
-----  
Iteration Number: 4616  
Loss: 38.01637823387779  
l2 norm of gradients: 0.2742382393801443  
l2 norm of weights: 5.95273279467369  
-----  
Iteration Number: 4617  
Loss: 38.01260923038712  
l2 norm of gradients: 0.27419379396521887  
l2 norm of weights: 5.9526412154063895  
-----  
Iteration Number: 4618  
Loss: 38.00884144197774  
l2 norm of gradients: 0.2741493707627458  
l2 norm of weights: 5.9525496460218035  
-----  
Iteration Number: 4619  
Loss: 38.00507486785458
```



```

12 norm of gradients: 0.27410496975438914
12 norm of weights: 5.95245808651695
-----
Iteration Number: 4620
Loss: 38.00130950723513
12 norm of gradients: 0.27406059092182944
12 norm of weights: 5.952366536888857
-----
Iteration Number: 4621
Loss: 37.99754535932017
12 norm of gradients: 0.2740162342467641
12 norm of weights: 5.952274997134547
-----
Iteration Number: 4622
Loss: 37.99378242331761
12 norm of gradients: 0.27397189971090735
12 norm of weights: 5.952183467251051
-----
Iteration Number: 4623
Loss: 37.99002069844206
12 norm of gradients: 0.27392758729599015
12 norm of weights: 5.952091947235397
-----
Iteration Number: 4624
Loss: 37.98626018390417
12 norm of gradients: 0.27388329698376024
12 norm of weights: 5.952000437084618
-----
Iteration Number: 4625
Loss: 37.98250087891144
12 norm of gradients: 0.273839028755982
12 norm of weights: 5.95190893679575
-----
Iteration Number: 4626
Loss: 37.978742782668604
12 norm of gradients: 0.27379478259443657
12 norm of weights: 5.95181744636583
-----
Iteration Number: 4627
Loss: 37.97498589440677
12 norm of gradients: 0.2737505584809218
12 norm of weights: 5.951725965791896
-----
Iteration Number: 4628
Loss: 37.97123021331884
12 norm of gradients: 0.27370635639725227
12 norm of weights: 5.95163449507099
-----

```

Iteration Number: 4629
Loss: 37.96747573863441
l2 norm of gradients: 0.27366217632525924
l2 norm of weights: 5.951543034200156

Iteration Number: 4630
Loss: 37.96372246955944
l2 norm of gradients: 0.2736180182467906
l2 norm of weights: 5.95145158317644

Iteration Number: 4631
Loss: 37.95997040531056
l2 norm of gradients: 0.273573882143711
l2 norm of weights: 5.951360141996891

Iteration Number: 4632
Loss: 37.95621954511139
l2 norm of gradients: 0.27352976799790146
l2 norm of weights: 5.951268710658558

Iteration Number: 4633
Loss: 37.95246988817098
l2 norm of gradients: 0.2734856757912598
l2 norm of weights: 5.951177289158493

Iteration Number: 4634
Loss: 37.94872143371364
l2 norm of gradients: 0.27344160550570057
l2 norm of weights: 5.951085877493752

Iteration Number: 4635
Loss: 37.944974180938964
l2 norm of gradients: 0.2733975571231546
l2 norm of weights: 5.950994475661392

Iteration Number: 4636
Loss: 37.94122812910111
l2 norm of gradients: 0.2733535306255698
l2 norm of weights: 5.9509030836584715

Iteration Number: 4637
Loss: 37.93748327738869
l2 norm of gradients: 0.27330952599491015
l2 norm of weights: 5.950811701482054

Iteration Number: 4638
Loss: 37.933739625043756
l2 norm of gradients: 0.2732655432131564

```
l2 norm of weights: 5.9507203291292
-----
Iteration Number: 4639
Loss: 37.92999717126841
l2 norm of gradients: 0.27322158226230586
l2 norm of weights: 5.950628966596976
-----
Iteration Number: 4640
Loss: 37.92625591530938
l2 norm of gradients: 0.27317764312437237
l2 norm of weights: 5.950537613882451
-----
Iteration Number: 4641
Loss: 37.92251585636903
l2 norm of gradients: 0.2731337257813862
l2 norm of weights: 5.950446270982695
-----
Iteration Number: 4642
Loss: 37.91877699367805
l2 norm of gradients: 0.27308983021539424
l2 norm of weights: 5.9503549378947795
-----
Iteration Number: 4643
Loss: 37.91503932646278
l2 norm of gradients: 0.2730459564084597
l2 norm of weights: 5.95026361461578
-----
Iteration Number: 4644
Loss: 37.91130285394731
l2 norm of gradients: 0.2730021043426626
l2 norm of weights: 5.950172301142772
-----
Iteration Number: 4645
Loss: 37.907567575358186
l2 norm of gradients: 0.272958274000099
l2 norm of weights: 5.950080997472835
-----
Iteration Number: 4646
Loss: 37.903833489932765
l2 norm of gradients: 0.2729144653628817
l2 norm of weights: 5.94998970360305
-----
Iteration Number: 4647
Loss: 37.90010059687239
l2 norm of gradients: 0.2728706784131399
l2 norm of weights: 5.9498984195305
-----
Iteration Number: 4648
```

```

Loss: 37.89636889543144
l2 norm of gradients: 0.272826913133019
l2 norm of weights: 5.9498071452522705
-----
Iteration Number: 4649
Loss: 37.89263838483075
l2 norm of gradients: 0.2727831695046811
l2 norm of weights: 5.949715880765448
-----
Iteration Number: 4650
Loss: 37.88890906429649
l2 norm of gradients: 0.27273944751030466
l2 norm of weights: 5.949624626067124
-----
Iteration Number: 4651
Loss: 37.885180933059985
l2 norm of gradients: 0.27269574713208433
l2 norm of weights: 5.9495333811543905
-----
Iteration Number: 4652
Loss: 37.881453990361805
l2 norm of gradients: 0.2726520683522313
l2 norm of weights: 5.94944214602434
-----
Iteration Number: 4653
Loss: 37.877728235422786
l2 norm of gradients: 0.2726084111529731
l2 norm of weights: 5.949350920674068
-----
Iteration Number: 4654
Loss: 37.8740036674798
l2 norm of gradients: 0.2725647755165534
l2 norm of weights: 5.949259705100674
-----
Iteration Number: 4655
Loss: 37.87028028576295
l2 norm of gradients: 0.27252116142523247
l2 norm of weights: 5.949168499301259
-----
Iteration Number: 4656
Loss: 37.866558089521575
l2 norm of gradients: 0.27247756886128693
l2 norm of weights: 5.949077303272925
-----
Iteration Number: 4657
Loss: 37.86283707797783
l2 norm of gradients: 0.2724339978070092
l2 norm of weights: 5.948986117012777

```

```

-----
Iteration Number: 4658
Loss: 37.85911725036199
l2 norm of gradients: 0.2723904482447088
l2 norm of weights: 5.948894940517921
-----
Iteration Number: 4659
Loss: 37.855398605923476
l2 norm of gradients: 0.27234692015671075
l2 norm of weights: 5.948803773785466
-----
Iteration Number: 4660
Loss: 37.85168114389734
l2 norm of gradients: 0.2723034135253567
l2 norm of weights: 5.948712616812525
-----
Iteration Number: 4661
Loss: 37.84796486351896
l2 norm of gradients: 0.27225992833300466
l2 norm of weights: 5.948621469596209
-----
Iteration Number: 4662
Loss: 37.84424976403025
l2 norm of gradients: 0.2722164645620287
l2 norm of weights: 5.948530332133635
-----
Iteration Number: 4663
Loss: 37.84053584466167
l2 norm of gradients: 0.27217302219481904
l2 norm of weights: 5.948439204421921
-----
Iteration Number: 4664
Loss: 37.83682310466429
l2 norm of gradients: 0.27212960121378227
l2 norm of weights: 5.948348086458185
-----
Iteration Number: 4665
Loss: 37.83311154328045
l2 norm of gradients: 0.27208620160134117
l2 norm of weights: 5.94825697823955
-----
Iteration Number: 4666
Loss: 37.82940115974663
l2 norm of gradients: 0.27204282333993457
l2 norm of weights: 5.948165879763138
-----
Iteration Number: 4667
Loss: 37.82569195330329

```

```

12 norm of gradients: 0.27199946641201744
12 norm of weights: 5.9480747910260785
-----
Iteration Number: 4668
Loss: 37.821983923195376
12 norm of gradients: 0.27195613080006137
12 norm of weights: 5.947983712025496
-----
Iteration Number: 4669
Loss: 37.81827706866958
12 norm of gradients: 0.27191281648655347
12 norm of weights: 5.947892642758524
-----
Iteration Number: 4670
Loss: 37.81457138896995
12 norm of gradients: 0.2718695234539974
12 norm of weights: 5.947801583222294
-----
Iteration Number: 4671
Loss: 37.81086688334003
12 norm of gradients: 0.2718262516849127
12 norm of weights: 5.947710533413939
-----
Iteration Number: 4672
Loss: 37.80716355102478
12 norm of gradients: 0.27178300116183507
12 norm of weights: 5.947619493330597
-----
Iteration Number: 4673
Loss: 37.8034613912812
12 norm of gradients: 0.27173977186731646
12 norm of weights: 5.947528462969406
-----
Iteration Number: 4674
Loss: 37.79976040334143
12 norm of gradients: 0.27169656378392465
12 norm of weights: 5.947437442327508
-----
Iteration Number: 4675
Loss: 37.796060586472976
12 norm of gradients: 0.2716533768942437
12 norm of weights: 5.947346431402044
-----
Iteration Number: 4676
Loss: 37.7923619399077
12 norm of gradients: 0.27161021118087364
12 norm of weights: 5.947255430190162
-----

```

Iteration Number: 4677
Loss: 37.78866446290794
l2 norm of gradients: 0.2715670666264306
l2 norm of weights: 5.947164438689006

Iteration Number: 4678
Loss: 37.784968154715855
l2 norm of gradients: 0.27152394321354645
l2 norm of weights: 5.947073456895727

Iteration Number: 4679
Loss: 37.78127301457376
l2 norm of gradients: 0.27148084092486946
l2 norm of weights: 5.946982484807476

Iteration Number: 4680
Loss: 37.777579041764696
l2 norm of gradients: 0.27143775974306367
l2 norm of weights: 5.946891522421406

Iteration Number: 4681
Loss: 37.773886235504946
l2 norm of gradients: 0.2713946996508092
l2 norm of weights: 5.946800569734673

Iteration Number: 4682
Loss: 37.77019459507779
l2 norm of gradients: 0.2713516606308021
l2 norm of weights: 5.946709626744434

Iteration Number: 4683
Loss: 37.76650411972654
l2 norm of gradients: 0.27130864266575444
l2 norm of weights: 5.94661869344785

Iteration Number: 4684
Loss: 37.76281480869905
l2 norm of gradients: 0.27126564573839396
l2 norm of weights: 5.946527769842081

Iteration Number: 4685
Loss: 37.75912666125281
l2 norm of gradients: 0.27122266983146487
l2 norm of weights: 5.946436855924293

Iteration Number: 4686
Loss: 37.755439676651086
l2 norm of gradients: 0.2711797149277268

```

12 norm of weights: 5.94634595169165
-----
Iteration Number: 4687
Loss: 37.751753854150365
12 norm of gradients: 0.27113678100995553
12 norm of weights: 5.946255057141319
-----
Iteration Number: 4688
Loss: 37.74806919301067
12 norm of gradients: 0.2710938680609426
12 norm of weights: 5.946164172270473
-----
Iteration Number: 4689
Loss: 37.74438569247969
12 norm of gradients: 0.27105097606349554
12 norm of weights: 5.946073297076283
-----
Iteration Number: 4690
Loss: 37.74070335183061
12 norm of gradients: 0.27100810500043787
12 norm of weights: 5.945982431555922
-----
Iteration Number: 4691
Loss: 37.73702217030376
12 norm of gradients: 0.2709652548546087
12 norm of weights: 5.945891575706568
-----
Iteration Number: 4692
Loss: 37.73334214718096
12 norm of gradients: 0.27092242560886304
12 norm of weights: 5.9458007295253985
-----
Iteration Number: 4693
Loss: 37.7296632817163
12 norm of gradients: 0.27087961724607185
12 norm of weights: 5.945709893009592
-----
Iteration Number: 4694
Loss: 37.72598557316336
12 norm of gradients: 0.2708368297491219
12 norm of weights: 5.945619066156334
-----
Iteration Number: 4695
Loss: 37.72230902080109
12 norm of gradients: 0.2707940631009156
12 norm of weights: 5.945528248962809
-----
Iteration Number: 4696

```


Loss: 37.718633623879285
l2 norm of gradients: 0.27075131728437135
l2 norm of weights: 5.945437441426201

Iteration Number: 4697
Loss: 37.714959381671406
l2 norm of gradients: 0.27070859228242317
l2 norm of weights: 5.945346643543699

Iteration Number: 4698
Loss: 37.71128629343487
l2 norm of gradients: 0.27066588807802094
l2 norm of weights: 5.945255855312495

Iteration Number: 4699
Loss: 37.70761435844012
l2 norm of gradients: 0.2706232046541302
l2 norm of weights: 5.945165076729781

Iteration Number: 4700
Loss: 37.70394357594596
l2 norm of gradients: 0.2705805419937324
l2 norm of weights: 5.945074307792752

Iteration Number: 4701
Loss: 37.700273945237626
l2 norm of gradients: 0.2705379000798244
l2 norm of weights: 5.944983548498604

Iteration Number: 4702
Loss: 37.69660546556482
l2 norm of gradients: 0.27049527889541924
l2 norm of weights: 5.944892798844537

Iteration Number: 4703
Loss: 37.69293813619843
l2 norm of gradients: 0.2704526784235452
l2 norm of weights: 5.94480205882775

Iteration Number: 4704
Loss: 37.689271956413926
l2 norm of gradients: 0.2704100986472465
l2 norm of weights: 5.944711328445448

Iteration Number: 4705
Loss: 37.685606925479895
l2 norm of gradients: 0.27036753954958304
l2 norm of weights: 5.9446206076948345

```

-----
Iteration Number: 4706
Loss: 37.68194304266114
l2 norm of gradients: 0.2703250011136302
l2 norm of weights: 5.9445298965731155
-----
Iteration Number: 4707
Loss: 37.67828030724279
l2 norm of gradients: 0.2702824833224792
l2 norm of weights: 5.944439195077502
-----
Iteration Number: 4708
Loss: 37.67461871848552
l2 norm of gradients: 0.2702399861592367
l2 norm of weights: 5.944348503205204
-----
Iteration Number: 4709
Loss: 37.670958275661086
l2 norm of gradients: 0.2701975096070253
l2 norm of weights: 5.944257820953435
-----
Iteration Number: 4710
Loss: 37.667298978044684
l2 norm of gradients: 0.27015505364898296
l2 norm of weights: 5.944167148319409
-----
Iteration Number: 4711
Loss: 37.66364082491448
l2 norm of gradients: 0.270112618268263
l2 norm of weights: 5.944076485300344
-----
Iteration Number: 4712
Loss: 37.65998381554487
l2 norm of gradients: 0.27007020344803495
l2 norm of weights: 5.943985831893459
-----
Iteration Number: 4713
Loss: 37.656327949205206
l2 norm of gradients: 0.2700278091714835
l2 norm of weights: 5.943895188095974
-----
Iteration Number: 4714
Loss: 37.652673225180955
l2 norm of gradients: 0.2699854354218088
l2 norm of weights: 5.943804553905112
-----
Iteration Number: 4715
Loss: 37.64901964273963

```

```

12 norm of gradients: 0.26994308218222696
12 norm of weights: 5.9437139293181005
-----
Iteration Number: 4716
Loss: 37.645367201169265
12 norm of gradients: 0.2699007494359691
12 norm of weights: 5.9436233143321635
-----
Iteration Number: 4717
Loss: 37.64171589973573
12 norm of gradients: 0.26985843716628233
12 norm of weights: 5.943532708944531
-----
Iteration Number: 4718
Loss: 37.63806573772859
12 norm of gradients: 0.269816145356429
12 norm of weights: 5.943442113152435
-----
Iteration Number: 4719
Loss: 37.63441671441725
12 norm of gradients: 0.26977387398968705
12 norm of weights: 5.943351526953108
-----
Iteration Number: 4720
Loss: 37.630768829090556
12 norm of gradients: 0.26973162304934983
12 norm of weights: 5.943260950343785
-----
Iteration Number: 4721
Loss: 37.62712208103355
12 norm of gradients: 0.26968939251872615
12 norm of weights: 5.943170383321703
-----
Iteration Number: 4722
Loss: 37.62347646951062
12 norm of gradients: 0.26964718238114044
12 norm of weights: 5.9430798258841016
-----
Iteration Number: 4723
Loss: 37.61983199382504
12 norm of gradients: 0.26960499261993237
12 norm of weights: 5.942989278028221
-----
Iteration Number: 4724
Loss: 37.61618865325092
12 norm of gradients: 0.2695628232184572
12 norm of weights: 5.942898739751304
-----

```

Iteration Number: 4725
Loss: 37.61254644706795
l2 norm of gradients: 0.2695206741600855
l2 norm of weights: 5.942808211050596

Iteration Number: 4726
Loss: 37.60890537456016
l2 norm of gradients: 0.2694785454282033
l2 norm of weights: 5.942717691923344

Iteration Number: 4727
Loss: 37.60526543501891
l2 norm of gradients: 0.2694364370062119
l2 norm of weights: 5.942627182366799

Iteration Number: 4728
Loss: 37.60162662773523
l2 norm of gradients: 0.26939434887752817
l2 norm of weights: 5.942536682378209

Iteration Number: 4729
Loss: 37.597988951973754
l2 norm of gradients: 0.26935228102558423
l2 norm of weights: 5.942446191954827

Iteration Number: 4730
Loss: 37.59435240704697
l2 norm of gradients: 0.26931023343382754
l2 norm of weights: 5.942355711093909

Iteration Number: 4731
Loss: 37.590716992233666
l2 norm of gradients: 0.26926820608572094
l2 norm of weights: 5.942265239792714

Iteration Number: 4732
Loss: 37.58708270681312
l2 norm of gradients: 0.26922619896474276
l2 norm of weights: 5.9421747780484955

Iteration Number: 4733
Loss: 37.58344955008601
l2 norm of gradients: 0.2691842120543862
l2 norm of weights: 5.942084325858518

Iteration Number: 4734
Loss: 37.579817521339145
l2 norm of gradients: 0.2691422453381602

```

12 norm of weights: 5.941993883220044
-----
Iteration Number: 4735
Loss: 37.5761866198663
12 norm of gradients: 0.26910029879958886
12 norm of weights: 5.941903450130338
-----
Iteration Number: 4736
Loss: 37.57255684495029
12 norm of gradients: 0.26905837242221137
12 norm of weights: 5.941813026586667
-----
Iteration Number: 4737
Loss: 37.56892819588681
12 norm of gradients: 0.26901646618958247
12 norm of weights: 5.941722612586298
-----
Iteration Number: 4738
Loss: 37.56530067197904
12 norm of gradients: 0.2689745800852721
12 norm of weights: 5.941632208126503
-----
Iteration Number: 4739
Loss: 37.561674272500845
12 norm of gradients: 0.2689327140928651
12 norm of weights: 5.941541813204553
-----
Iteration Number: 4740
Loss: 37.55804899676612
12 norm of gradients: 0.26889086819596203
12 norm of weights: 5.941451427817725
-----
Iteration Number: 4741
Loss: 37.55442484404801
12 norm of gradients: 0.26884904237817836
12 norm of weights: 5.941361051963294
-----
Iteration Number: 4742
Loss: 37.55080181366186
12 norm of gradients: 0.2688072366231448
12 norm of weights: 5.941270685638538
-----
Iteration Number: 4743
Loss: 37.54717990489996
12 norm of gradients: 0.26876545091450726
12 norm of weights: 5.941180328840739
-----
Iteration Number: 4744

```

Loss: 37.543559117046044
l2 norm of gradients: 0.26872368523592693
l2 norm of weights: 5.941089981567178

Iteration Number: 4745
Loss: 37.53993944941004
l2 norm of gradients: 0.26868193957108
l2 norm of weights: 5.940999643815139

Iteration Number: 4746
Loss: 37.536320901287176
l2 norm of gradients: 0.26864021390365794
l2 norm of weights: 5.940909315581907

Iteration Number: 4747
Loss: 37.53270347196989
l2 norm of gradients: 0.26859850821736725
l2 norm of weights: 5.940818996864774

Iteration Number: 4748
Loss: 37.52908716077256
l2 norm of gradients: 0.26855682249592966
l2 norm of weights: 5.940728687661026

Iteration Number: 4749
Loss: 37.52547196697862
l2 norm of gradients: 0.2685151567230819
l2 norm of weights: 5.940638387967956

Iteration Number: 4750
Loss: 37.521857889898776
l2 norm of gradients: 0.268473510882576
l2 norm of weights: 5.94054809778286

Iteration Number: 4751
Loss: 37.51824492882747
l2 norm of gradients: 0.2684318849581788
l2 norm of weights: 5.940457817103031

Iteration Number: 4752
Loss: 37.51463308308197
l2 norm of gradients: 0.2683902789336725
l2 norm of weights: 5.94036754592577

Iteration Number: 4753
Loss: 37.51102235194551
l2 norm of gradients: 0.268348692792854
l2 norm of weights: 5.940277284248372

```

-----
Iteration Number: 4754
Loss: 37.50741273473356
l2 norm of gradients: 0.26830712651953564
l2 norm of weights: 5.940187032068143
-----
Iteration Number: 4755
Loss: 37.50380423074341
l2 norm of gradients: 0.26826558009754464
l2 norm of weights: 5.940096789382384
-----
Iteration Number: 4756
Loss: 37.50019683928276
l2 norm of gradients: 0.26822405351072315
l2 norm of weights: 5.9400065561884015
-----
Iteration Number: 4757
Loss: 37.496590559664256
l2 norm of gradients: 0.26818254674292835
l2 norm of weights: 5.939916332483502
-----
Iteration Number: 4758
Loss: 37.49298539118411
l2 norm of gradients: 0.26814105977803265
l2 norm of weights: 5.939826118264994
-----
Iteration Number: 4759
Loss: 37.489381333147385
l2 norm of gradients: 0.2680995925999233
l2 norm of weights: 5.939735913530192
-----
Iteration Number: 4760
Loss: 37.48577838487175
l2 norm of gradients: 0.26805814519250226
l2 norm of weights: 5.939645718276404
-----
Iteration Number: 4761
Loss: 37.48217654565681
l2 norm of gradients: 0.26801671753968687
l2 norm of weights: 5.93955553250095
-----
Iteration Number: 4762
Loss: 37.47857581481697
l2 norm of gradients: 0.2679753096254093
l2 norm of weights: 5.939465356201144
-----
Iteration Number: 4763
Loss: 37.47497619165399

```

```

12 norm of gradients: 0.2679339214336165
12 norm of weights: 5.939375189374305
-----
Iteration Number: 4764
Loss: 37.47137767548935
12 norm of gradients: 0.2678925529482704
12 norm of weights: 5.939285032017754
-----
Iteration Number: 4765
Loss: 37.467780265626764
12 norm of gradients: 0.2678512041533481
12 norm of weights: 5.939194884128813
-----
Iteration Number: 4766
Loss: 37.46418396136847
12 norm of gradients: 0.2678098750328412
12 norm of weights: 5.939104745704808
-----
Iteration Number: 4767
Loss: 37.46058876203861
12 norm of gradients: 0.26776856557075646
12 norm of weights: 5.939014616743064
-----
Iteration Number: 4768
Loss: 37.45699466695746
12 norm of gradients: 0.26772727575111543
12 norm of weights: 5.938924497240909
-----
Iteration Number: 4769
Loss: 37.4534016754204
12 norm of gradients: 0.2676860055579545
12 norm of weights: 5.938834387195674
-----
Iteration Number: 4770
Loss: 37.44980978674634
12 norm of gradients: 0.267644754975325
12 norm of weights: 5.93874428660469
-----
Iteration Number: 4771
Loss: 37.446219000258985
12 norm of gradients: 0.267603523987293
12 norm of weights: 5.938654195465292
-----
Iteration Number: 4772
Loss: 37.44262931526228
12 norm of gradients: 0.2675623125779394
12 norm of weights: 5.938564113774815
-----

```


Iteration Number: 4773
Loss: 37.439040731075785
l2 norm of gradients: 0.26752112073136003
l2 norm of weights: 5.938474041530597

Iteration Number: 4774
Loss: 37.43545324701869
l2 norm of gradients: 0.2674799484316653
l2 norm of weights: 5.938383978729978

Iteration Number: 4775
Loss: 37.43186686240985
l2 norm of gradients: 0.2674387956629807
l2 norm of weights: 5.9382939253702975

Iteration Number: 4776
Loss: 37.428281576563165
l2 norm of gradients: 0.2673976624094462
l2 norm of weights: 5.9382038814489

Iteration Number: 4777
Loss: 37.42469738879907
l2 norm of gradients: 0.26735654865521674
l2 norm of weights: 5.938113846963131

Iteration Number: 4778
Loss: 37.421114298435455
l2 norm of gradients: 0.267315454384462
l2 norm of weights: 5.938023821910337

Iteration Number: 4779
Loss: 37.417532304787585
l2 norm of gradients: 0.2672743795813663
l2 norm of weights: 5.937933806287867

Iteration Number: 4780
Loss: 37.41395140718465
l2 norm of gradients: 0.26723332423012874
l2 norm of weights: 5.937843800093072

Iteration Number: 4781
Loss: 37.410371604941936
l2 norm of gradients: 0.2671922883149631
l2 norm of weights: 5.937753803323304

Iteration Number: 4782
Loss: 37.40679289738273
l2 norm of gradients: 0.26715127182009785

```

12 norm of weights: 5.937663815975918
-----
Iteration Number: 4783
Loss: 37.403215283830285
12 norm of gradients: 0.2671102747297763
12 norm of weights: 5.937573838048269
-----
Iteration Number: 4784
Loss: 37.399638763606625
12 norm of gradients: 0.2670692970282562
12 norm of weights: 5.937483869537718
-----
Iteration Number: 4785
Loss: 37.39606333603201
12 norm of gradients: 0.2670283386998102
12 norm of weights: 5.937393910441622
-----
Iteration Number: 4786
Loss: 37.39248900044126
12 norm of gradients: 0.2669873997287255
12 norm of weights: 5.937303960757344
-----
Iteration Number: 4787
Loss: 37.38891575615457
12 norm of gradients: 0.2669464800993038
12 norm of weights: 5.937214020482247
-----
Iteration Number: 4788
Loss: 37.385343602484205
12 norm of gradients: 0.26690557979586166
12 norm of weights: 5.937124089613698
-----
Iteration Number: 4789
Loss: 37.381772538779806
12 norm of gradients: 0.26686469880273
12 norm of weights: 5.937034168149064
-----
Iteration Number: 4790
Loss: 37.378202564346566
12 norm of gradients: 0.2668238371042548
12 norm of weights: 5.936944256085714
-----
Iteration Number: 4791
Loss: 37.37463367852541
12 norm of gradients: 0.2667829946847961
12 norm of weights: 5.936854353421019
-----
Iteration Number: 4792

```

Loss: 37.371065880641204
l2 norm of gradients: 0.2667421715287289
l2 norm of weights: 5.936764460152352

Iteration Number: 4793
Loss: 37.36749917002076
l2 norm of gradients: 0.26670136762044255
l2 norm of weights: 5.936674576277087

Iteration Number: 4794
Loss: 37.363933545993824
l2 norm of gradients: 0.266660582944341
l2 norm of weights: 5.936584701792602

Iteration Number: 4795
Loss: 37.360369007886796
l2 norm of gradients: 0.2666198174848427
l2 norm of weights: 5.936494836696275

Iteration Number: 4796
Loss: 37.35680555504525
l2 norm of gradients: 0.26657907122638086
l2 norm of weights: 5.936404980985484

Iteration Number: 4797
Loss: 37.35324318678585
l2 norm of gradients: 0.26653834415340294
l2 norm of weights: 5.936315134657615

Iteration Number: 4798
Loss: 37.34968190243995
l2 norm of gradients: 0.26649763625037104
l2 norm of weights: 5.936225297710051

Iteration Number: 4799
Loss: 37.34612170135155
l2 norm of gradients: 0.2664569475017617
l2 norm of weights: 5.936135470140174

Iteration Number: 4800
Loss: 37.34256258284302
l2 norm of gradients: 0.266416277892066
l2 norm of weights: 5.936045651945376

Iteration Number: 4801
Loss: 37.33900454625368
l2 norm of gradients: 0.26637562740578935
l2 norm of weights: 5.935955843123044

```

-----
Iteration Number: 4802
Loss: 37.335447590910896
l2 norm of gradients: 0.2663349960274519
l2 norm of weights: 5.93586604367057
-----
Iteration Number: 4803
Loss: 37.33189171616796
l2 norm of gradients: 0.26629438374158804
l2 norm of weights: 5.935776253585347
-----
Iteration Number: 4804
Loss: 37.328336921339506
l2 norm of gradients: 0.26625379053274645
l2 norm of weights: 5.93568647286477
-----
Iteration Number: 4805
Loss: 37.324783205768426
l2 norm of gradients: 0.2662132163854905
l2 norm of weights: 5.935596701506234
-----
Iteration Number: 4806
Loss: 37.321230568794334
l2 norm of gradients: 0.26617266128439787
l2 norm of weights: 5.93550693950714
-----
Iteration Number: 4807
Loss: 37.317679009757725
l2 norm of gradients: 0.2661321252140607
l2 norm of weights: 5.935417186864887
-----
Iteration Number: 4808
Loss: 37.31412852799145
l2 norm of gradients: 0.26609160815908545
l2 norm of weights: 5.935327443576877
-----
Iteration Number: 4809
Loss: 37.310579122837055
l2 norm of gradients: 0.2660511101040928
l2 norm of weights: 5.935237709640514
-----
Iteration Number: 4810
Loss: 37.30703079362668
l2 norm of gradients: 0.266010631033718
l2 norm of weights: 5.935147985053202
-----
Iteration Number: 4811
Loss: 37.30348353971509

```

```

12 norm of gradients: 0.2659701709326107
12 norm of weights: 5.935058269812352
-----
Iteration Number: 4812
Loss: 37.299937360436694
12 norm of gradients: 0.26592972978543467
12 norm of weights: 5.934968563915371
-----
Iteration Number: 4813
Loss: 37.2963922551182
12 norm of gradients: 0.2658893075768682
12 norm of weights: 5.93487886735967
-----
Iteration Number: 4814
Loss: 37.29284822312524
12 norm of gradients: 0.26584890429160357
12 norm of weights: 5.934789180142663
-----
Iteration Number: 4815
Loss: 37.28930526378305
12 norm of gradients: 0.2658085199143479
12 norm of weights: 5.934699502261762
-----
Iteration Number: 4816
Loss: 37.285763376442645
12 norm of gradients: 0.26576815442982216
12 norm of weights: 5.934609833714386
-----
Iteration Number: 4817
Loss: 37.2822225604436
12 norm of gradients: 0.2657278078227616
12 norm of weights: 5.934520174497954
-----
Iteration Number: 4818
Loss: 37.2786828151421
12 norm of gradients: 0.2656874800779161
12 norm of weights: 5.934430524609883
-----
Iteration Number: 4819
Loss: 37.27514413986046
12 norm of gradients: 0.2656471711800493
12 norm of weights: 5.9343408840475975
-----
Iteration Number: 4820
Loss: 37.27160653397146
12 norm of gradients: 0.26560688111393954
12 norm of weights: 5.93425125280852
-----

```

Iteration Number: 4821
Loss: 37.268069996798815
l2 norm of gradients: 0.2655666098643791
l2 norm of weights: 5.934161630890074

Iteration Number: 4822
Loss: 37.264534527699425
l2 norm of gradients: 0.26552635741617436
l2 norm of weights: 5.93407201828969

Iteration Number: 4823
Loss: 37.2610001260206
l2 norm of gradients: 0.2654861237541464
l2 norm of weights: 5.933982415004794

Iteration Number: 4824
Loss: 37.25746679111554
l2 norm of gradients: 0.26544590886313
l2 norm of weights: 5.93389282103282

Iteration Number: 4825
Loss: 37.253934522319796
l2 norm of gradients: 0.26540571272797436
l2 norm of weights: 5.933803236371196

Iteration Number: 4826
Loss: 37.25040331898958
l2 norm of gradients: 0.2653655353335428
l2 norm of weights: 5.933713661017359

Iteration Number: 4827
Loss: 37.24687318046976
l2 norm of gradients: 0.26532537666471273
l2 norm of weights: 5.933624094968745

Iteration Number: 4828
Loss: 37.24334410612473
l2 norm of gradients: 0.2652852367063758
l2 norm of weights: 5.9335345382227915

Iteration Number: 4829
Loss: 37.239816095294834
l2 norm of gradients: 0.2652451154434378
l2 norm of weights: 5.9334449907769375

Iteration Number: 4830
Loss: 37.236289147331334
l2 norm of gradients: 0.2652050128608185

```

12 norm of weights: 5.933355452628625
-----
Iteration Number: 4831
Loss: 37.23276326159171
12 norm of gradients: 0.2651649289434519
12 norm of weights: 5.933265923775296
-----
Iteration Number: 4832
Loss: 37.22923843742926
12 norm of gradients: 0.26512486367628607
12 norm of weights: 5.933176404214397
-----
Iteration Number: 4833
Loss: 37.22571467419526
12 norm of gradients: 0.26508481704428316
12 norm of weights: 5.933086893943372
-----
Iteration Number: 4834
Loss: 37.222191971241394
12 norm of gradients: 0.2650447890324194
12 norm of weights: 5.932997392959672
-----
Iteration Number: 4835
Loss: 37.218670327925814
12 norm of gradients: 0.265004779625685
12 norm of weights: 5.932907901260745
-----
Iteration Number: 4836
Loss: 37.215149743592924
12 norm of gradients: 0.26496478880908436
12 norm of weights: 5.932818418844044
-----
Iteration Number: 4837
Loss: 37.21163021761884
12 norm of gradients: 0.26492481656763583
12 norm of weights: 5.932728945707021
-----
Iteration Number: 4838
Loss: 37.208111749345235
12 norm of gradients: 0.26488486288637175
12 norm of weights: 5.932639481847134
-----
Iteration Number: 4839
Loss: 37.20459433813965
12 norm of gradients: 0.2648449277503386
12 norm of weights: 5.932550027261838
-----
Iteration Number: 4840

```

Loss: 37.20107798334766
l2 norm of gradients: 0.2648050111445967
l2 norm of weights: 5.932460581948591

Iteration Number: 4841
Loss: 37.19756268433434
l2 norm of gradients: 0.2647651130542204
l2 norm of weights: 5.932371145904856

Iteration Number: 4842
Loss: 37.1940484404651
l2 norm of gradients: 0.2647252334642982
l2 norm of weights: 5.932281719128094

Iteration Number: 4843
Loss: 37.19053525108088
l2 norm of gradients: 0.2646853723599322
l2 norm of weights: 5.9321923016157685

Iteration Number: 4844
Loss: 37.18702311556626
l2 norm of gradients: 0.264645529726239
l2 norm of weights: 5.932102893365345

Iteration Number: 4845
Loss: 37.18351203326357
l2 norm of gradients: 0.2646057055483486
l2 norm of weights: 5.932013494374293

Iteration Number: 4846
Loss: 37.180002003539585
l2 norm of gradients: 0.2645658998114051
l2 norm of weights: 5.931924104640079

Iteration Number: 4847
Loss: 37.17649302575994
l2 norm of gradients: 0.26452611250056673
l2 norm of weights: 5.931834724160176

Iteration Number: 4848
Loss: 37.172985099288354
l2 norm of gradients: 0.26448634360100537
l2 norm of weights: 5.931745352932056

Iteration Number: 4849
Loss: 37.16947822347336
l2 norm of gradients: 0.26444659309790686
l2 norm of weights: 5.9316559909531925


```

-----
Iteration Number: 4850
Loss: 37.165972397695455
l2 norm of gradients: 0.26440686097647104
l2 norm of weights: 5.931566638221064
-----
Iteration Number: 4851
Loss: 37.16246762131405
l2 norm of gradients: 0.26436714722191146
l2 norm of weights: 5.931477294733145
-----
Iteration Number: 4852
Loss: 37.15896389369185
l2 norm of gradients: 0.26432745181945555
l2 norm of weights: 5.931387960486917
-----
Iteration Number: 4853
Loss: 37.15546121419804
l2 norm of gradients: 0.26428777475434473
l2 norm of weights: 5.93129863547986
-----
Iteration Number: 4854
Loss: 37.15195958219224
l2 norm of gradients: 0.26424811601183407
l2 norm of weights: 5.931209319709458
-----
Iteration Number: 4855
Loss: 37.1484589970496
l2 norm of gradients: 0.26420847557719257
l2 norm of weights: 5.931120013173197
-----
Iteration Number: 4856
Loss: 37.14495945813674
l2 norm of gradients: 0.264168853435703
l2 norm of weights: 5.93103071586856
-----
Iteration Number: 4857
Loss: 37.14146096481354
l2 norm of gradients: 0.264129249572662
l2 norm of weights: 5.930941427793038
-----
Iteration Number: 4858
Loss: 37.137963516453496
l2 norm of gradients: 0.2640896639733799
l2 norm of weights: 5.930852148944119
-----
Iteration Number: 4859
Loss: 37.13446711242406

```

```

12 norm of gradients: 0.26405009662318085
12 norm of weights: 5.930762879319296
-----
Iteration Number: 4860
Loss: 37.13097175209506
12 norm of gradients: 0.2640105475074028
12 norm of weights: 5.930673618916061
-----
Iteration Number: 4861
Loss: 37.12747743484149
12 norm of gradients: 0.2639710166113973
12 norm of weights: 5.9305843677319094
-----
Iteration Number: 4862
Loss: 37.12398416002849
12 norm of gradients: 0.2639315039205299
12 norm of weights: 5.930495125764337
-----
Iteration Number: 4863
Loss: 37.120491927035225
12 norm of gradients: 0.26389200942017965
12 norm of weights: 5.930405893010843
-----
Iteration Number: 4864
Loss: 37.11700073522587
12 norm of gradients: 0.2638525330957394
12 norm of weights: 5.930316669468929
-----
Iteration Number: 4865
Loss: 37.11351058397878
12 norm of gradients: 0.26381307493261574
12 norm of weights: 5.930227455136093
-----
Iteration Number: 4866
Loss: 37.11002147266022
12 norm of gradients: 0.2637736349162289
12 norm of weights: 5.93013825000984
-----
Iteration Number: 4867
Loss: 37.106533400652346
12 norm of gradients: 0.2637342130320128
12 norm of weights: 5.930049054087676
-----
Iteration Number: 4868
Loss: 37.10304636731435
12 norm of gradients: 0.263694809265415
12 norm of weights: 5.929959867367106
-----

```

Iteration Number: 4869
Loss: 37.09956037204111
l2 norm of gradients: 0.2636554236018968
l2 norm of weights: 5.92987068984564

Iteration Number: 4870
Loss: 37.09607541419341
l2 norm of gradients: 0.26361605602693305
l2 norm of weights: 5.929781521520788

Iteration Number: 4871
Loss: 37.09259149316083
l2 norm of gradients: 0.26357670652601245
l2 norm of weights: 5.929692362390061

Iteration Number: 4872
Loss: 37.08910860830489
l2 norm of gradients: 0.263537375084637
l2 norm of weights: 5.9296032124509725

Iteration Number: 4873
Loss: 37.08562675901689
l2 norm of gradients: 0.26349806168832246
l2 norm of weights: 5.929514071701039

Iteration Number: 4874
Loss: 37.082145944659516
l2 norm of gradients: 0.26345876632259835
l2 norm of weights: 5.929424940137776

Iteration Number: 4875
Loss: 37.07866616462716
l2 norm of gradients: 0.2634194889730075
l2 norm of weights: 5.929335817758702

Iteration Number: 4876
Loss: 37.07518741828986
l2 norm of gradients: 0.2633802296251065
l2 norm of weights: 5.929246704561337

Iteration Number: 4877
Loss: 37.07170970502395
l2 norm of gradients: 0.2633409882644654
l2 norm of weights: 5.929157600543204

Iteration Number: 4878
Loss: 37.06823302421849
l2 norm of gradients: 0.26330176487666795

```

12 norm of weights: 5.929068505701825
-----
Iteration Number: 4879
Loss: 37.06475737525507
12 norm of gradients: 0.2632625594473113
12 norm of weights: 5.928979420034728
-----
Iteration Number: 4880
Loss: 37.06128275750536
12 norm of gradients: 0.2632233719620061
12 norm of weights: 5.928890343539437
-----
Iteration Number: 4881
Loss: 37.05780917036012
12 norm of gradients: 0.26318420240637674
12 norm of weights: 5.9288012762134805
-----
Iteration Number: 4882
Loss: 37.05433661319387
12 norm of gradients: 0.2631450507660609
12 norm of weights: 5.928712218054391
-----
Iteration Number: 4883
Loss: 37.05086508539294
12 norm of gradients: 0.2631059170267099
12 norm of weights: 5.928623169059698
-----
Iteration Number: 4884
Loss: 37.04739458634072
12 norm of gradients: 0.26306680117398823
12 norm of weights: 5.928534129226937
-----
Iteration Number: 4885
Loss: 37.04392511542703
12 norm of gradients: 0.2630277031935743
12 norm of weights: 5.928445098553641
-----
Iteration Number: 4886
Loss: 37.04045667203122
12 norm of gradients: 0.26298862307115983
12 norm of weights: 5.928356077037348
-----
Iteration Number: 4887
Loss: 37.03698925554509
12 norm of gradients: 0.26294956079244985
12 norm of weights: 5.928267064675596
-----
Iteration Number: 4888

```

Loss: 37.03352286533701
l2 norm of gradients: 0.26291051634316287
l2 norm of weights: 5.9281780614659265

Iteration Number: 4889
Loss: 37.03005750081109
l2 norm of gradients: 0.2628714897090309
l2 norm of weights: 5.92808906740588

Iteration Number: 4890
Loss: 37.02659316134768
l2 norm of gradients: 0.2628324808757993
l2 norm of weights: 5.928000082493

Iteration Number: 4891
Loss: 37.02312984633906
l2 norm of gradients: 0.26279348982922696
l2 norm of weights: 5.927911106724832

Iteration Number: 4892
Loss: 37.01966755517
l2 norm of gradients: 0.262754516555086
l2 norm of weights: 5.927822140098921

Iteration Number: 4893
Loss: 37.016206287224435
l2 norm of gradients: 0.262715561039162
l2 norm of weights: 5.927733182612818

Iteration Number: 4894
Loss: 37.01274604189986
l2 norm of gradients: 0.2626766232672538
l2 norm of weights: 5.927644234264071

Iteration Number: 4895
Loss: 37.00928681858715
l2 norm of gradients: 0.2626377032251739
l2 norm of weights: 5.927555295050232

Iteration Number: 4896
Loss: 37.00582861666243
l2 norm of gradients: 0.2625988008987477
l2 norm of weights: 5.927466364968855

Iteration Number: 4897
Loss: 37.00237143553234
l2 norm of gradients: 0.2625599162738143
l2 norm of weights: 5.927377444017495

```

-----
Iteration Number: 4898
Loss: 36.99891527458486
l2 norm of gradients: 0.26252104933622605
l2 norm of weights: 5.927288532193706
-----
Iteration Number: 4899
Loss: 36.995460133211914
l2 norm of gradients: 0.26248220007184836
l2 norm of weights: 5.92719962949505
-----
Iteration Number: 4900
Loss: 36.99200601079551
l2 norm of gradients: 0.2624433684665603
l2 norm of weights: 5.927110735919084
-----
Iteration Number: 4901
Loss: 36.988552906738114
l2 norm of gradients: 0.262404554506254
l2 norm of weights: 5.927021851463369
-----
Iteration Number: 4902
Loss: 36.985100820435086
l2 norm of gradients: 0.26236575817683505
l2 norm of weights: 5.926932976125471
-----
Iteration Number: 4903
Loss: 36.98164975128431
l2 norm of gradients: 0.2623269794642219
l2 norm of weights: 5.926844109902952
-----
Iteration Number: 4904
Loss: 36.978199698672036
l2 norm of gradients: 0.2622882183543468
l2 norm of weights: 5.926755252793378
-----
Iteration Number: 4905
Loss: 36.974750661998115
l2 norm of gradients: 0.26224947483315486
l2 norm of weights: 5.92666640479432
-----
Iteration Number: 4906
Loss: 36.97130264066265
l2 norm of gradients: 0.2622107488866045
l2 norm of weights: 5.926577565903345
-----
Iteration Number: 4907
Loss: 36.96785563405389

```

```

12 norm of gradients: 0.2621720405006675
12 norm of weights: 5.9264887361180225
-----
Iteration Number: 4908
Loss: 36.96440964156766
12 norm of gradients: 0.26213334966132873
12 norm of weights: 5.926399915435929
-----
Iteration Number: 4909
Loss: 36.960964662607964
12 norm of gradients: 0.26209467635458616
12 norm of weights: 5.926311103854637
-----
Iteration Number: 4910
Loss: 36.95752069657373
12 norm of gradients: 0.26205602056645116
12 norm of weights: 5.926222301371722
-----
Iteration Number: 4911
Loss: 36.95407774286038
12 norm of gradients: 0.26201738228294813
12 norm of weights: 5.926133507984762
-----
Iteration Number: 4912
Loss: 36.950635800869875
12 norm of gradients: 0.26197876149011473
12 norm of weights: 5.926044723691337
-----
Iteration Number: 4913
Loss: 36.947194870001645
12 norm of gradients: 0.2619401581740016
12 norm of weights: 5.925955948489028
-----
Iteration Number: 4914
Loss: 36.94375494965474
12 norm of gradients: 0.26190157232067274
12 norm of weights: 5.9258671823754145
-----
Iteration Number: 4915
Loss: 36.94031603922614
12 norm of gradients: 0.2618630039162051
12 norm of weights: 5.925778425348082
-----
Iteration Number: 4916
Loss: 36.93687813812835
12 norm of gradients: 0.2618244529466889
12 norm of weights: 5.925689677404618
-----

```

Iteration Number: 4917
Loss: 36.933441245762026
l2 norm of gradients: 0.26178591939822726
l2 norm of weights: 5.925600938542607

Iteration Number: 4918
Loss: 36.930005361518994
l2 norm of gradients: 0.2617474032569366
l2 norm of weights: 5.92551220875964

Iteration Number: 4919
Loss: 36.92657048481096
l2 norm of gradients: 0.26170890450894635
l2 norm of weights: 5.925423488053306

Iteration Number: 4920
Loss: 36.92313661503663
l2 norm of gradients: 0.26167042314039884
l2 norm of weights: 5.925334776421197

Iteration Number: 4921
Loss: 36.91970375160317
l2 norm of gradients: 0.26163195913744985
l2 norm of weights: 5.925246073860906

Iteration Number: 4922
Loss: 36.91627189392098
l2 norm of gradients: 0.2615935124862679
l2 norm of weights: 5.925157380370029

Iteration Number: 4923
Loss: 36.91284104139068
l2 norm of gradients: 0.2615550831730345
l2 norm of weights: 5.925068695946163

Iteration Number: 4924
Loss: 36.90941119341323
l2 norm of gradients: 0.2615166711839445
l2 norm of weights: 5.924980020586906

Iteration Number: 4925
Loss: 36.905982349402926
l2 norm of gradients: 0.26147827650520555
l2 norm of weights: 5.924891354289858

Iteration Number: 4926
Loss: 36.90255450876367
l2 norm of gradients: 0.26143989912303817


```

12 norm of weights: 5.92480269705262
-----
Iteration Number: 4927
Loss: 36.89912767090029
12 norm of gradients: 0.2614015390236761
12 norm of weights: 5.9247140488727945
-----
Iteration Number: 4928
Loss: 36.895701835233766
12 norm of gradients: 0.26136319619336607
12 norm of weights: 5.924625409747987
-----
Iteration Number: 4929
Loss: 36.89227700114863
12 norm of gradients: 0.2613248706183677
12 norm of weights: 5.924536779675804
-----
Iteration Number: 4930
Loss: 36.888853168077915
12 norm of gradients: 0.2612865622849534
12 norm of weights: 5.924448158653853
-----
Iteration Number: 4931
Loss: 36.88543033542189
12 norm of gradients: 0.2612482711794089
12 norm of weights: 5.924359546679743
-----
Iteration Number: 4932
Loss: 36.882008502590416
12 norm of gradients: 0.26120999728803257
12 norm of weights: 5.924270943751085
-----
Iteration Number: 4933
Loss: 36.87858766898857
12 norm of gradients: 0.2611717405971358
12 norm of weights: 5.924182349865492
-----
Iteration Number: 4934
Loss: 36.87516783404783
12 norm of gradients: 0.26113350109304295
12 norm of weights: 5.924093765020578
-----
Iteration Number: 4935
Loss: 36.8717489971566
12 norm of gradients: 0.2610952787620911
12 norm of weights: 5.924005189213957
-----
Iteration Number: 4936

```

Loss: 36.868331157739284
l2 norm of gradients: 0.26105707359063046
l2 norm of weights: 5.923916622443248

Iteration Number: 4937
Loss: 36.86491431521596
l2 norm of gradients: 0.261018885565024
l2 norm of weights: 5.923828064706069

Iteration Number: 4938
Loss: 36.86149846897804
l2 norm of gradients: 0.26098071467164746
l2 norm of weights: 5.923739516000041

Iteration Number: 4939
Loss: 36.85808361846834
l2 norm of gradients: 0.26094256089688966
l2 norm of weights: 5.923650976322786

Iteration Number: 4940
Loss: 36.854669763075336
l2 norm of gradients: 0.26090442422715204
l2 norm of weights: 5.9235624456719265

Iteration Number: 4941
Loss: 36.851256902228215
l2 norm of gradients: 0.2608663046488491
l2 norm of weights: 5.923473924045087

Iteration Number: 4942
Loss: 36.84784503534378
l2 norm of gradients: 0.2608282021484078
l2 norm of weights: 5.923385411439896

Iteration Number: 4943
Loss: 36.844434161830385
l2 norm of gradients: 0.2607901167122684
l2 norm of weights: 5.92329690785398

Iteration Number: 4944
Loss: 36.84102428111235
l2 norm of gradients: 0.2607520483268837
l2 norm of weights: 5.923208413284969

Iteration Number: 4945
Loss: 36.837615392596426
l2 norm of gradients: 0.2607139969787192
l2 norm of weights: 5.923119927730494

```

-----
Iteration Number: 4946
Loss: 36.834207495714665
l2 norm of gradients: 0.2606759626542534
l2 norm of weights: 5.923031451188187
-----
Iteration Number: 4947
Loss: 36.83080058987922
l2 norm of gradients: 0.2606379453399773
l2 norm of weights: 5.922942983655686
-----
Iteration Number: 4948
Loss: 36.82739467449876
l2 norm of gradients: 0.2605999450223949
l2 norm of weights: 5.922854525130622
-----
Iteration Number: 4949
Loss: 36.823989749017144
l2 norm of gradients: 0.2605619616880228
l2 norm of weights: 5.922766075610635
-----
Iteration Number: 4950
Loss: 36.82058581283111
l2 norm of gradients: 0.2605239953233906
l2 norm of weights: 5.922677635093364
-----
Iteration Number: 4951
Loss: 36.81718286536759
l2 norm of gradients: 0.26048604591504015
l2 norm of weights: 5.922589203576448
-----
Iteration Number: 4952
Loss: 36.81378090605558
l2 norm of gradients: 0.2604481134495263
l2 norm of weights: 5.922500781057531
-----
Iteration Number: 4953
Loss: 36.81037993431321
l2 norm of gradients: 0.26041019791341674
l2 norm of weights: 5.922412367534253
-----
Iteration Number: 4954
Loss: 36.80697994955636
l2 norm of gradients: 0.2603722992932916
l2 norm of weights: 5.9223239630042634
-----
Iteration Number: 4955
Loss: 36.80358095120392

```

```

12 norm of gradients: 0.26033441757574366
12 norm of weights: 5.922235567465207
-----
Iteration Number: 4956
Loss: 36.80018293870122
12 norm of gradients: 0.2602965527473787
12 norm of weights: 5.92214718091473
-----
Iteration Number: 4957
Loss: 36.79678591144619
12 norm of gradients: 0.2602587047948148
12 norm of weights: 5.922058803350485
-----
Iteration Number: 4958
Loss: 36.793389868883416
12 norm of gradients: 0.2602208737046828
12 norm of weights: 5.921970434770121
-----
Iteration Number: 4959
Loss: 36.78999481042225
12 norm of gradients: 0.26018305946362624
12 norm of weights: 5.921882075171292
-----
Iteration Number: 4960
Loss: 36.78660073550334
12 norm of gradients: 0.2601452620583012
12 norm of weights: 5.921793724551652
-----
Iteration Number: 4961
Loss: 36.78320764353719
12 norm of gradients: 0.26010748147537655
12 norm of weights: 5.921705382908857
-----
Iteration Number: 4962
Loss: 36.779815533955826
12 norm of gradients: 0.2600697177015335
12 norm of weights: 5.921617050240563
-----
Iteration Number: 4963
Loss: 36.77642440619219
12 norm of gradients: 0.26003197072346607
12 norm of weights: 5.921528726544429
-----
Iteration Number: 4964
Loss: 36.77303425966396
12 norm of gradients: 0.2599942405278805
12 norm of weights: 5.921440411818116
-----

```

Iteration Number: 4965
Loss: 36.769645093809835
l2 norm of gradients: 0.2599565271014962
l2 norm of weights: 5.921352106059286

Iteration Number: 4966
Loss: 36.766256908048476
l2 norm of gradients: 0.25991883043104463
l2 norm of weights: 5.921263809265603

Iteration Number: 4967
Loss: 36.76286970180546
l2 norm of gradients: 0.25988115050326993
l2 norm of weights: 5.921175521434729

Iteration Number: 4968
Loss: 36.759483474530626
l2 norm of gradients: 0.25984348730492896
l2 norm of weights: 5.921087242564333

Iteration Number: 4969
Loss: 36.75609822563334
l2 norm of gradients: 0.2598058408227908
l2 norm of weights: 5.920998972652082

Iteration Number: 4970
Loss: 36.752713954554665
l2 norm of gradients: 0.25976821104363723
l2 norm of weights: 5.920910711695646

Iteration Number: 4971
Loss: 36.74933066071574
l2 norm of gradients: 0.25973059795426257
l2 norm of weights: 5.920822459692694

Iteration Number: 4972
Loss: 36.74594834356696
l2 norm of gradients: 0.25969300154147346
l2 norm of weights: 5.9207342166409

Iteration Number: 4973
Loss: 36.742567002515315
l2 norm of gradients: 0.25965542179208917
l2 norm of weights: 5.920645982537937

Iteration Number: 4974
Loss: 36.73918663701122
l2 norm of gradients: 0.2596178586929413

```

12 norm of weights: 5.920557757381482
-----
Iteration Number: 4975
Loss: 36.735807246486566
12 norm of gradients: 0.2595803122308742
12 norm of weights: 5.9204695411692105
-----
Iteration Number: 4976
Loss: 36.73242883036625
12 norm of gradients: 0.25954278239274425
12 norm of weights: 5.9203813338988
-----
Iteration Number: 4977
Loss: 36.72905138809141
12 norm of gradients: 0.25950526916542066
12 norm of weights: 5.920293135567931
-----
Iteration Number: 4978
Loss: 36.7256749190947
12 norm of gradients: 0.2594677725357847
12 norm of weights: 5.9202049461742865
-----
Iteration Number: 4979
Loss: 36.72229942281043
12 norm of gradients: 0.2594302924907303
12 norm of weights: 5.920116765715548
-----
Iteration Number: 4980
Loss: 36.71892489867704
12 norm of gradients: 0.25939282901716376
12 norm of weights: 5.920028594189398
-----
Iteration Number: 4981
Loss: 36.715551346133076
12 norm of gradients: 0.2593553821020037
12 norm of weights: 5.919940431593526
-----
Iteration Number: 4982
Loss: 36.71217876460636
12 norm of gradients: 0.25931795173218136
12 norm of weights: 5.919852277925618
-----
Iteration Number: 4983
Loss: 36.7088071535408
12 norm of gradients: 0.25928053789463984
12 norm of weights: 5.919764133183361
-----
Iteration Number: 4984

```

Loss: 36.705436512369964
l2 norm of gradients: 0.2592431405763351
l2 norm of weights: 5.919675997364448

Iteration Number: 4985
Loss: 36.70206684053738
l2 norm of gradients: 0.2592057597642352
l2 norm of weights: 5.919587870466567

Iteration Number: 4986
Loss: 36.69869813747703
l2 norm of gradients: 0.2591683954453206
l2 norm of weights: 5.919499752487415

Iteration Number: 4987
Loss: 36.69533040263474
l2 norm of gradients: 0.2591310476065842
l2 norm of weights: 5.919411643424685

Iteration Number: 4988
Loss: 36.6919636354375
l2 norm of gradients: 0.2590937162350309
l2 norm of weights: 5.919323543276073

Iteration Number: 4989
Loss: 36.68859783534028
l2 norm of gradients: 0.2590564013176783
l2 norm of weights: 5.9192354520392785

Iteration Number: 4990
Loss: 36.685233001770946
l2 norm of gradients: 0.259019102841556
l2 norm of weights: 5.919147369711998

Iteration Number: 4991
Loss: 36.68186913418064
l2 norm of gradients: 0.258981820793706
l2 norm of weights: 5.919059296291934

Iteration Number: 4992
Loss: 36.67850623201012
l2 norm of gradients: 0.2589445551611825
l2 norm of weights: 5.918971231776788

Iteration Number: 4993
Loss: 36.675144294690824
l2 norm of gradients: 0.2589073059310521
l2 norm of weights: 5.918883176164263

```

-----
Iteration Number: 4994
Loss: 36.67178332168714
l2 norm of gradients: 0.25887007309039356
l2 norm of weights: 5.918795129452064
-----
Iteration Number: 4995
Loss: 36.668423312415044
l2 norm of gradients: 0.25883285662629785
l2 norm of weights: 5.918707091637899
-----
Iteration Number: 4996
Loss: 36.66506426634173
l2 norm of gradients: 0.2587956565258683
l2 norm of weights: 5.9186190627194755
-----
Iteration Number: 4997
Loss: 36.661706182906165
l2 norm of gradients: 0.25875847277622027
l2 norm of weights: 5.918531042694501
-----
Iteration Number: 4998
Loss: 36.65834906153746
l2 norm of gradients: 0.2587213053644814
l2 norm of weights: 5.918443031560688
-----
Iteration Number: 4999
Loss: 36.654992901700034
l2 norm of gradients: 0.2586841542777917
l2 norm of weights: 5.918355029315749
-----
Iteration Number: 5000
Loss: 36.6516377028298
l2 norm of gradients: 0.258647019503303
l2 norm of weights: 5.918267035957397
-----
Iteration Number: 5001
Loss: 36.648283464379425
l2 norm of gradients: 0.25860990102817966
l2 norm of weights: 5.918179051483348
-----
Iteration Number: 5002
Loss: 36.64493018578912
l2 norm of gradients: 0.25857279883959805
l2 norm of weights: 5.918091075891318
-----
Iteration Number: 5003
Loss: 36.64157786650906

```



```

12 norm of gradients: 0.25853571292474675
12 norm of weights: 5.9180031091790255
-----
Iteration Number: 5004
Loss: 36.63822650598788
12 norm of gradients: 0.2584986432708263
12 norm of weights: 5.91791515134419
-----
Iteration Number: 5005
Loss: 36.634876103673335
12 norm of gradients: 0.2584615898650497
12 norm of weights: 5.917827202384532
-----
Iteration Number: 5006
Loss: 36.6315266590138
12 norm of gradients: 0.2584245526946417
12 norm of weights: 5.9177392622977765
-----
Iteration Number: 5007
Loss: 36.62817817145784
12 norm of gradients: 0.2583875317468395
12 norm of weights: 5.917651331081644
-----
Iteration Number: 5008
Loss: 36.62483064045779
12 norm of gradients: 0.2583505270088921
12 norm of weights: 5.917563408733862
-----
Iteration Number: 5009
Loss: 36.62148406546596
12 norm of gradients: 0.2583135384680608
12 norm of weights: 5.917475495252157
-----
Iteration Number: 5010
Loss: 36.61813844592696
12 norm of gradients: 0.2582765661116189
12 norm of weights: 5.917387590634256
-----
Iteration Number: 5011
Loss: 36.614793781290174
12 norm of gradients: 0.2582396099268518
12 norm of weights: 5.91729969487789
-----
Iteration Number: 5012
Loss: 36.611450071023754
12 norm of gradients: 0.258202669901057
12 norm of weights: 5.917211807980791
-----

```

Iteration Number: 5013
Loss: 36.608107314565586
l2 norm of gradients: 0.2581657460215438
l2 norm of weights: 5.91712392994069

Iteration Number: 5014
Loss: 36.60476551136184
l2 norm of gradients: 0.2581288382756338
l2 norm of weights: 5.917036060755321

Iteration Number: 5015
Loss: 36.60142466088773
l2 norm of gradients: 0.25809194665066054
l2 norm of weights: 5.9169482004224205

Iteration Number: 5016
Loss: 36.59808476257703
l2 norm of gradients: 0.2580550711339697
l2 norm of weights: 5.916860348939725

Iteration Number: 5017
Loss: 36.594745815885624
l2 norm of gradients: 0.2580182117129186
l2 norm of weights: 5.916772506304972

Iteration Number: 5018
Loss: 36.59140782029083
l2 norm of gradients: 0.25798136837487695
l2 norm of weights: 5.9166846725159035

Iteration Number: 5019
Loss: 36.5880707752184
l2 norm of gradients: 0.25794454110722626
l2 norm of weights: 5.916596847570257

Iteration Number: 5020
Loss: 36.5847346801425
l2 norm of gradients: 0.25790772989735994
l2 norm of weights: 5.916509031465779

Iteration Number: 5021
Loss: 36.58139953451012
l2 norm of gradients: 0.2578709347326836
l2 norm of weights: 5.916421224200211

Iteration Number: 5022
Loss: 36.57806533778642
l2 norm of gradients: 0.25783415560061446

```
12 norm of weights: 5.9163334257713
-----
Iteration Number: 5023
Loss: 36.574732089428025
12 norm of gradients: 0.257797392488582
12 norm of weights: 5.916245636176792
-----
Iteration Number: 5024
Loss: 36.57139978888161
12 norm of gradients: 0.2577606453840275
12 norm of weights: 5.916157855414435
-----
Iteration Number: 5025
Loss: 36.56806843561459
12 norm of gradients: 0.25772391427440405
12 norm of weights: 5.916070083481979
-----
Iteration Number: 5026
Loss: 36.56473802908329
12 norm of gradients: 0.25768719914717686
12 norm of weights: 5.915982320377176
-----
Iteration Number: 5027
Loss: 36.56140856874367
12 norm of gradients: 0.2576504999898229
12 norm of weights: 5.915894566097777
-----
Iteration Number: 5028
Loss: 36.55808005406774
12 norm of gradients: 0.25761381678983103
12 norm of weights: 5.915806820641538
-----
Iteration Number: 5029
Loss: 36.55475248449972
12 norm of gradients: 0.25757714953470195
12 norm of weights: 5.915719084006212
-----
Iteration Number: 5030
Loss: 36.55142585950978
12 norm of gradients: 0.25754049821194847
12 norm of weights: 5.915631356189557
-----
Iteration Number: 5031
Loss: 36.548100178556
12 norm of gradients: 0.2575038628090949
12 norm of weights: 5.915543637189331
-----
Iteration Number: 5032
```

```

Loss: 36.54477544110097
12 norm of gradients: 0.25746724331367754
12 norm of weights: 5.915455927003294
-----
Iteration Number: 5033
Loss: 36.54145164660765
12 norm of gradients: 0.2574306397132446
12 norm of weights: 5.915368225629208
-----
Iteration Number: 5034
Loss: 36.53812879453304
12 norm of gradients: 0.25739405199535614
12 norm of weights: 5.915280533064833
-----
Iteration Number: 5035
Loss: 36.534806884349734
12 norm of gradients: 0.25735748014758386
12 norm of weights: 5.9151928493079335
-----
Iteration Number: 5036
Loss: 36.53148591551422
12 norm of gradients: 0.25732092415751123
12 norm of weights: 5.915105174356276
-----
Iteration Number: 5037
Loss: 36.52816588749543
12 norm of gradients: 0.25728438401273385
12 norm of weights: 5.915017508207626
-----
Iteration Number: 5038
Loss: 36.52484679974895
12 norm of gradients: 0.25724785970085867
12 norm of weights: 5.914929850859752
-----
Iteration Number: 5039
Loss: 36.52152865174579
12 norm of gradients: 0.2572113512095048
12 norm of weights: 5.914842202310424
-----
Iteration Number: 5040
Loss: 36.51821144295484
12 norm of gradients: 0.25717485852630284
12 norm of weights: 5.914754562557411
-----
Iteration Number: 5041
Loss: 36.51489517283771
12 norm of gradients: 0.2571383816388952
12 norm of weights: 5.914666931598486

```

```

-----
Iteration Number: 5042
Loss: 36.51157984086151
l2 norm of gradients: 0.2571019205349361
l2 norm of weights: 5.9145793094314225
-----
Iteration Number: 5043
Loss: 36.50826544648779
l2 norm of gradients: 0.25706547520209133
l2 norm of weights: 5.914491696053997
-----
Iteration Number: 5044
Loss: 36.50495198919863
l2 norm of gradients: 0.2570290456280386
l2 norm of weights: 5.914404091463984
-----
Iteration Number: 5045
Loss: 36.50163946844217
l2 norm of gradients: 0.2569926318004672
l2 norm of weights: 5.914316495659162
-----
Iteration Number: 5046
Loss: 36.49832788370496
l2 norm of gradients: 0.2569562337070782
l2 norm of weights: 5.91422890863731
-----
Iteration Number: 5047
Loss: 36.495017234448014
l2 norm of gradients: 0.2569198513355842
l2 norm of weights: 5.91414133039621
-----
Iteration Number: 5048
Loss: 36.491707520135016
l2 norm of gradients: 0.25688348467370953
l2 norm of weights: 5.9140537609336405
-----
Iteration Number: 5049
Loss: 36.48839874024273
l2 norm of gradients: 0.2568471337091905
l2 norm of weights: 5.913966200247388
-----
Iteration Number: 5050
Loss: 36.48509089425032
l2 norm of gradients: 0.2568107984297745
l2 norm of weights: 5.913878648335235
-----
Iteration Number: 5051
Loss: 36.481783981605076

```

```

12 norm of gradients: 0.25677447882322096
12 norm of weights: 5.91379110519497
-----
Iteration Number: 5052
Loss: 36.478478001795125
12 norm of gradients: 0.25673817487730094
12 norm of weights: 5.913703570824378
-----
Iteration Number: 5053
Loss: 36.47517295428819
12 norm of gradients: 0.256701886579797
12 norm of weights: 5.913616045221248
-----
Iteration Number: 5054
Loss: 36.47186883855595
12 norm of gradients: 0.2566656139185032
12 norm of weights: 5.913528528383373
-----
Iteration Number: 5055
Loss: 36.46856565407341
12 norm of gradients: 0.2566293568812254
12 norm of weights: 5.913441020308541
-----
Iteration Number: 5056
Loss: 36.46526340031048
12 norm of gradients: 0.25659311545578106
12 norm of weights: 5.913353520994547
-----
Iteration Number: 5057
Loss: 36.46196207674698
12 norm of gradients: 0.2565568896299991
12 norm of weights: 5.913266030439185
-----
Iteration Number: 5058
Loss: 36.45866168283909
12 norm of gradients: 0.25652067939172013
12 norm of weights: 5.91317854864025
-----
Iteration Number: 5059
Loss: 36.45536221808215
12 norm of gradients: 0.2564844847287962
12 norm of weights: 5.913091075595539
-----
Iteration Number: 5060
Loss: 36.45206368194657
12 norm of gradients: 0.2564483056290909
12 norm of weights: 5.913003611302852
-----

```

Iteration Number: 5061
Loss: 36.44876607389591
l2 norm of gradients: 0.25641214208047947
l2 norm of weights: 5.912916155759986

Iteration Number: 5062
Loss: 36.445469393423
l2 norm of gradients: 0.2563759940708487
l2 norm of weights: 5.912828708964746

Iteration Number: 5063
Loss: 36.44217363998667
l2 norm of gradients: 0.2563398615880967
l2 norm of weights: 5.91274127091493

Iteration Number: 5064
Loss: 36.43887881308187
l2 norm of gradients: 0.2563037446201332
l2 norm of weights: 5.9126538416083445

Iteration Number: 5065
Loss: 36.43558491216905
l2 norm of gradients: 0.2562676431548796
l2 norm of weights: 5.912566421042794

Iteration Number: 5066
Loss: 36.43229193672773
l2 norm of gradients: 0.25623155718026847
l2 norm of weights: 5.912479009216087

Iteration Number: 5067
Loss: 36.42899988625185
l2 norm of gradients: 0.2561954866842441
l2 norm of weights: 5.912391606126028

Iteration Number: 5068
Loss: 36.425708760207
l2 norm of gradients: 0.25615943165476207
l2 norm of weights: 5.9123042117704285

Iteration Number: 5069
Loss: 36.422418558079386
l2 norm of gradients: 0.25612339207978957
l2 norm of weights: 5.912216826147098

Iteration Number: 5070
Loss: 36.41912927933851
l2 norm of gradients: 0.25608736794730513

```

12 norm of weights: 5.91212944925385
-----
Iteration Number: 5071
Loss: 36.41584092347542
12 norm of gradients: 0.25605135924529876
12 norm of weights: 5.912042081088497
-----
Iteration Number: 5072
Loss: 36.412553489961006
12 norm of gradients: 0.25601536596177193
12 norm of weights: 5.911954721648853
-----
Iteration Number: 5073
Loss: 36.40926697828465
12 norm of gradients: 0.25597938808473747
12 norm of weights: 5.911867370932735
-----
Iteration Number: 5074
Loss: 36.405981387922694
12 norm of gradients: 0.2559434256022195
12 norm of weights: 5.91178002893796
-----
Iteration Number: 5075
Loss: 36.40269671835824
12 norm of gradients: 0.25590747850225376
12 norm of weights: 5.911692695662346
-----
Iteration Number: 5076
Loss: 36.399412969080025
12 norm of gradients: 0.2558715467728873
12 norm of weights: 5.911605371103714
-----
Iteration Number: 5077
Loss: 36.39613013955079
12 norm of gradients: 0.25583563040217844
12 norm of weights: 5.911518055259885
-----
Iteration Number: 5078
Loss: 36.392848229278506
12 norm of gradients: 0.25579972937819695
12 norm of weights: 5.911430748128682
-----
Iteration Number: 5079
Loss: 36.38956723773519
12 norm of gradients: 0.25576384368902394
12 norm of weights: 5.911343449707928
-----
Iteration Number: 5080

```


Loss: 36.38628716440571
l2 norm of gradients: 0.25572797332275193
l2 norm of weights: 5.91125615999545

Iteration Number: 5081
Loss: 36.383008008779356
l2 norm of gradients: 0.2556921182674845
l2 norm of weights: 5.9111688789890735

Iteration Number: 5082
Loss: 36.3797297703297
l2 norm of gradients: 0.255656278511337
l2 norm of weights: 5.9110816066866265

Iteration Number: 5083
Loss: 36.37645244855599
l2 norm of gradients: 0.2556204540424357
l2 norm of weights: 5.91099434308594

Iteration Number: 5084
Loss: 36.37317604293269
l2 norm of gradients: 0.2555846448489184
l2 norm of weights: 5.910907088184843

Iteration Number: 5085
Loss: 36.36990055294793
l2 norm of gradients: 0.2555488509189339
l2 norm of weights: 5.910819841981169

Iteration Number: 5086
Loss: 36.366625978100636
l2 norm of gradients: 0.2555130722406427
l2 norm of weights: 5.9107326044727495

Iteration Number: 5087
Loss: 36.36335231786679
l2 norm of gradients: 0.25547730880221625
l2 norm of weights: 5.91064537565742

Iteration Number: 5088
Loss: 36.36007957173599
l2 norm of gradients: 0.2554415605918374
l2 norm of weights: 5.910558155533018

Iteration Number: 5089
Loss: 36.356807739199844
l2 norm of gradients: 0.2554058275977002
l2 norm of weights: 5.91047094409738

```

-----
Iteration Number: 5090
Loss: 36.353536819748356
l2 norm of gradients: 0.25537010980800995
l2 norm of weights: 5.910383741348343
-----
Iteration Number: 5091
Loss: 36.3502668128565
l2 norm of gradients: 0.2553344072109831
l2 norm of weights: 5.910296547283749
-----
Iteration Number: 5092
Loss: 36.34699771803779
l2 norm of gradients: 0.2552987197948476
l2 norm of weights: 5.910209361901438
-----
Iteration Number: 5093
Loss: 36.34372953476163
l2 norm of gradients: 0.2552630475478422
l2 norm of weights: 5.910122185199254
-----
Iteration Number: 5094
Loss: 36.3404622625333
l2 norm of gradients: 0.25522739045821713
l2 norm of weights: 5.91003501717504
-----
Iteration Number: 5095
Loss: 36.33719590083029
l2 norm of gradients: 0.2551917485142337
l2 norm of weights: 5.909947857826642
-----
Iteration Number: 5096
Loss: 36.33393044915304
l2 norm of gradients: 0.2551561217041645
l2 norm of weights: 5.909860707151905
-----
Iteration Number: 5097
Loss: 36.33066590699449
l2 norm of gradients: 0.2551205100162932
l2 norm of weights: 5.909773565148678
-----
Iteration Number: 5098
Loss: 36.327402273845614
l2 norm of gradients: 0.25508491343891465
l2 norm of weights: 5.90968643181481
-----
Iteration Number: 5099
Loss: 36.32413954918821

```

```

12 norm of gradients: 0.2550493319603348
12 norm of weights: 5.909599307148153
-----
Iteration Number: 5100
Loss: 36.32087773253157
12 norm of gradients: 0.2550137655688708
12 norm of weights: 5.909512191146557
-----
Iteration Number: 5101
Loss: 36.31761682336419
12 norm of gradients: 0.2549782142528509
12 norm of weights: 5.909425083807875
-----
Iteration Number: 5102
Loss: 36.31435682117654
12 norm of gradients: 0.2549426780006145
12 norm of weights: 5.909337985129962
-----
Iteration Number: 5103
Loss: 36.31109772546877
12 norm of gradients: 0.25490715680051207
12 norm of weights: 5.9092508951106755
-----
Iteration Number: 5104
Loss: 36.30783953572758
12 norm of gradients: 0.25487165064090517
12 norm of weights: 5.909163813747869
-----
Iteration Number: 5105
Loss: 36.30458225145569
12 norm of gradients: 0.2548361595101664
12 norm of weights: 5.909076741039403
-----
Iteration Number: 5106
Loss: 36.301325872150144
12 norm of gradients: 0.2548006833966797
12 norm of weights: 5.908989676983136
-----
Iteration Number: 5107
Loss: 36.29807039729826
12 norm of gradients: 0.25476522228883974
12 norm of weights: 5.908902621576932
-----
Iteration Number: 5108
Loss: 36.294815826404516
12 norm of gradients: 0.2547297761750524
12 norm of weights: 5.908815574818649
-----

```

Iteration Number: 5109
Loss: 36.291562158966556
l2 norm of gradients: 0.2546943450437345
l2 norm of weights: 5.908728536706152

Iteration Number: 5110
Loss: 36.28830939447676
l2 norm of gradients: 0.25465892888331415
l2 norm of weights: 5.9086415072373075

Iteration Number: 5111
Loss: 36.28505753244384
l2 norm of gradients: 0.2546235276822302
l2 norm of weights: 5.90855448640998

Iteration Number: 5112
Loss: 36.28180657234563
l2 norm of gradients: 0.25458814142893266
l2 norm of weights: 5.908467474222038

Iteration Number: 5113
Loss: 36.27855651371015
l2 norm of gradients: 0.2545527701118826
l2 norm of weights: 5.908380470671348

Iteration Number: 5114
Loss: 36.27530735601044
l2 norm of gradients: 0.2545174137195519
l2 norm of weights: 5.908293475755784

Iteration Number: 5115
Loss: 36.27205909876564
l2 norm of gradients: 0.2544820722404235
l2 norm of weights: 5.908206489473212

Iteration Number: 5116
Loss: 36.268811741471026
l2 norm of gradients: 0.2544467456629916
l2 norm of weights: 5.908119511821509

Iteration Number: 5117
Loss: 36.265565283614116
l2 norm of gradients: 0.25441143397576077
l2 norm of weights: 5.908032542798546

Iteration Number: 5118
Loss: 36.26231972470797
l2 norm of gradients: 0.25437613716724716

```

12 norm of weights: 5.9079455824022
-----
Iteration Number: 5119
Loss: 36.25907506425774
12 norm of gradients: 0.25434085522597744
12 norm of weights: 5.907858630630346
-----
Iteration Number: 5120
Loss: 36.255831301757
12 norm of gradients: 0.25430558814048937
12 norm of weights: 5.907771687480863
-----
Iteration Number: 5121
Loss: 36.25258843671609
12 norm of gradients: 0.2542703358993316
12 norm of weights: 5.907684752951629
-----
Iteration Number: 5122
Loss: 36.24934646862776
12 norm of gradients: 0.25423509849106385
12 norm of weights: 5.907597827040524
-----
Iteration Number: 5123
Loss: 36.24610539700738
12 norm of gradients: 0.25419987590425647
12 norm of weights: 5.907510909745431
-----
Iteration Number: 5124
Loss: 36.24286522134996
12 norm of gradients: 0.25416466812749094
12 norm of weights: 5.90742400106423
-----
Iteration Number: 5125
Loss: 36.23962594116297
12 norm of gradients: 0.25412947514935946
12 norm of weights: 5.907337100994808
-----
Iteration Number: 5126
Loss: 36.236387555952206
12 norm of gradients: 0.2540942969584652
12 norm of weights: 5.907250209535049
-----
Iteration Number: 5127
Loss: 36.233150065227385
12 norm of gradients: 0.2540591335434223
12 norm of weights: 5.9071633266828405
-----
Iteration Number: 5128

```

Loss: 36.22991346848314
l2 norm of gradients: 0.25402398489285544
l2 norm of weights: 5.907076452436069

Iteration Number: 5129
Loss: 36.22667776522991
l2 norm of gradients: 0.2539888509954004
l2 norm of weights: 5.9069895867926245

Iteration Number: 5130
Loss: 36.223442954980214
l2 norm of gradients: 0.25395373183970377
l2 norm of weights: 5.906902729750397

Iteration Number: 5131
Loss: 36.22020903723351
l2 norm of gradients: 0.2539186274144229
l2 norm of weights: 5.906815881307278

Iteration Number: 5132
Loss: 36.216976011499185
l2 norm of gradients: 0.25388353770822597
l2 norm of weights: 5.906729041461161

Iteration Number: 5133
Loss: 36.213743877284834
l2 norm of gradients: 0.253848462709792
l2 norm of weights: 5.906642210209941

Iteration Number: 5134
Loss: 36.21051263410057
l2 norm of gradients: 0.2538134024078108
l2 norm of weights: 5.906555387551511

Iteration Number: 5135
Loss: 36.20728228145478
l2 norm of gradients: 0.2537783567909829
l2 norm of weights: 5.906468573483769

Iteration Number: 5136
Loss: 36.204052818861484
l2 norm of gradients: 0.2537433258480196
l2 norm of weights: 5.906381768004614

Iteration Number: 5137
Loss: 36.20082424581137
l2 norm of gradients: 0.2537083095676432
l2 norm of weights: 5.906294971111944

```

-----
Iteration Number: 5138
Loss: 36.19759656183575
l2 norm of gradients: 0.2536733079385864
l2 norm of weights: 5.90620818280366
-----
Iteration Number: 5139
Loss: 36.19436976644021
l2 norm of gradients: 0.2536383209495928
l2 norm of weights: 5.9061214030776625
-----
Iteration Number: 5140
Loss: 36.1911438591286
l2 norm of gradients: 0.2536033485894169
l2 norm of weights: 5.906034631931857
-----
Iteration Number: 5141
Loss: 36.18791883941654
l2 norm of gradients: 0.25356839084682375
l2 norm of weights: 5.905947869364145
-----
Iteration Number: 5142
Loss: 36.18469470681262
l2 norm of gradients: 0.253533447710589
l2 norm of weights: 5.905861115372435
-----
Iteration Number: 5143
Loss: 36.18147146082822
l2 norm of gradients: 0.2534985191694994
l2 norm of weights: 5.905774369954631
-----
Iteration Number: 5144
Loss: 36.17824910098671
l2 norm of gradients: 0.2534636052123519
l2 norm of weights: 5.905687633108641
-----
Iteration Number: 5145
Loss: 36.17502762678731
l2 norm of gradients: 0.2534287058279545
l2 norm of weights: 5.905600904832378
-----
Iteration Number: 5146
Loss: 36.17180703774923
l2 norm of gradients: 0.25339382100512575
l2 norm of weights: 5.905514185123747
-----
Iteration Number: 5147
Loss: 36.16858733339531

```

```

12 norm of gradients: 0.25335895073269493
12 norm of weights: 5.905427473980665
-----
Iteration Number: 5148
Loss: 36.16536851322641
12 norm of gradients: 0.2533240949995018
12 norm of weights: 5.90534077140104
-----
Iteration Number: 5149
Loss: 36.16215057675534
12 norm of gradients: 0.253289253794397
12 norm of weights: 5.905254077382789
-----
Iteration Number: 5150
Loss: 36.15893352350669
12 norm of gradients: 0.2532544271062417
12 norm of weights: 5.905167391923829
-----
Iteration Number: 5151
Loss: 36.15571735299831
12 norm of gradients: 0.2532196149239076
12 norm of weights: 5.905080715022073
-----
Iteration Number: 5152
Loss: 36.152502064734975
12 norm of gradients: 0.25318481723627717
12 norm of weights: 5.904994046675442
-----
Iteration Number: 5153
Loss: 36.14928765823881
12 norm of gradients: 0.2531500340322435
12 norm of weights: 5.904907386881852
-----
Iteration Number: 5154
Loss: 36.14607413303163
12 norm of gradients: 0.2531152653007101
12 norm of weights: 5.904820735639227
-----
Iteration Number: 5155
Loss: 36.142861488616774
12 norm of gradients: 0.2530805110305912
12 norm of weights: 5.9047340929454855
-----
Iteration Number: 5156
Loss: 36.139649724527935
12 norm of gradients: 0.25304577121081157
12 norm of weights: 5.904647458798553
-----

```


Iteration Number: 5157
Loss: 36.1364388402778
l2 norm of gradients: 0.25301104583030665
l2 norm of weights: 5.904560833196351

Iteration Number: 5158
Loss: 36.133228835374304
l2 norm of gradients: 0.2529763348780223
l2 norm of weights: 5.904474216136806

Iteration Number: 5159
Loss: 36.13001970934869
l2 norm of gradients: 0.25294163834291494
l2 norm of weights: 5.9043876076178465

Iteration Number: 5160
Loss: 36.12681146172377
l2 norm of gradients: 0.25290695621395154
l2 norm of weights: 5.904301007637398

Iteration Number: 5161
Loss: 36.1236040919976
l2 norm of gradients: 0.25287228848010973
l2 norm of weights: 5.904214416193389

Iteration Number: 5162
Loss: 36.12039759971691
l2 norm of gradients: 0.25283763513037755
l2 norm of weights: 5.904127833283751

Iteration Number: 5163
Loss: 36.11719198437902
l2 norm of gradients: 0.2528029961537535
l2 norm of weights: 5.904041258906415

Iteration Number: 5164
Loss: 36.11398724552732
l2 norm of gradients: 0.2527683715392467
l2 norm of weights: 5.903954693059314

Iteration Number: 5165
Loss: 36.110783382664465
l2 norm of gradients: 0.2527337612758766
l2 norm of weights: 5.903868135740382

Iteration Number: 5166
Loss: 36.10758039531958
l2 norm of gradients: 0.25269916535267334

```

12 norm of weights: 5.903781586947553
-----
Iteration Number: 5167
Loss: 36.104378283019
12 norm of gradients: 0.2526645837586773
12 norm of weights: 5.903695046678766
-----
Iteration Number: 5168
Loss: 36.10117704528464
12 norm of gradients: 0.2526300164829395
12 norm of weights: 5.903608514931954
-----
Iteration Number: 5169
Loss: 36.09797668162866
12 norm of gradients: 0.25259546351452133
12 norm of weights: 5.90352199170506
-----
Iteration Number: 5170
Loss: 36.09477719157914
12 norm of gradients: 0.25256092484249465
12 norm of weights: 5.903435476996021
-----
Iteration Number: 5171
Loss: 36.09157857467224
12 norm of gradients: 0.2525264004559418
12 norm of weights: 5.9033489708027815
-----
Iteration Number: 5172
Loss: 36.088380830415915
12 norm of gradients: 0.2524918903439554
12 norm of weights: 5.903262473123281
-----
Iteration Number: 5173
Loss: 36.08518395834665
12 norm of gradients: 0.25245739449563864
12 norm of weights: 5.903175983955464
-----
Iteration Number: 5174
Loss: 36.08198795797946
12 norm of gradients: 0.25242291290010493
12 norm of weights: 5.9030895032972746
-----
Iteration Number: 5175
Loss: 36.07879282884764
12 norm of gradients: 0.25238844554647827
12 norm of weights: 5.90300303114666
-----
Iteration Number: 5176

```

Loss: 36.075598570474845
l2 norm of gradients: 0.252353992423893
l2 norm of weights: 5.902916567501568

Iteration Number: 5177
Loss: 36.072405182387065
l2 norm of gradients: 0.2523195535214937
l2 norm of weights: 5.902830112359945

Iteration Number: 5178
Loss: 36.06921266411094
l2 norm of gradients: 0.25228512882843546
l2 norm of weights: 5.902743665719742

Iteration Number: 5179
Loss: 36.066021015175814
l2 norm of gradients: 0.2522507183338836
l2 norm of weights: 5.902657227578909

Iteration Number: 5180
Loss: 36.0628302351077
l2 norm of gradients: 0.25221632202701383
l2 norm of weights: 5.9025707979353985

Iteration Number: 5181
Loss: 36.05964032342568
l2 norm of gradients: 0.2521819398970124
l2 norm of weights: 5.902484376787164

Iteration Number: 5182
Loss: 36.056451279677276
l2 norm of gradients: 0.2521475719330755
l2 norm of weights: 5.9023979641321604

Iteration Number: 5183
Loss: 36.05326310338063
l2 norm of gradients: 0.25211321812440995
l2 norm of weights: 5.902311559968341

Iteration Number: 5184
Loss: 36.05007579406186
l2 norm of gradients: 0.2520788784602328
l2 norm of weights: 5.902225164293665

Iteration Number: 5185
Loss: 36.04688935125119
l2 norm of gradients: 0.25204455292977124
l2 norm of weights: 5.9021387771060905

```

-----
Iteration Number: 5186
Loss: 36.04370377448188
l2 norm of gradients: 0.2520102415222629
l2 norm of weights: 5.902052398403574
-----
Iteration Number: 5187
Loss: 36.040519063285224
l2 norm of gradients: 0.2519759442269557
l2 norm of weights: 5.90196602818408
-----
Iteration Number: 5188
Loss: 36.037335217190545
l2 norm of gradients: 0.25194166103310783
l2 norm of weights: 5.9018796664455655
-----
Iteration Number: 5189
Loss: 36.03415223573671
l2 norm of gradients: 0.2519073919299875
l2 norm of weights: 5.901793313185998
-----
Iteration Number: 5190
Loss: 36.0309701184398
l2 norm of gradients: 0.2518731369068736
l2 norm of weights: 5.901706968403338
-----
Iteration Number: 5191
Loss: 36.0277888648469
l2 norm of gradients: 0.25183889595305486
l2 norm of weights: 5.901620632095551
-----
Iteration Number: 5192
Loss: 36.02460847447479
l2 norm of gradients: 0.2518046690578305
l2 norm of weights: 5.901534304260605
-----
Iteration Number: 5193
Loss: 36.02142894686562
l2 norm of gradients: 0.25177045621050975
l2 norm of weights: 5.901447984896467
-----
Iteration Number: 5194
Loss: 36.01825028155822
l2 norm of gradients: 0.25173625740041217
l2 norm of weights: 5.901361674001104
-----
Iteration Number: 5195
Loss: 36.01507247807736

```

```

12 norm of gradients: 0.2517020726168676
12 norm of weights: 5.901275371572488
-----
Iteration Number: 5196
Loss: 36.01189553596009
12 norm of gradients: 0.25166790184921595
12 norm of weights: 5.901189077608589
-----
Iteration Number: 5197
Loss: 36.0087194547472
12 norm of gradients: 0.25163374508680725
12 norm of weights: 5.90110279210738
-----
Iteration Number: 5198
Loss: 36.00554423395962
12 norm of gradients: 0.25159960231900197
12 norm of weights: 5.901016515066835
-----
Iteration Number: 5199
Loss: 36.002369873142555
12 norm of gradients: 0.25156547353517034
12 norm of weights: 5.900930246484926
-----
Iteration Number: 5200
Loss: 35.99919637182393
12 norm of gradients: 0.2515313587246931
12 norm of weights: 5.900843986359631
-----
Iteration Number: 5201
Loss: 35.9960237295581
12 norm of gradients: 0.25149725787696103
12 norm of weights: 5.900757734688926
-----
Iteration Number: 5202
Loss: 35.992851945857296
12 norm of gradients: 0.25146317098137505
12 norm of weights: 5.90067149147079
-----
Iteration Number: 5203
Loss: 35.98968102027686
12 norm of gradients: 0.251429098027346
12 norm of weights: 5.900585256703202
-----
Iteration Number: 5204
Loss: 35.986510952344865
12 norm of gradients: 0.2513950390042951
12 norm of weights: 5.900499030384142
-----

```

Iteration Number: 5205
Loss: 35.98334174159824
l2 norm of gradients: 0.2513609939016537
l2 norm of weights: 5.900412812511592

Iteration Number: 5206
Loss: 35.98017338758896
l2 norm of gradients: 0.2513269627088629
l2 norm of weights: 5.900326603083535

Iteration Number: 5207
Loss: 35.97700588983647
l2 norm of gradients: 0.2512929454153743
l2 norm of weights: 5.900240402097953

Iteration Number: 5208
Loss: 35.97383924788988
l2 norm of gradients: 0.25125894201064924
l2 norm of weights: 5.900154209552835

Iteration Number: 5209
Loss: 35.97067346128349
l2 norm of gradients: 0.2512249524841594
l2 norm of weights: 5.900068025446164

Iteration Number: 5210
Loss: 35.967508529566544
l2 norm of gradients: 0.2511909768253864
l2 norm of weights: 5.89998184977593

Iteration Number: 5211
Loss: 35.964344452269316
l2 norm of gradients: 0.2511570150238218
l2 norm of weights: 5.899895682540119

Iteration Number: 5212
Loss: 35.96118122893286
l2 norm of gradients: 0.25112306706896736
l2 norm of weights: 5.899809523736724

Iteration Number: 5213
Loss: 35.9580188591057
l2 norm of gradients: 0.2510891329503349
l2 norm of weights: 5.899723373363732

Iteration Number: 5214
Loss: 35.95485734232117
l2 norm of gradients: 0.251055212657446

```
12 norm of weights: 5.899637231419138
-----
Iteration Number: 5215
Loss: 35.951696678125856
12 norm of gradients: 0.25102130617983254
12 norm of weights: 5.899551097900934
-----
Iteration Number: 5216
Loss: 35.948536866060756
12 norm of gradients: 0.2509874135070363
12 norm of weights: 5.899464972807116
-----
Iteration Number: 5217
Loss: 35.94537790566426
12 norm of gradients: 0.250953534628609
12 norm of weights: 5.899378856135678
-----
Iteration Number: 5218
Loss: 35.94221979648374
12 norm of gradients: 0.25091966953411243
12 norm of weights: 5.899292747884617
-----
Iteration Number: 5219
Loss: 35.93906253805971
12 norm of gradients: 0.25088581821311823
12 norm of weights: 5.899206648051932
-----
Iteration Number: 5220
Loss: 35.93590612994322
12 norm of gradients: 0.2508519806552081
12 norm of weights: 5.89912055663562
-----
Iteration Number: 5221
Loss: 35.93275057167112
12 norm of gradients: 0.2508181568499737
12 norm of weights: 5.899034473633684
-----
Iteration Number: 5222
Loss: 35.929595862784744
12 norm of gradients: 0.25078434678701667
12 norm of weights: 5.898948399044122
-----
Iteration Number: 5223
Loss: 35.92644200283582
12 norm of gradients: 0.25075055045594835
12 norm of weights: 5.8988623328649386
-----
Iteration Number: 5224
```

Loss: 35.92328899136779
l2 norm of gradients: 0.25071676784639024
l2 norm of weights: 5.898776275094137

Iteration Number: 5225
Loss: 35.92013682792404
l2 norm of gradients: 0.25068299894797386
l2 norm of weights: 5.898690225729723

Iteration Number: 5226
Loss: 35.916985512051035
l2 norm of gradients: 0.25064924375034037
l2 norm of weights: 5.8986041847697015

Iteration Number: 5227
Loss: 35.91383504329715
l2 norm of gradients: 0.2506155022431409
l2 norm of weights: 5.898518152212079

Iteration Number: 5228
Loss: 35.91068542121639
l2 norm of gradients: 0.2505817744160365
l2 norm of weights: 5.898432128054865

Iteration Number: 5229
Loss: 35.90753664533869
l2 norm of gradients: 0.2505480602586982
l2 norm of weights: 5.8983461122960685

Iteration Number: 5230
Loss: 35.904388715217756
l2 norm of gradients: 0.2505143597608067
l2 norm of weights: 5.898260104933699

Iteration Number: 5231
Loss: 35.90124163041319
l2 norm of gradients: 0.2504806729120527
l2 norm of weights: 5.8981741059657695

Iteration Number: 5232
Loss: 35.898095390456014
l2 norm of gradients: 0.25044699970213674
l2 norm of weights: 5.8980881153902915

Iteration Number: 5233
Loss: 35.894949994907705
l2 norm of gradients: 0.2504133401207691
l2 norm of weights: 5.8980021332052806


```

-----
Iteration Number: 5234
Loss: 35.89180544331311
l2 norm of gradients: 0.25037969415767014
l2 norm of weights: 5.897916159408751
-----
Iteration Number: 5235
Loss: 35.888661735220516
l2 norm of gradients: 0.2503460618025697
l2 norm of weights: 5.897830193998717
-----
Iteration Number: 5236
Loss: 35.88551887018144
l2 norm of gradients: 0.2503124430452077
l2 norm of weights: 5.897744236973199
-----
Iteration Number: 5237
Loss: 35.882376847744936
l2 norm of gradients: 0.2502788378753337
l2 norm of weights: 5.897658288330214
-----
Iteration Number: 5238
Loss: 35.87923566747068
l2 norm of gradients: 0.2502452462827073
l2 norm of weights: 5.897572348067781
-----
Iteration Number: 5239
Loss: 35.876095328888226
l2 norm of gradients: 0.25021166825709756
l2 norm of weights: 5.897486416183922
-----
Iteration Number: 5240
Loss: 35.87295583156968
l2 norm of gradients: 0.25017810378828353
l2 norm of weights: 5.897400492676658
-----
Iteration Number: 5241
Loss: 35.8698171750585
l2 norm of gradients: 0.25014455286605397
l2 norm of weights: 5.8973145775440114
-----
Iteration Number: 5242
Loss: 35.86667935890307
l2 norm of gradients: 0.2501110154802074
l2 norm of weights: 5.897228670784008
-----
Iteration Number: 5243
Loss: 35.86354238266603

```

```

12 norm of gradients: 0.25007749162055215
12 norm of weights: 5.897142772394671
-----
Iteration Number: 5244
Loss: 35.86040624588652
12 norm of gradients: 0.25004398127690614
12 norm of weights: 5.897056882374029
-----
Iteration Number: 5245
Loss: 35.85727094813135
12 norm of gradients: 0.2500104844390972
12 norm of weights: 5.896971000720108
-----
Iteration Number: 5246
Loss: 35.85413648895593
12 norm of gradients: 0.24997700109696266
12 norm of weights: 5.896885127430936
-----
Iteration Number: 5247
Loss: 35.851002867893925
12 norm of gradients: 0.24994353124034988
12 norm of weights: 5.896799262504544
-----
Iteration Number: 5248
Loss: 35.84787008452056
12 norm of gradients: 0.2499100748591156
12 norm of weights: 5.8967134059389625
-----
Iteration Number: 5249
Loss: 35.8447381383847
12 norm of gradients: 0.2498766319431265
12 norm of weights: 5.896627557732223
-----
Iteration Number: 5250
Loss: 35.84160702903607
12 norm of gradients: 0.24984320248225886
12 norm of weights: 5.896541717882358
-----
Iteration Number: 5251
Loss: 35.83847675603551
12 norm of gradients: 0.24980978646639848
12 norm of weights: 5.896455886387403
-----
Iteration Number: 5252
Loss: 35.83534731893498
12 norm of gradients: 0.24977638388544113
12 norm of weights: 5.896370063245392
-----

```

Iteration Number: 5253
Loss: 35.83221871729713
l2 norm of gradients: 0.24974299472929196
l2 norm of weights: 5.8962842484543625

Iteration Number: 5254
Loss: 35.82909095067462
l2 norm of gradients: 0.24970961898786592
l2 norm of weights: 5.896198442012352

Iteration Number: 5255
Loss: 35.825964018627744
l2 norm of gradients: 0.24967625665108753
l2 norm of weights: 5.896112643917399

Iteration Number: 5256
Loss: 35.82283792070168
l2 norm of gradients: 0.24964290770889108
l2 norm of weights: 5.896026854167542

Iteration Number: 5257
Loss: 35.81971265647268
l2 norm of gradients: 0.2496095721512203
l2 norm of weights: 5.895941072760824

Iteration Number: 5258
Loss: 35.816588225487465
l2 norm of gradients: 0.24957624996802855
l2 norm of weights: 5.895855299695284

Iteration Number: 5259
Loss: 35.81346462730888
l2 norm of gradients: 0.24954294114927889
l2 norm of weights: 5.8957695349689665

Iteration Number: 5260
Loss: 35.81034186149585
l2 norm of gradients: 0.24950964568494405
l2 norm of weights: 5.895683778579917

Iteration Number: 5261
Loss: 35.80721992760426
l2 norm of gradients: 0.24947636356500613
l2 norm of weights: 5.895598030526179

Iteration Number: 5262
Loss: 35.804098825196476
l2 norm of gradients: 0.24944309477945684

```

12 norm of weights: 5.895512290805799
-----
Iteration Number: 5263
Loss: 35.80097855383017
12 norm of gradients: 0.24940983931829758
12 norm of weights: 5.895426559416825
-----
Iteration Number: 5264
Loss: 35.79785911306883
12 norm of gradients: 0.24937659717153926
12 norm of weights: 5.895340836357307
-----
Iteration Number: 5265
Loss: 35.79474050247329
12 norm of gradients: 0.24934336832920231
12 norm of weights: 5.895255121625292
-----
Iteration Number: 5266
Loss: 35.79162272160046
12 norm of gradients: 0.2493101527813167
12 norm of weights: 5.895169415218831
-----
Iteration Number: 5267
Loss: 35.788505770014496
12 norm of gradients: 0.249276950517922
12 norm of weights: 5.895083717135978
-----
Iteration Number: 5268
Loss: 35.785389647284276
12 norm of gradients: 0.2492437615290672
12 norm of weights: 5.8949980273747835
-----
Iteration Number: 5269
Loss: 35.7822743529644
12 norm of gradients: 0.2492105858048108
12 norm of weights: 5.894912345933302
-----
Iteration Number: 5270
Loss: 35.77915988661275
12 norm of gradients: 0.24917742333522103
12 norm of weights: 5.89482667280959
-----
Iteration Number: 5271
Loss: 35.77604624780415
12 norm of gradients: 0.24914427411037518
12 norm of weights: 5.894741008001703
-----
Iteration Number: 5272

```

Loss: 35.77293343609539
l2 norm of gradients: 0.24911113812036048
l2 norm of weights: 5.894655351507699

Iteration Number: 5273
Loss: 35.7698214510539
l2 norm of gradients: 0.2490780153552733
l2 norm of weights: 5.894569703325635

Iteration Number: 5274
Loss: 35.76671029223848
l2 norm of gradients: 0.24904490580521968
l2 norm of weights: 5.894484063453572

Iteration Number: 5275
Loss: 35.76359995922143
l2 norm of gradients: 0.249011809460315
l2 norm of weights: 5.8943984318895675

Iteration Number: 5276
Loss: 35.76049045156011
l2 norm of gradients: 0.24897872631068424
l2 norm of weights: 5.894312808631686

Iteration Number: 5277
Loss: 35.75738176882413
l2 norm of gradients: 0.24894565634646162
l2 norm of weights: 5.894227193677991

Iteration Number: 5278
Loss: 35.75427391057095
l2 norm of gradients: 0.24891259955779088
l2 norm of weights: 5.894141587026544

Iteration Number: 5279
Loss: 35.751166876381404
l2 norm of gradients: 0.24887955593482533
l2 norm of weights: 5.894055988675412

Iteration Number: 5280
Loss: 35.74806066581078
l2 norm of gradients: 0.24884652546772731
l2 norm of weights: 5.893970398622658

Iteration Number: 5281
Loss: 35.74495527843101
l2 norm of gradients: 0.24881350814666905
l2 norm of weights: 5.893884816866352

```
-----  
Iteration Number: 5282  
Loss: 35.74185071380381  
l2 norm of gradients: 0.24878050396183168  
l2 norm of weights: 5.893799243404561  
-----  
Iteration Number: 5283  
Loss: 35.73874697150175  
l2 norm of gradients: 0.2487475129034062  
l2 norm of weights: 5.893713678235355  
-----  
Iteration Number: 5284  
Loss: 35.73564405109144  
l2 norm of gradients: 0.2487145349615926  
l2 norm of weights: 5.8936281213568025  
-----  
Iteration Number: 5285  
Loss: 35.73254195213889  
l2 norm of gradients: 0.24868157012660053  
l2 norm of weights: 5.893542572766977  
-----  
Iteration Number: 5286  
Loss: 35.72944067421791  
l2 norm of gradients: 0.24864861838864866  
l2 norm of weights: 5.8934570324639495  
-----  
Iteration Number: 5287  
Loss: 35.72634021689252  
l2 norm of gradients: 0.2486156797379653  
l2 norm of weights: 5.893371500445794  
-----  
Iteration Number: 5288  
Loss: 35.72324057973256  
l2 norm of gradients: 0.24858275416478812  
l2 norm of weights: 5.893285976710585  
-----  
Iteration Number: 5289  
Loss: 35.72014176230917  
l2 norm of gradients: 0.2485498416593638  
l2 norm of weights: 5.893200461256399  
-----  
Iteration Number: 5290  
Loss: 35.717043764194536  
l2 norm of gradients: 0.24851694221194873  
l2 norm of weights: 5.8931149540813115  
-----  
Iteration Number: 5291  
Loss: 35.713946584957164
```

```

12 norm of gradients: 0.24848405581280833
12 norm of weights: 5.8930294551834015
-----
Iteration Number: 5292
Loss: 35.71085022416935
12 norm of gradients: 0.24845118245221753
12 norm of weights: 5.892943964560747
-----
Iteration Number: 5293
Loss: 35.7077546813946
12 norm of gradients: 0.24841832212046033
12 norm of weights: 5.892858482211429
-----
Iteration Number: 5294
Loss: 35.70465995620977
12 norm of gradients: 0.24838547480783035
12 norm of weights: 5.892773008133528
-----
Iteration Number: 5295
Loss: 35.701566048194756
12 norm of gradients: 0.24835264050463
12 norm of weights: 5.892687542325124
-----
Iteration Number: 5296
Loss: 35.698472956901554
12 norm of gradients: 0.2483198192011715
12 norm of weights: 5.892602084784303
-----
Iteration Number: 5297
Loss: 35.695380681929436
12 norm of gradients: 0.2482870108877759
12 norm of weights: 5.892516635509148
-----
Iteration Number: 5298
Loss: 35.692289222828364
12 norm of gradients: 0.24825421555477378
12 norm of weights: 5.892431194497746
-----
Iteration Number: 5299
Loss: 35.68919857918203
12 norm of gradients: 0.24822143319250478
12 norm of weights: 5.892345761748182
-----
Iteration Number: 5300
Loss: 35.68610875056221
12 norm of gradients: 0.24818866379131793
12 norm of weights: 5.892260337258541
-----

```

Iteration Number: 5301
Loss: 35.68301973654958
l2 norm of gradients: 0.24815590734157136
l2 norm of weights: 5.892174921026916

Iteration Number: 5302
Loss: 35.67993153670571
l2 norm of gradients: 0.24812316383363253
l2 norm of weights: 5.892089513051394

Iteration Number: 5303
Loss: 35.67684415061891
l2 norm of gradients: 0.24809043325787788
l2 norm of weights: 5.892004113330065

Iteration Number: 5304
Loss: 35.67375757784798
l2 norm of gradients: 0.2480577156046934
l2 norm of weights: 5.891918721861023

Iteration Number: 5305
Loss: 35.67067181798311
l2 norm of gradients: 0.24802501086447398
l2 norm of weights: 5.891833338642358

Iteration Number: 5306
Loss: 35.66758687058518
l2 norm of gradients: 0.24799231902762378
l2 norm of weights: 5.891747963672166

Iteration Number: 5307
Loss: 35.66450273525192
l2 norm of gradients: 0.2479596400845562
l2 norm of weights: 5.89166259694854

Iteration Number: 5308
Loss: 35.661419411544315
l2 norm of gradients: 0.24792697402569377
l2 norm of weights: 5.891577238469576

Iteration Number: 5309
Loss: 35.65833689904256
l2 norm of gradients: 0.24789432084146806
l2 norm of weights: 5.891491888233372

Iteration Number: 5310
Loss: 35.65525519732389
l2 norm of gradients: 0.24786168052231988


```

12 norm of weights: 5.891406546238026
-----
Iteration Number: 5311
Loss: 35.652174305970874
12 norm of gradients: 0.24782905305869923
12 norm of weights: 5.891321212481638
-----
Iteration Number: 5312
Loss: 35.64909422454766
12 norm of gradients: 0.24779643844106516
12 norm of weights: 5.8912358869623045
-----
Iteration Number: 5313
Loss: 35.64601495265054
12 norm of gradients: 0.24776383665988597
12 norm of weights: 5.891150569678129
-----
Iteration Number: 5314
Loss: 35.642936489843
12 norm of gradients: 0.24773124770563876
12 norm of weights: 5.891065260627213
-----
Iteration Number: 5315
Loss: 35.639858835712495
12 norm of gradients: 0.24769867156881006
12 norm of weights: 5.890979959807661
-----
Iteration Number: 5316
Loss: 35.63678198983588
12 norm of gradients: 0.24766610823989532
12 norm of weights: 5.890894667217576
-----
Iteration Number: 5317
Loss: 35.633705951796514
12 norm of gradients: 0.24763355770939918
12 norm of weights: 5.890809382855062
-----
Iteration Number: 5318
Loss: 35.63063072116624
12 norm of gradients: 0.2476010199678352
12 norm of weights: 5.8907241067182285
-----
Iteration Number: 5319
Loss: 35.627556297529786
12 norm of gradients: 0.24756849500572625
12 norm of weights: 5.89063883880518
-----
Iteration Number: 5320

```

Loss: 35.624482680474394
l2 norm of gradients: 0.2475359828136039
l2 norm of weights: 5.890553579114027

Iteration Number: 5321
Loss: 35.621409869572624
l2 norm of gradients: 0.24750348338200917
l2 norm of weights: 5.8904683276428775

Iteration Number: 5322
Loss: 35.61833786440948
l2 norm of gradients: 0.2474709967014919
l2 norm of weights: 5.890383084389843

Iteration Number: 5323
Loss: 35.61526666456377
l2 norm of gradients: 0.24743852276261097
l2 norm of weights: 5.890297849353033

Iteration Number: 5324
Loss: 35.61219626962462
l2 norm of gradients: 0.24740606155593428
l2 norm of weights: 5.890212622530562

Iteration Number: 5325
Loss: 35.60912667916745
l2 norm of gradients: 0.2473736130720388
l2 norm of weights: 5.890127403920544

Iteration Number: 5326
Loss: 35.6060578927762
l2 norm of gradients: 0.2473411773015105
l2 norm of weights: 5.890042193521091

Iteration Number: 5327
Loss: 35.6029899100342
l2 norm of gradients: 0.2473087542349443
l2 norm of weights: 5.889956991330321

Iteration Number: 5328
Loss: 35.59992273053218
l2 norm of gradients: 0.24727634386294417
l2 norm of weights: 5.889871797346349

Iteration Number: 5329
Loss: 35.59685635384638
l2 norm of gradients: 0.24724394617612291
l2 norm of weights: 5.889786611567295

```

-----
Iteration Number: 5330
Loss: 35.59379077955843
l2 norm of gradients: 0.24721156116510257
l2 norm of weights: 5.889701433991275
-----
Iteration Number: 5331
Loss: 35.590726007260706
l2 norm of gradients: 0.24717918882051387
l2 norm of weights: 5.889616264616412
-----
Iteration Number: 5332
Loss: 35.587662036539086
l2 norm of gradients: 0.24714682913299665
l2 norm of weights: 5.889531103440823
-----
Iteration Number: 5333
Loss: 35.58459886696511
l2 norm of gradients: 0.2471144820931997
l2 norm of weights: 5.8894459504626315
-----
Iteration Number: 5334
Loss: 35.581536498143365
l2 norm of gradients: 0.24708214769178058
l2 norm of weights: 5.889360805679961
-----
Iteration Number: 5335
Loss: 35.57847492965003
l2 norm of gradients: 0.2470498259194061
l2 norm of weights: 5.889275669090935
-----
Iteration Number: 5336
Loss: 35.57541416106639
l2 norm of gradients: 0.24701751676675163
l2 norm of weights: 5.8891905406936775
-----
Iteration Number: 5337
Loss: 35.57235419198553
l2 norm of gradients: 0.24698522022450167
l2 norm of weights: 5.889105420486315
-----
Iteration Number: 5338
Loss: 35.569295022001796
l2 norm of gradients: 0.24695293628334955
l2 norm of weights: 5.889020308466975
-----
Iteration Number: 5339
Loss: 35.566236650687316

```

```

12 norm of gradients: 0.24692066493399756
12 norm of weights: 5.888935204633784
-----
Iteration Number: 5340
Loss: 35.563179077637784
12 norm of gradients: 0.2468884061671567
12 norm of weights: 5.8888501089848715
-----
Iteration Number: 5341
Loss: 35.560122302443716
12 norm of gradients: 0.24685615997354712
12 norm of weights: 5.888765021518369
-----
Iteration Number: 5342
Loss: 35.55706632468544
12 norm of gradients: 0.24682392634389763
12 norm of weights: 5.888679942232405
-----
Iteration Number: 5343
Loss: 35.55401114396114
12 norm of gradients: 0.2467917052689459
12 norm of weights: 5.888594871125113
-----
Iteration Number: 5344
Loss: 35.550956759853456
12 norm of gradients: 0.24675949673943856
12 norm of weights: 5.888509808194627
-----
Iteration Number: 5345
Loss: 35.54790317195003
12 norm of gradients: 0.24672730074613106
12 norm of weights: 5.888424753439078
-----
Iteration Number: 5346
Loss: 35.54485037985289
12 norm of gradients: 0.24669511727978757
12 norm of weights: 5.888339706856603
-----
Iteration Number: 5347
Loss: 35.541798383138264
12 norm of gradients: 0.24666294633118122
12 norm of weights: 5.888254668445338
-----
Iteration Number: 5348
Loss: 35.538747181403345
12 norm of gradients: 0.24663078789109386
12 norm of weights: 5.888169638203419
-----

```

Iteration Number: 5349
Loss: 35.535696774243284
l2 norm of gradients: 0.24659864195031628
l2 norm of weights: 5.888084616128987

Iteration Number: 5350
Loss: 35.53264716123047
l2 norm of gradients: 0.2465665084996479
l2 norm of weights: 5.8879996022201775

Iteration Number: 5351
Loss: 35.52959834197801
l2 norm of gradients: 0.24653438752989704
l2 norm of weights: 5.887914596475132

Iteration Number: 5352
Loss: 35.526550316070285
l2 norm of gradients: 0.24650227903188077
l2 norm of weights: 5.887829598891992

Iteration Number: 5353
Loss: 35.52350308309594
l2 norm of gradients: 0.24647018299642498
l2 norm of weights: 5.887744609468898

Iteration Number: 5354
Loss: 35.52045664265529
l2 norm of gradients: 0.24643809941436423
l2 norm of weights: 5.887659628203996

Iteration Number: 5355
Loss: 35.51741099432574
l2 norm of gradients: 0.2464060282765419
l2 norm of weights: 5.887574655095427

Iteration Number: 5356
Loss: 35.51436613772152
l2 norm of gradients: 0.24637396957381014
l2 norm of weights: 5.887489690141338

Iteration Number: 5357
Loss: 35.511322072416704
l2 norm of gradients: 0.24634192329702978
l2 norm of weights: 5.887404733339875

Iteration Number: 5358
Loss: 35.50827879801939
l2 norm of gradients: 0.24630988943707033

```

12 norm of weights: 5.887319784689184
-----
Iteration Number: 5359
Loss: 35.505236314118015
12 norm of gradients: 0.2462778679848103
12 norm of weights: 5.887234844187414
-----
Iteration Number: 5360
Loss: 35.50219462030882
12 norm of gradients: 0.24624585893113662
12 norm of weights: 5.887149911832712
-----
Iteration Number: 5361
Loss: 35.49915371617727
12 norm of gradients: 0.24621386226694494
12 norm of weights: 5.887064987623233
-----
Iteration Number: 5362
Loss: 35.49611360133073
12 norm of gradients: 0.24618187798313984
12 norm of weights: 5.886980071557122
-----
Iteration Number: 5363
Loss: 35.49307427536138
12 norm of gradients: 0.24614990607063428
12 norm of weights: 5.886895163632536
-----
Iteration Number: 5364
Loss: 35.490035737870116
12 norm of gradients: 0.24611794652035018
12 norm of weights: 5.886810263847626
-----
Iteration Number: 5365
Loss: 35.48699798843667
12 norm of gradients: 0.246085999323218
12 norm of weights: 5.886725372200544
-----
Iteration Number: 5366
Loss: 35.48396102667214
12 norm of gradients: 0.24605406447017683
12 norm of weights: 5.88664048868945
-----
Iteration Number: 5367
Loss: 35.48092485216744
12 norm of gradients: 0.24602214195217448
12 norm of weights: 5.886555613312496
-----
Iteration Number: 5368

```

Loss: 35.4778894645259
l2 norm of gradients: 0.24599023176016738
l2 norm of weights: 5.88647074606784

Iteration Number: 5369
Loss: 35.4748548633427
l2 norm of gradients: 0.2459583338851207
l2 norm of weights: 5.88638588695364

Iteration Number: 5370
Loss: 35.47182104821023
l2 norm of gradients: 0.24592644831800806
l2 norm of weights: 5.8863010359680565

Iteration Number: 5371
Loss: 35.46878801872581
l2 norm of gradients: 0.24589457504981171
l2 norm of weights: 5.886216193109248

Iteration Number: 5372
Loss: 35.465755774496486
l2 norm of gradients: 0.2458627140715228
l2 norm of weights: 5.8861313583753745

Iteration Number: 5373
Loss: 35.46272431512571
l2 norm of gradients: 0.24583086537414067
l2 norm of weights: 5.886046531764601

Iteration Number: 5374
Loss: 35.45969364019241
l2 norm of gradients: 0.24579902894867361
l2 norm of weights: 5.885961713275087

Iteration Number: 5375
Loss: 35.456663749315126
l2 norm of gradients: 0.24576720478613825
l2 norm of weights: 5.885876902904998

Iteration Number: 5376
Loss: 35.453634642081695
l2 norm of gradients: 0.24573539287756
l2 norm of weights: 5.885792100652501

Iteration Number: 5377
Loss: 35.45060631810388
l2 norm of gradients: 0.24570359321397267
l2 norm of weights: 5.885707306515759

```

-----
Iteration Number: 5378
Loss: 35.447578776968996
l2 norm of gradients: 0.2456718057864188
l2 norm of weights: 5.885622520492939
-----
Iteration Number: 5379
Loss: 35.44455201828564
l2 norm of gradients: 0.2456400305859493
l2 norm of weights: 5.88553774258221
-----
Iteration Number: 5380
Loss: 35.44152604165994
l2 norm of gradients: 0.24560826760362378
l2 norm of weights: 5.885452972781741
-----
Iteration Number: 5381
Loss: 35.43850084668191
l2 norm of gradients: 0.24557651683051024
l2 norm of weights: 5.885368211089701
-----
Iteration Number: 5382
Loss: 35.435476432956186
l2 norm of gradients: 0.24554477825768545
l2 norm of weights: 5.885283457504261
-----
Iteration Number: 5383
Loss: 35.4324528000929
l2 norm of gradients: 0.24551305187623437
l2 norm of weights: 5.885198712023593
-----
Iteration Number: 5384
Loss: 35.42942994768915
l2 norm of gradients: 0.24548133767725075
l2 norm of weights: 5.885113974645869
-----
Iteration Number: 5385
Loss: 35.426407875348566
l2 norm of gradients: 0.24544963565183683
l2 norm of weights: 5.885029245369265
-----
Iteration Number: 5386
Loss: 35.42338658266912
l2 norm of gradients: 0.24541794579110307
l2 norm of weights: 5.884944524191952
-----
Iteration Number: 5387
Loss: 35.42036606925852

```



```

12 norm of gradients: 0.24538626808616876
12 norm of weights: 5.884859811112109
-----
Iteration Number: 5388
Loss: 35.417346334729025
12 norm of gradients: 0.24535460252816146
12 norm of weights: 5.884775106127909
-----
Iteration Number: 5389
Loss: 35.414327378671516
12 norm of gradients: 0.2453229491082173
12 norm of weights: 5.884690409237533
-----
Iteration Number: 5390
Loss: 35.41130920069032
12 norm of gradients: 0.24529130781748082
12 norm of weights: 5.884605720439159
-----
Iteration Number: 5391
Loss: 35.40829180040366
12 norm of gradients: 0.24525967864710507
12 norm of weights: 5.8845210397309655
-----
Iteration Number: 5392
Loss: 35.40527517740697
12 norm of gradients: 0.24522806158825144
12 norm of weights: 5.8844363671111335
-----
Iteration Number: 5393
Loss: 35.40225933130507
12 norm of gradients: 0.2451964566320899
12 norm of weights: 5.8843517025778445
-----
Iteration Number: 5394
Loss: 35.39924426170273
12 norm of gradients: 0.24516486376979876
12 norm of weights: 5.884267046129279
-----
Iteration Number: 5395
Loss: 35.396229968219
12 norm of gradients: 0.24513328299256473
12 norm of weights: 5.884182397763624
-----
Iteration Number: 5396
Loss: 35.39321645043857
12 norm of gradients: 0.24510171429158303
12 norm of weights: 5.884097757479061
-----

```

Iteration Number: 5397
Loss: 35.390203707987105
l2 norm of gradients: 0.24507015765805723
l2 norm of weights: 5.884013125273777

Iteration Number: 5398
Loss: 35.387191740463074
l2 norm of gradients: 0.24503861308319932
l2 norm of weights: 5.883928501145957

Iteration Number: 5399
Loss: 35.38418054748108
l2 norm of gradients: 0.2450070805582296
l2 norm of weights: 5.883843885093789

Iteration Number: 5400
Loss: 35.381170128634515
l2 norm of gradients: 0.2449755600743768
l2 norm of weights: 5.88375927711546

Iteration Number: 5401
Loss: 35.378160483547305
l2 norm of gradients: 0.2449440516228781
l2 norm of weights: 5.883674677209161

Iteration Number: 5402
Loss: 35.37515161181872
l2 norm of gradients: 0.24491255519497904
l2 norm of weights: 5.883590085373081

Iteration Number: 5403
Loss: 35.37214351305897
l2 norm of gradients: 0.24488107078193325
l2 norm of weights: 5.883505501605411

Iteration Number: 5404
Loss: 35.36913618687808
l2 norm of gradients: 0.24484959837500317
l2 norm of weights: 5.8834209259043435

Iteration Number: 5405
Loss: 35.36612963288093
l2 norm of gradients: 0.24481813796545915
l2 norm of weights: 5.883336358268071

Iteration Number: 5406
Loss: 35.36312385068392
l2 norm of gradients: 0.24478668954458013

```

12 norm of weights: 5.883251798694788
-----
Iteration Number: 5407
Loss: 35.36011883989164
12 norm of gradients: 0.24475525310365323
12 norm of weights: 5.883167247182688
-----
Iteration Number: 5408
Loss: 35.3571146001236
12 norm of gradients: 0.24472382863397402
12 norm of weights: 5.8830827037299676
-----
Iteration Number: 5409
Loss: 35.35411113097445
12 norm of gradients: 0.24469241612684622
12 norm of weights: 5.882998168334825
-----
Iteration Number: 5410
Loss: 35.351108432066475
12 norm of gradients: 0.24466101557358208
12 norm of weights: 5.882913640995456
-----
Iteration Number: 5411
Loss: 35.348106503015345
12 norm of gradients: 0.2446296269655018
12 norm of weights: 5.882829121710061
-----
Iteration Number: 5412
Loss: 35.3451053434163
12 norm of gradients: 0.24459825029393428
12 norm of weights: 5.882744610476837
-----
Iteration Number: 5413
Loss: 35.3421049529004
12 norm of gradients: 0.24456688555021627
12 norm of weights: 5.882660107293987
-----
Iteration Number: 5414
Loss: 35.33910533106494
12 norm of gradients: 0.24453553272569312
12 norm of weights: 5.882575612159711
-----
Iteration Number: 5415
Loss: 35.33610647752439
12 norm of gradients: 0.24450419181171834
12 norm of weights: 5.882491125072211
-----
Iteration Number: 5416

```

Loss: 35.3331083919052
l2 norm of gradients: 0.2444728627996536
l2 norm of weights: 5.882406646029693

Iteration Number: 5417
Loss: 35.330111073805035
l2 norm of gradients: 0.24444154568086887
l2 norm of weights: 5.882322175030359

Iteration Number: 5418
Loss: 35.32711452283859
l2 norm of gradients: 0.24441024044674242
l2 norm of weights: 5.882237712072415

Iteration Number: 5419
Loss: 35.32411873862838
l2 norm of gradients: 0.24437894708866065
l2 norm of weights: 5.882153257154068

Iteration Number: 5420
Loss: 35.32112372078055
l2 norm of gradients: 0.24434766559801824
l2 norm of weights: 5.882068810273524

Iteration Number: 5421
Loss: 35.318129468914776
l2 norm of gradients: 0.2443163959662181
l2 norm of weights: 5.881984371428992

Iteration Number: 5422
Loss: 35.31513598264212
l2 norm of gradients: 0.24428513818467118
l2 norm of weights: 5.88189994061868

Iteration Number: 5423
Loss: 35.31214326158256
l2 norm of gradients: 0.24425389224479682
l2 norm of weights: 5.881815517840798

Iteration Number: 5424
Loss: 35.309151305344066
l2 norm of gradients: 0.2442226581380225
l2 norm of weights: 5.881731103093558

Iteration Number: 5425
Loss: 35.30616011355228
l2 norm of gradients: 0.24419143585578382
l2 norm of weights: 5.881646696375172

```
-----
Iteration Number: 5426
Loss: 35.30316968580743
l2 norm of gradients: 0.24416022538952467
l2 norm of weights: 5.881562297683851
-----
Iteration Number: 5427
Loss: 35.300180021743806
l2 norm of gradients: 0.2441290267306969
l2 norm of weights: 5.88147790701781
-----
Iteration Number: 5428
Loss: 35.29719112096604
l2 norm of gradients: 0.24409783987076064
l2 norm of weights: 5.881393524375263
-----
Iteration Number: 5429
Loss: 35.29420298309675
l2 norm of gradients: 0.2440666648011843
l2 norm of weights: 5.881309149754426
-----
Iteration Number: 5430
Loss: 35.29121560775302
l2 norm of gradients: 0.24403550151344405
l2 norm of weights: 5.881224783153515
-----
Iteration Number: 5431
Loss: 35.28822899454413
l2 norm of gradients: 0.24400434999902465
l2 norm of weights: 5.881140424570749
-----
Iteration Number: 5432
Loss: 35.28524314309605
l2 norm of gradients: 0.24397321024941865
l2 norm of weights: 5.881056074004344
-----
Iteration Number: 5433
Loss: 35.28225805302752
l2 norm of gradients: 0.24394208225612687
l2 norm of weights: 5.880971731452521
-----
Iteration Number: 5434
Loss: 35.2792737239546
l2 norm of gradients: 0.2439109660106581
l2 norm of weights: 5.880887396913501
-----
Iteration Number: 5435
Loss: 35.27629015549548
```

```

12 norm of gradients: 0.24387986150452945
12 norm of weights: 5.880803070385502
-----
Iteration Number: 5436
Loss: 35.273307347272606
12 norm of gradients: 0.2438487687292659
12 norm of weights: 5.880718751866749
-----
Iteration Number: 5437
Loss: 35.270325298899806
12 norm of gradients: 0.24381768767640064
12 norm of weights: 5.880634441355464
-----
Iteration Number: 5438
Loss: 35.267344009998176
12 norm of gradients: 0.2437866183374749
12 norm of weights: 5.88055013884987
-----
Iteration Number: 5439
Loss: 35.264363480192266
12 norm of gradients: 0.24375556070403795
12 norm of weights: 5.880465844348193
-----
Iteration Number: 5440
Loss: 35.261383709097686
12 norm of gradients: 0.24372451476764714
12 norm of weights: 5.880381557848659
-----
Iteration Number: 5441
Loss: 35.25840469633763
12 norm of gradients: 0.24369348051986797
12 norm of weights: 5.880297279349493
-----
Iteration Number: 5442
Loss: 35.25542644153092
12 norm of gradients: 0.24366245795227376
12 norm of weights: 5.880213008848925
-----
Iteration Number: 5443
Loss: 35.25244894430011
12 norm of gradients: 0.24363144705644607
12 norm of weights: 5.880128746345182
-----
Iteration Number: 5444
Loss: 35.249472204271314
12 norm of gradients: 0.2436004478239744
12 norm of weights: 5.880044491836493
-----

```

Iteration Number: 5445
Loss: 35.246496221059985
l2 norm of gradients: 0.24356946024645626
l2 norm of weights: 5.879960245321088

Iteration Number: 5446
Loss: 35.24352099428765
l2 norm of gradients: 0.24353848431549724
l2 norm of weights: 5.879876006797201

Iteration Number: 5447
Loss: 35.240546523577805
l2 norm of gradients: 0.24350752002271084
l2 norm of weights: 5.879791776263063

Iteration Number: 5448
Loss: 35.23757280856009
l2 norm of gradients: 0.2434765673597186
l2 norm of weights: 5.879707553716905

Iteration Number: 5449
Loss: 35.23459984884891
l2 norm of gradients: 0.24344562631815012
l2 norm of weights: 5.8796233391569634

Iteration Number: 5450
Loss: 35.231627644068354
l2 norm of gradients: 0.24341469688964282
l2 norm of weights: 5.879539132581471

Iteration Number: 5451
Loss: 35.22865619385023
l2 norm of gradients: 0.2433837790658423
l2 norm of weights: 5.879454933988666

Iteration Number: 5452
Loss: 35.2256854978054
l2 norm of gradients: 0.2433528728384018
l2 norm of weights: 5.879370743376783

Iteration Number: 5453
Loss: 35.22271555557125
l2 norm of gradients: 0.24332197819898294
l2 norm of weights: 5.879286560744062

Iteration Number: 5454
Loss: 35.21974636677135
l2 norm of gradients: 0.24329109513925498

```
12 norm of weights: 5.87920238608874
-----
Iteration Number: 5455
Loss: 35.21677793102045
12 norm of gradients: 0.24326022365089517
12 norm of weights: 5.879118219409056
-----
Iteration Number: 5456
Loss: 35.213810247949034
12 norm of gradients: 0.24322936372558882
12 norm of weights: 5.879034060703252
-----
Iteration Number: 5457
Loss: 35.21084331718343
12 norm of gradients: 0.24319851535502904
12 norm of weights: 5.878949909969568
-----
Iteration Number: 5458
Loss: 35.20787713834814
12 norm of gradients: 0.24316767853091698
12 norm of weights: 5.878865767206246
-----
Iteration Number: 5459
Loss: 35.20491171107454
12 norm of gradients: 0.24313685324496145
12 norm of weights: 5.87878163241153
-----
Iteration Number: 5460
Loss: 35.20194703498247
12 norm of gradients: 0.2431060394888795
12 norm of weights: 5.8786975055836646
-----
Iteration Number: 5461
Loss: 35.19898310969634
12 norm of gradients: 0.24307523725439578
12 norm of weights: 5.878613386720893
-----
Iteration Number: 5462
Loss: 35.19601993485652
12 norm of gradients: 0.24304444653324303
12 norm of weights: 5.878529275821462
-----
Iteration Number: 5463
Loss: 35.1930575100694
12 norm of gradients: 0.24301366731716179
12 norm of weights: 5.878445172883618
-----
Iteration Number: 5464
```


Loss: 35.19009583498609
l2 norm of gradients: 0.24298289959790043
l2 norm of weights: 5.878361077905608

Iteration Number: 5465
Loss: 35.18713490922151
l2 norm of gradients: 0.24295214336721527
l2 norm of weights: 5.878276990885682

Iteration Number: 5466
Loss: 35.184174732408295
l2 norm of gradients: 0.24292139861687034
l2 norm of weights: 5.8781929118220875

Iteration Number: 5467
Loss: 35.18121530416401
l2 norm of gradients: 0.24289066533863773
l2 norm of weights: 5.878108840713077

Iteration Number: 5468
Loss: 35.17825662413122
l2 norm of gradients: 0.24285994352429724
l2 norm of weights: 5.878024777556901

Iteration Number: 5469
Loss: 35.17529869193354
l2 norm of gradients: 0.24282923316563643
l2 norm of weights: 5.877940722351809

Iteration Number: 5470
Loss: 35.17234150719974
l2 norm of gradients: 0.24279853425445083
l2 norm of weights: 5.877856675096058

Iteration Number: 5471
Loss: 35.169385069558565
l2 norm of gradients: 0.24276784678254373
l2 norm of weights: 5.8777726357879

Iteration Number: 5472
Loss: 35.16642937864623
l2 norm of gradients: 0.24273717074172624
l2 norm of weights: 5.877688604425589

Iteration Number: 5473
Loss: 35.16347443407959
l2 norm of gradients: 0.24270650612381728
l2 norm of weights: 5.877604581007382

```
-----  
Iteration Number: 5474  
Loss: 35.16052023551236  
l2 norm of gradients: 0.24267585292064342  
l2 norm of weights: 5.877520565531534  
-----  
Iteration Number: 5475  
Loss: 35.157566782551534  
l2 norm of gradients: 0.24264521112403933  
l2 norm of weights: 5.877436557996304  
-----  
Iteration Number: 5476  
Loss: 35.1546140748423  
l2 norm of gradients: 0.24261458072584707  
l2 norm of weights: 5.8773525583999495  
-----  
Iteration Number: 5477  
Loss: 35.15166211201228  
l2 norm of gradients: 0.2425839617179168  
l2 norm of weights: 5.877268566740731  
-----  
Iteration Number: 5478  
Loss: 35.14871089368775  
l2 norm of gradients: 0.24255335409210627  
l2 norm of weights: 5.8771845830169065  
-----  
Iteration Number: 5479  
Loss: 35.14576041950754  
l2 norm of gradients: 0.242522757840281  
l2 norm of weights: 5.877100607226739  
-----  
Iteration Number: 5480  
Loss: 35.14281068911214  
l2 norm of gradients: 0.24249217295431424  
l2 norm of weights: 5.877016639368489  
-----  
Iteration Number: 5481  
Loss: 35.13986170211654  
l2 norm of gradients: 0.2424615994260872  
l2 norm of weights: 5.8769326794404195  
-----  
Iteration Number: 5482  
Loss: 35.13691345816026  
l2 norm of gradients: 0.24243103724748843  
l2 norm of weights: 5.8768487274407955  
-----  
Iteration Number: 5483  
Loss: 35.133965956886044
```

```
12 norm of gradients: 0.24240048641041456
12 norm of weights: 5.876764783367879
-----
Iteration Number: 5484
Loss: 35.131019197915386
12 norm of gradients: 0.24236994690676975
12 norm of weights: 5.876680847219938
-----
Iteration Number: 5485
Loss: 35.12807318088298
12 norm of gradients: 0.2423394187284659
12 norm of weights: 5.8765969189952365
-----
Iteration Number: 5486
Loss: 35.125127905433715
12 norm of gradients: 0.2423089018674226
12 norm of weights: 5.876512998692045
-----
Iteration Number: 5487
Loss: 35.12218337119045
12 norm of gradients: 0.24227839631556716
12 norm of weights: 5.876429086308627
-----
Iteration Number: 5488
Loss: 35.119239577796414
12 norm of gradients: 0.2422479020648347
12 norm of weights: 5.876345181843255
-----
Iteration Number: 5489
Loss: 35.116296524881434
12 norm of gradients: 0.24221741910716774
12 norm of weights: 5.876261285294198
-----
Iteration Number: 5490
Loss: 35.11335421207785
12 norm of gradients: 0.24218694743451666
12 norm of weights: 5.876177396659725
-----
Iteration Number: 5491
Loss: 35.11041263903068
12 norm of gradients: 0.24215648703883952
12 norm of weights: 5.87609351593811
-----
Iteration Number: 5492
Loss: 35.10747180537254
12 norm of gradients: 0.24212603791210194
12 norm of weights: 5.876009643127624
-----
```

Iteration Number: 5493
Loss: 35.10453171073315
l2 norm of gradients: 0.24209560004627723
l2 norm of weights: 5.875925778226541

Iteration Number: 5494
Loss: 35.10159235475648
l2 norm of gradients: 0.24206517343334635
l2 norm of weights: 5.875841921233135

Iteration Number: 5495
Loss: 35.09865373707427
l2 norm of gradients: 0.24203475806529792
l2 norm of weights: 5.875758072145682

Iteration Number: 5496
Loss: 35.09571585732592
l2 norm of gradients: 0.24200435393412806
l2 norm of weights: 5.8756742309624554

Iteration Number: 5497
Loss: 35.0927787151529
l2 norm of gradients: 0.24197396103184077
l2 norm of weights: 5.8755903976817345

Iteration Number: 5498
Loss: 35.08984231018978
l2 norm of gradients: 0.24194357935044725
l2 norm of weights: 5.875506572301795

Iteration Number: 5499
Loss: 35.0869066420611
l2 norm of gradients: 0.24191320888196677
l2 norm of weights: 5.875422754820918

Iteration Number: 5500
Loss: 35.08397171042821
l2 norm of gradients: 0.2418828496184258
l2 norm of weights: 5.87533894523738

Iteration Number: 5501
Loss: 35.081037514920055
l2 norm of gradients: 0.24185250155185864
l2 norm of weights: 5.875255143549464

Iteration Number: 5502
Loss: 35.07810405517243
l2 norm of gradients: 0.24182216467430714

```

12 norm of weights: 5.875171349755449
-----
Iteration Number: 5503
Loss: 35.0751713308258
12 norm of gradients: 0.24179183897782058
12 norm of weights: 5.875087563853618
-----
Iteration Number: 5504
Loss: 35.07223934152172
12 norm of gradients: 0.24176152445445603
12 norm of weights: 5.875003785842254
-----
Iteration Number: 5505
Loss: 35.06930808689389
12 norm of gradients: 0.24173122109627782
12 norm of weights: 5.874920015719641
-----
Iteration Number: 5506
Loss: 35.066377566593374
12 norm of gradients: 0.24170092889535816
12 norm of weights: 5.874836253484063
-----
Iteration Number: 5507
Loss: 35.06344778025239
12 norm of gradients: 0.24167064784377656
12 norm of weights: 5.874752499133805
-----
Iteration Number: 5508
Loss: 35.060518727509276
12 norm of gradients: 0.24164037793362014
12 norm of weights: 5.874668752667155
-----
Iteration Number: 5509
Loss: 35.057590408015535
12 norm of gradients: 0.2416101191569837
12 norm of weights: 5.874585014082399
-----
Iteration Number: 5510
Loss: 35.054662821402665
12 norm of gradients: 0.2415798715059692
12 norm of weights: 5.874501283377826
-----
Iteration Number: 5511
Loss: 35.05173596731159
12 norm of gradients: 0.2415496349726865
12 norm of weights: 5.874417560551723
-----
Iteration Number: 5512

```

```

Loss: 35.048809845393166
l2 norm of gradients: 0.24151940954925272
l2 norm of weights: 5.874333845602382
-----
Iteration Number: 5513
Loss: 35.045884455276806
l2 norm of gradients: 0.24148919522779264
l2 norm of weights: 5.874250138528092
-----
Iteration Number: 5514
Loss: 35.0429597966153
l2 norm of gradients: 0.2414589920004383
l2 norm of weights: 5.874166439327145
-----
Iteration Number: 5515
Loss: 35.04003586904854
l2 norm of gradients: 0.24142879985932944
l2 norm of weights: 5.874082747997835
-----
Iteration Number: 5516
Loss: 35.03711267221253
l2 norm of gradients: 0.2413986187966133
l2 norm of weights: 5.873999064538453
-----
Iteration Number: 5517
Loss: 35.0341902057651
l2 norm of gradients: 0.24136844880444439
l2 norm of weights: 5.873915388947294
-----
Iteration Number: 5518
Loss: 35.03126846933453
l2 norm of gradients: 0.24133828987498485
l2 norm of weights: 5.873831721222653
-----
Iteration Number: 5519
Loss: 35.02834746257125
l2 norm of gradients: 0.24130814200040412
l2 norm of weights: 5.873748061362825
-----
Iteration Number: 5520
Loss: 35.02542718511892
l2 norm of gradients: 0.24127800517287917
l2 norm of weights: 5.873664409366109
-----
Iteration Number: 5521
Loss: 35.02250763662634
l2 norm of gradients: 0.24124787938459458
l2 norm of weights: 5.873580765230799

```

```

-----
Iteration Number: 5522
Loss: 35.0195888167288
l2 norm of gradients: 0.241217764627742
l2 norm of weights: 5.873497128955195
-----
Iteration Number: 5523
Loss: 35.01667072507036
l2 norm of gradients: 0.2411876608945208
l2 norm of weights: 5.873413500537597
-----
Iteration Number: 5524
Loss: 35.013753361313924
l2 norm of gradients: 0.24115756817713763
l2 norm of weights: 5.873329879976305
-----
Iteration Number: 5525
Loss: 35.0108367250851
l2 norm of gradients: 0.24112748646780663
l2 norm of weights: 5.873246267269619
-----
Iteration Number: 5526
Loss: 35.00792081603673
l2 norm of gradients: 0.24109741575874924
l2 norm of weights: 5.87316266241584
-----
Iteration Number: 5527
Loss: 35.005005633817916
l2 norm of gradients: 0.24106735604219434
l2 norm of weights: 5.873079065413272
-----
Iteration Number: 5528
Loss: 35.00209117807
l2 norm of gradients: 0.24103730731037826
l2 norm of weights: 5.872995476260218
-----
Iteration Number: 5529
Loss: 34.99917744844242
l2 norm of gradients: 0.24100726955554472
l2 norm of weights: 5.872911894954982
-----
Iteration Number: 5530
Loss: 34.99626444457829
l2 norm of gradients: 0.24097724276994456
l2 norm of weights: 5.87282832149587
-----
Iteration Number: 5531
Loss: 34.993352166128275

```

```

12 norm of gradients: 0.24094722694583634
12 norm of weights: 5.872744755881187
-----
Iteration Number: 5532
Loss: 34.990440612738965
12 norm of gradients: 0.24091722207548574
12 norm of weights: 5.87266119810924
-----
Iteration Number: 5533
Loss: 34.987529784062886
12 norm of gradients: 0.24088722815116595
12 norm of weights: 5.872577648178338
-----
Iteration Number: 5534
Loss: 34.984619679735005
12 norm of gradients: 0.24085724516515725
12 norm of weights: 5.872494106086787
-----
Iteration Number: 5535
Loss: 34.981710299416974
12 norm of gradients: 0.24082727310974755
12 norm of weights: 5.872410571832898
-----
Iteration Number: 5536
Loss: 34.97880164274551
12 norm of gradients: 0.24079731197723198
12 norm of weights: 5.8723270454149805
-----
Iteration Number: 5537
Loss: 34.97589370938276
12 norm of gradients: 0.2407673617599129
12 norm of weights: 5.872243526831346
-----
Iteration Number: 5538
Loss: 34.97298649897012
12 norm of gradients: 0.24073742245010005
12 norm of weights: 5.872160016080306
-----
Iteration Number: 5539
Loss: 34.97008001115677
12 norm of gradients: 0.24070749404011055
12 norm of weights: 5.872076513160173
-----
Iteration Number: 5540
Loss: 34.96717424558594
12 norm of gradients: 0.24067757652226876
12 norm of weights: 5.871993018069261
-----

```


Iteration Number: 5541
Loss: 34.96426920192161
l2 norm of gradients: 0.24064766988890618
l2 norm of weights: 5.871909530805885

Iteration Number: 5542
Loss: 34.96136487980731
l2 norm of gradients: 0.24061777413236193
l2 norm of weights: 5.871826051368359

Iteration Number: 5543
Loss: 34.958461278891996
l2 norm of gradients: 0.24058788924498212
l2 norm of weights: 5.871742579754999

Iteration Number: 5544
Loss: 34.95555839882961
l2 norm of gradients: 0.2405580152191203
l2 norm of weights: 5.871659115964123

Iteration Number: 5545
Loss: 34.952656239258495
l2 norm of gradients: 0.24052815204713704
l2 norm of weights: 5.871575659994047

Iteration Number: 5546
Loss: 34.94975479984708
l2 norm of gradients: 0.24049829972140055
l2 norm of weights: 5.8714922118430914

Iteration Number: 5547
Loss: 34.946854080248215
l2 norm of gradients: 0.24046845823428592
l2 norm of weights: 5.871408771509574

Iteration Number: 5548
Loss: 34.94395408009163
l2 norm of gradients: 0.24043862757817572
l2 norm of weights: 5.8713253389918165

Iteration Number: 5549
Loss: 34.9410547990495
l2 norm of gradients: 0.24040880774545964
l2 norm of weights: 5.871241914288139

Iteration Number: 5550
Loss: 34.938156236770794
l2 norm of gradients: 0.24037899872853463

```

12 norm of weights: 5.871158497396863
-----
Iteration Number: 5551
Loss: 34.93525839290584
12 norm of gradients: 0.24034920051980488
12 norm of weights: 5.871075088316311
-----
Iteration Number: 5552
Loss: 34.932361267102216
12 norm of gradients: 0.24031941311168176
12 norm of weights: 5.870991687044808
-----
Iteration Number: 5553
Loss: 34.929464859020925
12 norm of gradients: 0.2402896364965838
12 norm of weights: 5.870908293580677
-----
Iteration Number: 5554
Loss: 34.92656916831209
12 norm of gradients: 0.24025987066693694
12 norm of weights: 5.870824907922244
-----
Iteration Number: 5555
Loss: 34.92367419462751
12 norm of gradients: 0.240230115615174
12 norm of weights: 5.870741530067835
-----
Iteration Number: 5556
Loss: 34.92077993762343
12 norm of gradients: 0.24020037133373517
12 norm of weights: 5.8706581600157755
-----
Iteration Number: 5557
Loss: 34.91788639695345
12 norm of gradients: 0.2401706378150679
12 norm of weights: 5.870574797764394
-----
Iteration Number: 5558
Loss: 34.91499357227999
12 norm of gradients: 0.24014091505162657
12 norm of weights: 5.8704914433120186
-----
Iteration Number: 5559
Loss: 34.9121014632389
12 norm of gradients: 0.24011120303587294
12 norm of weights: 5.87040809665698
-----
Iteration Number: 5560

```

Loss: 34.909210069507765
l2 norm of gradients: 0.2400815017602757
l2 norm of weights: 5.870324757797606

Iteration Number: 5561
Loss: 34.906319390726615
l2 norm of gradients: 0.24005181121731098
l2 norm of weights: 5.870241426732229

Iteration Number: 5562
Loss: 34.90342942655661
l2 norm of gradients: 0.24002213139946182
l2 norm of weights: 5.870158103459181

Iteration Number: 5563
Loss: 34.900540176652086
l2 norm of gradients: 0.2399924622992185
l2 norm of weights: 5.870074787976792

Iteration Number: 5564
Loss: 34.89765164067204
l2 norm of gradients: 0.23996280390907834
l2 norm of weights: 5.8699914802833995

Iteration Number: 5565
Loss: 34.894763818267165
l2 norm of gradients: 0.23993315622154596
l2 norm of weights: 5.869908180377334

Iteration Number: 5566
Loss: 34.89187670909874
l2 norm of gradients: 0.23990351922913283
l2 norm of weights: 5.869824888256933

Iteration Number: 5567
Loss: 34.888990312823296
l2 norm of gradients: 0.23987389292435768
l2 norm of weights: 5.869741603920531

Iteration Number: 5568
Loss: 34.886104629099584
l2 norm of gradients: 0.2398442772997464
l2 norm of weights: 5.869658327366465

Iteration Number: 5569
Loss: 34.88321965757989
l2 norm of gradients: 0.2398146723478318
l2 norm of weights: 5.869575058593072

```

-----
Iteration Number: 5570
Loss: 34.88033539792815
l2 norm of gradients: 0.23978507806115398
l2 norm of weights: 5.869491797598691
-----
Iteration Number: 5571
Loss: 34.87745184979759
l2 norm of gradients: 0.23975549443225994
l2 norm of weights: 5.869408544381661
-----
Iteration Number: 5572
Loss: 34.87456901284387
l2 norm of gradients: 0.23972592145370378
l2 norm of weights: 5.86932529894032
-----
Iteration Number: 5573
Loss: 34.87168688673746
l2 norm of gradients: 0.23969635911804674
l2 norm of weights: 5.869242061273011
-----
Iteration Number: 5574
Loss: 34.868805471126876
l2 norm of gradients: 0.23966680741785715
l2 norm of weights: 5.869158831378076
-----
Iteration Number: 5575
Loss: 34.86592476567424
l2 norm of gradients: 0.23963726634571014
l2 norm of weights: 5.869075609253855
-----
Iteration Number: 5576
Loss: 34.86304477003608
l2 norm of gradients: 0.2396077358941882
l2 norm of weights: 5.868992394898693
-----
Iteration Number: 5577
Loss: 34.86016548388041
l2 norm of gradients: 0.23957821605588067
l2 norm of weights: 5.868909188310932
-----
Iteration Number: 5578
Loss: 34.857286906857524
l2 norm of gradients: 0.23954870682338394
l2 norm of weights: 5.868825989488918
-----
Iteration Number: 5579
Loss: 34.85440903862972

```

```

12 norm of gradients: 0.2395192081893015
12 norm of weights: 5.868742798430997
-----
Iteration Number: 5580
Loss: 34.851531878862865
12 norm of gradients: 0.23948972014624376
12 norm of weights: 5.868659615135514
-----
Iteration Number: 5581
Loss: 34.848655427211675
12 norm of gradients: 0.23946024268682808
12 norm of weights: 5.868576439600817
-----
Iteration Number: 5582
Loss: 34.845779683340055
12 norm of gradients: 0.23943077580367905
12 norm of weights: 5.868493271825255
-----
Iteration Number: 5583
Loss: 34.842904646915876
12 norm of gradients: 0.239401319489428
12 norm of weights: 5.868410111807174
-----
Iteration Number: 5584
Loss: 34.84003031758368
12 norm of gradients: 0.2393718737367135
12 norm of weights: 5.868326959544926
-----
Iteration Number: 5585
Loss: 34.83715669502465
12 norm of gradients: 0.23934243853818077
12 norm of weights: 5.868243815036861
-----
Iteration Number: 5586
Loss: 34.834283778878394
12 norm of gradients: 0.23931301388648243
12 norm of weights: 5.868160678281329
-----
Iteration Number: 5587
Loss: 34.83141156883312
12 norm of gradients: 0.23928359977427757
12 norm of weights: 5.868077549276682
-----
Iteration Number: 5588
Loss: 34.82854006452794
12 norm of gradients: 0.23925419619423266
12 norm of weights: 5.867994428021274
-----

```

Iteration Number: 5589
Loss: 34.825669265646866
l2 norm of gradients: 0.2392248031390209
l2 norm of weights: 5.867911314513457

Iteration Number: 5590
Loss: 34.8227991718313
l2 norm of gradients: 0.23919542060132248
l2 norm of weights: 5.867828208751589

Iteration Number: 5591
Loss: 34.819929782766934
l2 norm of gradients: 0.2391660485738246
l2 norm of weights: 5.8677451107340195

Iteration Number: 5592
Loss: 34.81706109809878
l2 norm of gradients: 0.23913668704922122
l2 norm of weights: 5.867662020459109

Iteration Number: 5593
Loss: 34.814193117499954
l2 norm of gradients: 0.23910733602021342
l2 norm of weights: 5.867578937925213

Iteration Number: 5594
Loss: 34.8113258406356
l2 norm of gradients: 0.239077995479509
l2 norm of weights: 5.867495863130688

Iteration Number: 5595
Loss: 34.80845926716103
l2 norm of gradients: 0.23904866541982295
l2 norm of weights: 5.867412796073893

Iteration Number: 5596
Loss: 34.805593396752954
l2 norm of gradients: 0.23901934583387682
l2 norm of weights: 5.867329736753186

Iteration Number: 5597
Loss: 34.80272822906652
l2 norm of gradients: 0.23899003671439936
l2 norm of weights: 5.86724668516693

Iteration Number: 5598
Loss: 34.79986376376972
l2 norm of gradients: 0.238960738054126

```

12 norm of weights: 5.867163641313482
-----
Iteration Number: 5599
Loss: 34.79700000053319
12 norm of gradients: 0.23893144984579906
12 norm of weights: 5.867080605191207
-----
Iteration Number: 5600
Loss: 34.794136939019815
12 norm of gradients: 0.23890217208216782
12 norm of weights: 5.8669975767984655
-----
Iteration Number: 5601
Loss: 34.79127457889324
12 norm of gradients: 0.23887290475598855
12 norm of weights: 5.86691455613362
-----
Iteration Number: 5602
Loss: 34.7884129198169
12 norm of gradients: 0.23884364786002407
12 norm of weights: 5.866831543195035
-----
Iteration Number: 5603
Loss: 34.78555196146547
12 norm of gradients: 0.2388144013870443
12 norm of weights: 5.866748537981076
-----
Iteration Number: 5604
Loss: 34.782691703498074
12 norm of gradients: 0.23878516532982594
12 norm of weights: 5.866665540490107
-----
Iteration Number: 5605
Loss: 34.77983214558596
12 norm of gradients: 0.23875593968115252
12 norm of weights: 5.866582550720495
-----
Iteration Number: 5606
Loss: 34.776973287398825
12 norm of gradients: 0.23872672443381435
12 norm of weights: 5.866499568670608
-----
Iteration Number: 5607
Loss: 34.77411512859661
12 norm of gradients: 0.23869751958060872
12 norm of weights: 5.866416594338811
-----
Iteration Number: 5608

```

Loss: 34.77125766885924
l2 norm of gradients: 0.23866832511433955
l2 norm of weights: 5.866333627723476

Iteration Number: 5609
Loss: 34.76840090784176
l2 norm of gradients: 0.23863914102781764
l2 norm of weights: 5.86625066882297

Iteration Number: 5610
Loss: 34.765544845217654
l2 norm of gradients: 0.23860996731386072
l2 norm of weights: 5.866167717635664

Iteration Number: 5611
Loss: 34.76268948065327
l2 norm of gradients: 0.23858080396529316
l2 norm of weights: 5.866084774159929

Iteration Number: 5612
Loss: 34.759834813822145
l2 norm of gradients: 0.23855165097494624
l2 norm of weights: 5.866001838394136

Iteration Number: 5613
Loss: 34.756980844391336
l2 norm of gradients: 0.23852250833565788
l2 norm of weights: 5.865918910336659

Iteration Number: 5614
Loss: 34.75412757203047
l2 norm of gradients: 0.2384933760402729
l2 norm of weights: 5.86583598998587

Iteration Number: 5615
Loss: 34.75127499640336
l2 norm of gradients: 0.2384642540816429
l2 norm of weights: 5.865753077340144

Iteration Number: 5616
Loss: 34.74842311719179
l2 norm of gradients: 0.2384351424526262
l2 norm of weights: 5.865670172397855

Iteration Number: 5617
Loss: 34.74557193405465
l2 norm of gradients: 0.23840604114608785
l2 norm of weights: 5.865587275157378


```
-----
Iteration Number: 5618
Loss: 34.74272144666504
l2 norm of gradients: 0.23837695015489968
l2 norm of weights: 5.865504385617092
-----
Iteration Number: 5619
Loss: 34.739871654698035
l2 norm of gradients: 0.2383478694719403
l2 norm of weights: 5.865421503775371
-----
Iteration Number: 5620
Loss: 34.73702255782432
l2 norm of gradients: 0.2383187990900951
l2 norm of weights: 5.865338629630595
-----
Iteration Number: 5621
Loss: 34.734174155704785
l2 norm of gradients: 0.238289739002256
l2 norm of weights: 5.865255763181143
-----
Iteration Number: 5622
Loss: 34.73132644802408
l2 norm of gradients: 0.23826068920132185
l2 norm of weights: 5.865172904425393
-----
Iteration Number: 5623
Loss: 34.72847943444994
l2 norm of gradients: 0.23823164968019817
l2 norm of weights: 5.865090053361726
-----
Iteration Number: 5624
Loss: 34.72563311464746
l2 norm of gradients: 0.23820262043179713
l2 norm of weights: 5.865007209988523
-----
Iteration Number: 5625
Loss: 34.72278748829747
l2 norm of gradients: 0.23817360144903765
l2 norm of weights: 5.864924374304167
-----
Iteration Number: 5626
Loss: 34.719942555068116
l2 norm of gradients: 0.2381445927248453
l2 norm of weights: 5.864841546307038
-----
Iteration Number: 5627
Loss: 34.71709831463149
```

```

12 norm of gradients: 0.23811559425215245
12 norm of weights: 5.864758725995522
-----
Iteration Number: 5628
Loss: 34.7142547666623
12 norm of gradients: 0.2380866060238981
12 norm of weights: 5.864675913368001
-----
Iteration Number: 5629
Loss: 34.71141191083257
12 norm of gradients: 0.23805762803302785
12 norm of weights: 5.864593108422861
-----
Iteration Number: 5630
Loss: 34.70856974681512
12 norm of gradients: 0.238028660272494
12 norm of weights: 5.864510311158488
-----
Iteration Number: 5631
Loss: 34.70572827428468
12 norm of gradients: 0.2379997027352556
12 norm of weights: 5.864427521573269
-----
Iteration Number: 5632
Loss: 34.70288749292028
12 norm of gradients: 0.23797075541427845
12 norm of weights: 5.864344739665589
-----
Iteration Number: 5633
Loss: 34.700047402385245
12 norm of gradients: 0.23794181830253464
12 norm of weights: 5.864261965433838
-----
Iteration Number: 5634
Loss: 34.69720800236135
12 norm of gradients: 0.23791289139300323
12 norm of weights: 5.864179198876403
-----
Iteration Number: 5635
Loss: 34.69436929252091
12 norm of gradients: 0.23788397467866979
12 norm of weights: 5.864096439991676
-----
Iteration Number: 5636
Loss: 34.691531272538015
12 norm of gradients: 0.23785506815252652
12 norm of weights: 5.864013688778044
-----

```

Iteration Number: 5637
Loss: 34.688693942091156
l2 norm of gradients: 0.23782617180757232
l2 norm of weights: 5.863930945233901

Iteration Number: 5638
Loss: 34.68585730085383
l2 norm of gradients: 0.23779728563681263
l2 norm of weights: 5.863848209357638

Iteration Number: 5639
Loss: 34.683021348499885
l2 norm of gradients: 0.2377684096332595
l2 norm of weights: 5.863765481147647

Iteration Number: 5640
Loss: 34.680186084712275
l2 norm of gradients: 0.23773954378993165
l2 norm of weights: 5.863682760602321

Iteration Number: 5641
Loss: 34.677351509158086
l2 norm of gradients: 0.23771068809985432
l2 norm of weights: 5.863600047720055

Iteration Number: 5642
Loss: 34.67451762152098
l2 norm of gradients: 0.23768184255605948
l2 norm of weights: 5.863517342499243

Iteration Number: 5643
Loss: 34.671684421470914
l2 norm of gradients: 0.2376530071515855
l2 norm of weights: 5.863434644938282

Iteration Number: 5644
Loss: 34.66885190868973
l2 norm of gradients: 0.23762418187947737
l2 norm of weights: 5.863351955035567

Iteration Number: 5645
Loss: 34.66602008285102
l2 norm of gradients: 0.23759536673278678
l2 norm of weights: 5.863269272789496

Iteration Number: 5646
Loss: 34.6631889436341
l2 norm of gradients: 0.23756656170457183

```

12 norm of weights: 5.863186598198466
-----
Iteration Number: 5647
Loss: 34.66035849072011
12 norm of gradients: 0.23753776678789734
12 norm of weights: 5.863103931260876
-----
Iteration Number: 5648
Loss: 34.65752872378006
12 norm of gradients: 0.23750898197583448
12 norm of weights: 5.863021271975126
-----
Iteration Number: 5649
Loss: 34.65469964249867
12 norm of gradients: 0.23748020726146124
12 norm of weights: 5.862938620339615
-----
Iteration Number: 5650
Loss: 34.65187124655055
12 norm of gradients: 0.2374514426378618
12 norm of weights: 5.862855976352745
-----
Iteration Number: 5651
Loss: 34.64904353561444
12 norm of gradients: 0.23742268809812717
12 norm of weights: 5.862773340012918
-----
Iteration Number: 5652
Loss: 34.646216509368806
12 norm of gradients: 0.23739394363535468
12 norm of weights: 5.8626907113185345
-----
Iteration Number: 5653
Loss: 34.64339016749691
12 norm of gradients: 0.23736520924264837
12 norm of weights: 5.862608090268
-----
Iteration Number: 5654
Loss: 34.64056450966628
12 norm of gradients: 0.23733648491311862
12 norm of weights: 5.862525476859716
-----
Iteration Number: 5655
Loss: 34.63773953557592
12 norm of gradients: 0.23730777063988237
12 norm of weights: 5.86244287109209
-----
Iteration Number: 5656

```

Loss: 34.63491524489193
l2 norm of gradients: 0.2372790664160632
l2 norm of weights: 5.8623602729635245

Iteration Number: 5657
Loss: 34.63209163729411
l2 norm of gradients: 0.23725037223479095
l2 norm of weights: 5.8622776824724285

Iteration Number: 5658
Loss: 34.62926871247535
l2 norm of gradients: 0.23722168808920208
l2 norm of weights: 5.862195099617206

Iteration Number: 5659
Loss: 34.626446470100106
l2 norm of gradients: 0.2371930139724395
l2 norm of weights: 5.862112524396267

Iteration Number: 5660
Loss: 34.623624909856446
l2 norm of gradients: 0.23716434987765259
l2 norm of weights: 5.862029956808019

Iteration Number: 5661
Loss: 34.62080403142866
l2 norm of gradients: 0.23713569579799723
l2 norm of weights: 5.861947396850871

Iteration Number: 5662
Loss: 34.61798383449326
l2 norm of gradients: 0.23710705172663576
l2 norm of weights: 5.861864844523233

Iteration Number: 5663
Loss: 34.61516431872771
l2 norm of gradients: 0.2370784176567369
l2 norm of weights: 5.861782299823517

Iteration Number: 5664
Loss: 34.61234548382662
l2 norm of gradients: 0.2370497935814759
l2 norm of weights: 5.861699762750132

Iteration Number: 5665
Loss: 34.60952732946742
l2 norm of gradients: 0.23702117949403437
l2 norm of weights: 5.861617233301493

```

-----
Iteration Number: 5666
Loss: 34.606709855323416
l2 norm of gradients: 0.23699257538760043
l2 norm of weights: 5.8615347114760095
-----
Iteration Number: 5667
Loss: 34.60389306108617
l2 norm of gradients: 0.2369639812553686
l2 norm of weights: 5.861452197272098
-----
Iteration Number: 5668
Loss: 34.60107694643306
l2 norm of gradients: 0.23693539709053985
l2 norm of weights: 5.861369690688172
-----
Iteration Number: 5669
Loss: 34.59826151105585
l2 norm of gradients: 0.23690682288632142
l2 norm of weights: 5.861287191722646
-----
Iteration Number: 5670
Loss: 34.59544675463385
l2 norm of gradients: 0.23687825863592718
l2 norm of weights: 5.8612047003739365
-----
Iteration Number: 5671
Loss: 34.59263267684403
l2 norm of gradients: 0.23684970433257732
l2 norm of weights: 5.86112221664046
-----
Iteration Number: 5672
Loss: 34.589819277378346
l2 norm of gradients: 0.23682115996949832
l2 norm of weights: 5.861039740520633
-----
Iteration Number: 5673
Loss: 34.58700655591534
l2 norm of gradients: 0.23679262553992317
l2 norm of weights: 5.860957272012876
-----
Iteration Number: 5674
Loss: 34.584194512141174
l2 norm of gradients: 0.23676410103709114
l2 norm of weights: 5.860874811115605
-----
Iteration Number: 5675
Loss: 34.58138314573782

```

```

12 norm of gradients: 0.236735586454248
12 norm of weights: 5.860792357827241
-----
Iteration Number: 5676
Loss: 34.57857245639539
12 norm of gradients: 0.23670708178464583
12 norm of weights: 5.860709912146205
-----
Iteration Number: 5677
Loss: 34.57576244379714
12 norm of gradients: 0.236678587021543
12 norm of weights: 5.860627474070916
-----
Iteration Number: 5678
Loss: 34.57295310762572
12 norm of gradients: 0.23665010215820442
12 norm of weights: 5.860545043599797
-----
Iteration Number: 5679
Loss: 34.57014444756918
12 norm of gradients: 0.23662162718790114
12 norm of weights: 5.860462620731271
-----
Iteration Number: 5680
Loss: 34.567336463315314
12 norm of gradients: 0.23659316210391074
12 norm of weights: 5.86038020546376
-----
Iteration Number: 5681
Loss: 34.56452915454493
12 norm of gradients: 0.23656470689951695
12 norm of weights: 5.860297797795688
-----
Iteration Number: 5682
Loss: 34.56172252094877
12 norm of gradients: 0.23653626156800991
12 norm of weights: 5.860215397725482
-----
Iteration Number: 5683
Loss: 34.558916562204466
12 norm of gradients: 0.23650782610268628
12 norm of weights: 5.860133005251565
-----
Iteration Number: 5684
Loss: 34.55611127800624
12 norm of gradients: 0.2364794004968487
12 norm of weights: 5.8600506203723635
-----

```

Iteration Number: 5685
Loss: 34.55330666804308
l2 norm of gradients: 0.23645098474380633
l2 norm of weights: 5.859968243086305

Iteration Number: 5686
Loss: 34.550502731996794
l2 norm of gradients: 0.23642257883687465
l2 norm of weights: 5.859885873391819

Iteration Number: 5687
Loss: 34.547699469560136
l2 norm of gradients: 0.23639418276937532
l2 norm of weights: 5.859803511287329

Iteration Number: 5688
Loss: 34.54489688041598
l2 norm of gradients: 0.23636579653463638
l2 norm of weights: 5.859721156771268

Iteration Number: 5689
Loss: 34.54209496425242
l2 norm of gradients: 0.2363374201259921
l2 norm of weights: 5.859638809842065

Iteration Number: 5690
Loss: 34.53929372075738
l2 norm of gradients: 0.23630905353678308
l2 norm of weights: 5.859556470498151

Iteration Number: 5691
Loss: 34.53649314962134
l2 norm of gradients: 0.23628069676035615
l2 norm of weights: 5.859474138737955

Iteration Number: 5692
Loss: 34.53369325053184
l2 norm of gradients: 0.23625234979006451
l2 norm of weights: 5.859391814559912

Iteration Number: 5693
Loss: 34.53089402317849
l2 norm of gradients: 0.23622401261926748
l2 norm of weights: 5.859309497962453

Iteration Number: 5694
Loss: 34.52809546725065
l2 norm of gradients: 0.23619568524133072


```

12 norm of weights: 5.859227188944012
-----
Iteration Number: 5695
Loss: 34.52529758243669
12 norm of gradients: 0.23616736764962612
12 norm of weights: 5.859144887503023
-----
Iteration Number: 5696
Loss: 34.522500368424026
12 norm of gradients: 0.23613905983753178
12 norm of weights: 5.859062593637921
-----
Iteration Number: 5697
Loss: 34.51970382490527
12 norm of gradients: 0.23611076179843216
12 norm of weights: 5.858980307347141
-----
Iteration Number: 5698
Loss: 34.51690795156875
12 norm of gradients: 0.23608247352571776
12 norm of weights: 5.85889802862912
-----
Iteration Number: 5699
Loss: 34.514112748106754
12 norm of gradients: 0.2360541950127854
12 norm of weights: 5.858815757482296
-----
Iteration Number: 5700
Loss: 34.511318214206355
12 norm of gradients: 0.2360259262530382
12 norm of weights: 5.858733493905105
-----
Iteration Number: 5701
Loss: 34.508524349562435
12 norm of gradients: 0.23599766723988533
12 norm of weights: 5.8586512378959865
-----
Iteration Number: 5702
Loss: 34.50573115386237
12 norm of gradients: 0.23596941796674228
12 norm of weights: 5.85856898945338
-----
Iteration Number: 5703
Loss: 34.502938626798134
12 norm of gradients: 0.23594117842703075
12 norm of weights: 5.858486748575724
-----
Iteration Number: 5704

```

```

Loss: 34.50014676806462
l2 norm of gradients: 0.23591294861417847
l2 norm of weights: 5.858404515261461
-----
Iteration Number: 5705
Loss: 34.49735557734889
l2 norm of gradients: 0.2358847285216196
l2 norm of weights: 5.858322289509031
-----
Iteration Number: 5706
Loss: 34.49456505434325
l2 norm of gradients: 0.2358565181427943
l2 norm of weights: 5.858240071316877
-----
Iteration Number: 5707
Loss: 34.49177519873687
l2 norm of gradients: 0.23582831747114893
l2 norm of weights: 5.858157860683442
-----
Iteration Number: 5708
Loss: 34.488986010232274
l2 norm of gradients: 0.23580012650013604
l2 norm of weights: 5.858075657607168
-----
Iteration Number: 5709
Loss: 34.48619748851149
l2 norm of gradients: 0.2357719452232144
l2 norm of weights: 5.857993462086501
-----
Iteration Number: 5710
Loss: 34.48340963327415
l2 norm of gradients: 0.2357437736338488
l2 norm of weights: 5.8579112741198855
-----
Iteration Number: 5711
Loss: 34.48062244421497
l2 norm of gradients: 0.23571561172551037
l2 norm of weights: 5.857829093705767
-----
Iteration Number: 5712
Loss: 34.47783592101981
l2 norm of gradients: 0.23568745949167616
l2 norm of weights: 5.857746920842592
-----
Iteration Number: 5713
Loss: 34.47505006338018
l2 norm of gradients: 0.2356593169258295
l2 norm of weights: 5.857664755528807

```

```

-----
Iteration Number: 5714
Loss: 34.47226487100049
l2 norm of gradients: 0.23563118402145983
l2 norm of weights: 5.857582597762861
-----
Iteration Number: 5715
Loss: 34.46948034356776
l2 norm of gradients: 0.23560306077206275
l2 norm of weights: 5.8575004475432015
-----
Iteration Number: 5716
Loss: 34.46669648077124
l2 norm of gradients: 0.23557494717113983
l2 norm of weights: 5.857418304868279
-----
Iteration Number: 5717
Loss: 34.46391328231752
l2 norm of gradients: 0.23554684321219888
l2 norm of weights: 5.857336169736541
-----
Iteration Number: 5718
Loss: 34.46113074789365
l2 norm of gradients: 0.23551874888875393
l2 norm of weights: 5.857254042146442
-----
Iteration Number: 5719
Loss: 34.458348877197366
l2 norm of gradients: 0.23549066419432474
l2 norm of weights: 5.8571719220964304
-----
Iteration Number: 5720
Loss: 34.45556766992199
l2 norm of gradients: 0.23546258912243748
l2 norm of weights: 5.857089809584958
-----
Iteration Number: 5721
Loss: 34.45278712575849
l2 norm of gradients: 0.23543452366662437
l2 norm of weights: 5.85700770461048
-----
Iteration Number: 5722
Loss: 34.4500072444156
l2 norm of gradients: 0.23540646782042357
l2 norm of weights: 5.856925607171448
-----
Iteration Number: 5723
Loss: 34.44722802557339

```

```

12 norm of gradients: 0.23537842157737943
12 norm of weights: 5.8568435172663165
-----
Iteration Number: 5724
Loss: 34.444449468937776
12 norm of gradients: 0.2353503849310424
12 norm of weights: 5.856761434893542
-----
Iteration Number: 5725
Loss: 34.44167157420665
12 norm of gradients: 0.23532235787496886
12 norm of weights: 5.856679360051578
-----
Iteration Number: 5726
Loss: 34.43889434106691
12 norm of gradients: 0.23529434040272135
12 norm of weights: 5.856597292738882
-----
Iteration Number: 5727
Loss: 34.43611776922161
12 norm of gradients: 0.23526633250786838
12 norm of weights: 5.85651523295391
-----
Iteration Number: 5728
Loss: 34.433341858366475
12 norm of gradients: 0.23523833418398465
12 norm of weights: 5.856433180695121
-----
Iteration Number: 5729
Loss: 34.430566608194866
12 norm of gradients: 0.23521034542465075
12 norm of weights: 5.856351135960973
-----
Iteration Number: 5730
Loss: 34.427792018410415
12 norm of gradients: 0.2351823662234533
12 norm of weights: 5.856269098749926
-----
Iteration Number: 5731
Loss: 34.42501808871352
12 norm of gradients: 0.23515439657398501
12 norm of weights: 5.856187069060438
-----
Iteration Number: 5732
Loss: 34.422244818790986
12 norm of gradients: 0.23512643646984469
12 norm of weights: 5.85610504689097
-----

```

Iteration Number: 5733
Loss: 34.41947220834748
l2 norm of gradients: 0.23509848590463694
l2 norm of weights: 5.856023032239984

Iteration Number: 5734
Loss: 34.41670025708302
l2 norm of gradients: 0.23507054487197268
l2 norm of weights: 5.855941025105943

Iteration Number: 5735
Loss: 34.41392896468985
l2 norm of gradients: 0.23504261336546847
l2 norm of weights: 5.855859025487307

Iteration Number: 5736
Loss: 34.411158330870066
l2 norm of gradients: 0.2350146913787472
l2 norm of weights: 5.855777033382542

Iteration Number: 5737
Loss: 34.40838835532782
l2 norm of gradients: 0.23498677890543743
l2 norm of weights: 5.855695048790109

Iteration Number: 5738
Loss: 34.40561903775614
l2 norm of gradients: 0.23495887593917397
l2 norm of weights: 5.855613071708476

Iteration Number: 5739
Loss: 34.4028503778536
l2 norm of gradients: 0.23493098247359745
l2 norm of weights: 5.855531102136106

Iteration Number: 5740
Loss: 34.40008237532288
l2 norm of gradients: 0.23490309850235463
l2 norm of weights: 5.855449140071466

Iteration Number: 5741
Loss: 34.39731502985857
l2 norm of gradients: 0.23487522401909802
l2 norm of weights: 5.855367185513023

Iteration Number: 5742
Loss: 34.39454834117134
l2 norm of gradients: 0.23484735901748632

```

12 norm of weights: 5.855285238459245
-----
Iteration Number: 5743
Loss: 34.39178230895054
12 norm of gradients: 0.2348195034911839
12 norm of weights: 5.855203298908599
-----
Iteration Number: 5744
Loss: 34.38901693290097
12 norm of gradients: 0.23479165743386132
12 norm of weights: 5.855121366859555
-----
Iteration Number: 5745
Loss: 34.386252212730156
12 norm of gradients: 0.23476382083919514
12 norm of weights: 5.855039442310582
-----
Iteration Number: 5746
Loss: 34.38348814812527
12 norm of gradients: 0.23473599370086742
12 norm of weights: 5.85495752526015
-----
Iteration Number: 5747
Loss: 34.38072473880402
12 norm of gradients: 0.23470817601256672
12 norm of weights: 5.854875615706731
-----
Iteration Number: 5748
Loss: 34.37796198444846
12 norm of gradients: 0.23468036776798717
12 norm of weights: 5.854793713648795
-----
Iteration Number: 5749
Loss: 34.37519988477583
12 norm of gradients: 0.2346525689608288
12 norm of weights: 5.854711819084817
-----
Iteration Number: 5750
Loss: 34.37243843948482
12 norm of gradients: 0.23462477958479783
12 norm of weights: 5.854629932013268
-----
Iteration Number: 5751
Loss: 34.36967764827061
12 norm of gradients: 0.23459699963360597
12 norm of weights: 5.854548052432623
-----
Iteration Number: 5752

```

Loss: 34.366917510836984
l2 norm of gradients: 0.23456922910097122
l2 norm of weights: 5.854466180341355

Iteration Number: 5753
Loss: 34.36415802689546
l2 norm of gradients: 0.23454146798061729
l2 norm of weights: 5.854384315737939

Iteration Number: 5754
Loss: 34.36139919614344
l2 norm of gradients: 0.23451371626627376
l2 norm of weights: 5.854302458620853

Iteration Number: 5755
Loss: 34.358641018282704
l2 norm of gradients: 0.2344859739516762
l2 norm of weights: 5.854220608988573

Iteration Number: 5756
Loss: 34.355883493015654
l2 norm of gradients: 0.23445824103056595
l2 norm of weights: 5.854138766839574

Iteration Number: 5757
Loss: 34.35312662005038
l2 norm of gradients: 0.2344305174966902
l2 norm of weights: 5.854056932172337

Iteration Number: 5758
Loss: 34.35037039908947
l2 norm of gradients: 0.23440280334380217
l2 norm of weights: 5.853975104985337

Iteration Number: 5759
Loss: 34.34761482983113
l2 norm of gradients: 0.23437509856566077
l2 norm of weights: 5.8538932852770555

Iteration Number: 5760
Loss: 34.34485991197984
l2 norm of gradients: 0.2343474031560309
l2 norm of weights: 5.853811473045973

Iteration Number: 5761
Loss: 34.342105645246654
l2 norm of gradients: 0.23431971710868307
l2 norm of weights: 5.853729668290568

```

-----
Iteration Number: 5762
Loss: 34.33935202933613
l2 norm of gradients: 0.2342920404173939
l2 norm of weights: 5.853647871009324
-----
Iteration Number: 5763
Loss: 34.33659906394478
l2 norm of gradients: 0.23426437307594575
l2 norm of weights: 5.8535660812007215
-----
Iteration Number: 5764
Loss: 34.33384674878417
l2 norm of gradients: 0.23423671507812666
l2 norm of weights: 5.853484298863243
-----
Iteration Number: 5765
Loss: 34.33109508355503
l2 norm of gradients: 0.2342090664177307
l2 norm of weights: 5.853402523995373
-----
Iteration Number: 5766
Loss: 34.32834406796682
l2 norm of gradients: 0.2341814270885577
l2 norm of weights: 5.853320756595595
-----
Iteration Number: 5767
Loss: 34.32559370172771
l2 norm of gradients: 0.23415379708441328
l2 norm of weights: 5.8532389966623946
-----
Iteration Number: 5768
Loss: 34.322843984534416
l2 norm of gradients: 0.2341261763991087
l2 norm of weights: 5.853157244194255
-----
Iteration Number: 5769
Loss: 34.32009491610208
l2 norm of gradients: 0.23409856502646137
l2 norm of weights: 5.853075499189664
-----
Iteration Number: 5770
Loss: 34.31734649613089
l2 norm of gradients: 0.23407096296029425
l2 norm of weights: 5.852993761647109
-----
Iteration Number: 5771
Loss: 34.31459872432802

```



```

12 norm of gradients: 0.23404337019443613
12 norm of weights: 5.852912031565075
-----
Iteration Number: 5772
Loss: 34.311851600407095
12 norm of gradients: 0.23401578672272152
12 norm of weights: 5.8528303089420515
-----
Iteration Number: 5773
Loss: 34.30910512406725
12 norm of gradients: 0.23398821253899094
12 norm of weights: 5.852748593776528
-----
Iteration Number: 5774
Loss: 34.30635929501284
12 norm of gradients: 0.23396064763709026
12 norm of weights: 5.8526668860669915
-----
Iteration Number: 5775
Loss: 34.30361411296339
12 norm of gradients: 0.23393309201087156
12 norm of weights: 5.852585185811934
-----
Iteration Number: 5776
Loss: 34.30086957761271
12 norm of gradients: 0.23390554565419241
12 norm of weights: 5.852503493009847
-----
Iteration Number: 5777
Loss: 34.298125688681345
12 norm of gradients: 0.2338780085609163
12 norm of weights: 5.85242180765922
-----
Iteration Number: 5778
Loss: 34.29538244587186
12 norm of gradients: 0.2338504807249123
12 norm of weights: 5.852340129758547
-----
Iteration Number: 5779
Loss: 34.29263984888952
12 norm of gradients: 0.23382296214005516
12 norm of weights: 5.85225845930632
-----
Iteration Number: 5780
Loss: 34.28989789744326
12 norm of gradients: 0.2337954528002257
12 norm of weights: 5.852176796301032
-----

```

Iteration Number: 5781
Loss: 34.287156591248284
l2 norm of gradients: 0.2337679526993102
l2 norm of weights: 5.852095140741178

Iteration Number: 5782
Loss: 34.28441593000547
l2 norm of gradients: 0.23374046183120065
l2 norm of weights: 5.852013492625252

Iteration Number: 5783
Loss: 34.281675913426696
l2 norm of gradients: 0.2337129801897949
l2 norm of weights: 5.85193185195175

Iteration Number: 5784
Loss: 34.27893654122288
l2 norm of gradients: 0.23368550776899644
l2 norm of weights: 5.851850218719168

Iteration Number: 5785
Loss: 34.276197813106066
l2 norm of gradients: 0.23365804456271444
l2 norm of weights: 5.851768592926004

Iteration Number: 5786
Loss: 34.273459728781816
l2 norm of gradients: 0.2336305905648638
l2 norm of weights: 5.851686974570753

Iteration Number: 5787
Loss: 34.27072228795819
l2 norm of gradients: 0.2336031457693651
l2 norm of weights: 5.851605363651916

Iteration Number: 5788
Loss: 34.2679854903465
l2 norm of gradients: 0.2335757101701447
l2 norm of weights: 5.851523760167989

Iteration Number: 5789
Loss: 34.26524933566312
l2 norm of gradients: 0.2335482837611344
l2 norm of weights: 5.851442164117475

Iteration Number: 5790
Loss: 34.26251382361079
l2 norm of gradients: 0.23352086653627202

```

12 norm of weights: 5.8513605754988705
-----
Iteration Number: 5791
Loss: 34.25977895391023
12 norm of gradients: 0.23349345848950076
12 norm of weights: 5.851278994310679
-----
Iteration Number: 5792
Loss: 34.257044726259494
12 norm of gradients: 0.23346605961476963
12 norm of weights: 5.8511974205514
-----
Iteration Number: 5793
Loss: 34.25431114037283
12 norm of gradients: 0.23343866990603315
12 norm of weights: 5.851115854219537
-----
Iteration Number: 5794
Loss: 34.25157819597078
12 norm of gradients: 0.23341128935725175
12 norm of weights: 5.851034295313593
-----
Iteration Number: 5795
Loss: 34.24884589275943
12 norm of gradients: 0.2333839179623913
12 norm of weights: 5.850952743832071
-----
Iteration Number: 5796
Loss: 34.246114230450736
12 norm of gradients: 0.23335655571542344
12 norm of weights: 5.850871199773475
-----
Iteration Number: 5797
Loss: 34.243383208756256
12 norm of gradients: 0.23332920261032525
12 norm of weights: 5.85078966313631
-----
Iteration Number: 5798
Loss: 34.24065282738771
12 norm of gradients: 0.23330185864107963
12 norm of weights: 5.850708133919082
-----
Iteration Number: 5799
Loss: 34.23792308606
12 norm of gradients: 0.23327452380167513
12 norm of weights: 5.850626612120296
-----
Iteration Number: 5800

```

Loss: 34.23519398448556
l2 norm of gradients: 0.23324719808610567
l2 norm of weights: 5.850545097738459

Iteration Number: 5801
Loss: 34.23246552237018
l2 norm of gradients: 0.23321988148837103
l2 norm of weights: 5.850463590772081

Iteration Number: 5802
Loss: 34.229737699440065
l2 norm of gradients: 0.23319257400247653
l2 norm of weights: 5.850382091219667

Iteration Number: 5803
Loss: 34.22701051540024
l2 norm of gradients: 0.23316527562243305
l2 norm of weights: 5.850300599079726

Iteration Number: 5804
Loss: 34.22428396996199
l2 norm of gradients: 0.23313798634225708
l2 norm of weights: 5.850219114350769

Iteration Number: 5805
Loss: 34.22155806284595
l2 norm of gradients: 0.23311070615597076
l2 norm of weights: 5.850137637031305

Iteration Number: 5806
Loss: 34.21883279376813
l2 norm of gradients: 0.23308343505760176
l2 norm of weights: 5.850056167119846

Iteration Number: 5807
Loss: 34.2161081624273
l2 norm of gradients: 0.23305617304118326
l2 norm of weights: 5.849974704614902

Iteration Number: 5808
Loss: 34.21338416855777
l2 norm of gradients: 0.23302892010075416
l2 norm of weights: 5.849893249514986

Iteration Number: 5809
Loss: 34.21066081185527
l2 norm of gradients: 0.23300167623035883
l2 norm of weights: 5.8498118018186105

```

-----
Iteration Number: 5810
Loss: 34.20793809204652
l2 norm of gradients: 0.2329744414240472
l2 norm of weights: 5.849730361524288
-----
Iteration Number: 5811
Loss: 34.20521600884516
l2 norm of gradients: 0.23294721567587487
l2 norm of weights: 5.849648928630534
-----
Iteration Number: 5812
Loss: 34.202494561967846
l2 norm of gradients: 0.23291999897990281
l2 norm of weights: 5.849567503135862
-----
Iteration Number: 5813
Loss: 34.199773751121676
l2 norm of gradients: 0.23289279133019772
l2 norm of weights: 5.849486085038788
-----
Iteration Number: 5814
Loss: 34.19705357603302
l2 norm of gradients: 0.2328655927208317
l2 norm of weights: 5.849404674337829
-----
Iteration Number: 5815
Loss: 34.19433403641211
l2 norm of gradients: 0.2328384031458824
l2 norm of weights: 5.8493232710314995
-----
Iteration Number: 5816
Loss: 34.191615131975354
l2 norm of gradients: 0.23281122259943313
l2 norm of weights: 5.8492418751183175
-----
Iteration Number: 5817
Loss: 34.18889686244292
l2 norm of gradients: 0.2327840510755725
l2 norm of weights: 5.849160486596801
-----
Iteration Number: 5818
Loss: 34.186179227522594
l2 norm of gradients: 0.23275688856839488
l2 norm of weights: 5.849079105465471
-----
Iteration Number: 5819
Loss: 34.1834622269362

```

```

12 norm of gradients: 0.23272973507199998
12 norm of weights: 5.848997731722843
-----
Iteration Number: 5820
Loss: 34.18074586040202
12 norm of gradients: 0.2327025905804931
12 norm of weights: 5.848916365367439
-----
Iteration Number: 5821
Loss: 34.17803012763556
12 norm of gradients: 0.23267545508798498
12 norm of weights: 5.848835006397779
-----
Iteration Number: 5822
Loss: 34.17531502835451
12 norm of gradients: 0.23264832858859189
12 norm of weights: 5.848753654812385
-----
Iteration Number: 5823
Loss: 34.17260056227696
12 norm of gradients: 0.23262121107643563
12 norm of weights: 5.848672310609777
-----
Iteration Number: 5824
Loss: 34.16988672912108
12 norm of gradients: 0.23259410254564328
12 norm of weights: 5.848590973788478
-----
Iteration Number: 5825
Loss: 34.16717352860482
12 norm of gradients: 0.23256700299034777
12 norm of weights: 5.848509644347014
-----
Iteration Number: 5826
Loss: 34.16446096044492
12 norm of gradients: 0.23253991240468713
12 norm of weights: 5.848428322283905
-----
Iteration Number: 5827
Loss: 34.16174902435662
12 norm of gradients: 0.23251283078280507
12 norm of weights: 5.848347007597678
-----
Iteration Number: 5828
Loss: 34.15903772006583
12 norm of gradients: 0.23248575811885078
12 norm of weights: 5.848265700286856
-----

```

Iteration Number: 5829
Loss: 34.156327047286936
l2 norm of gradients: 0.23245869440697875
l2 norm of weights: 5.8481844003499655

Iteration Number: 5830
Loss: 34.153617005738305
l2 norm of gradients: 0.23243163964134894
l2 norm of weights: 5.848103107785534

Iteration Number: 5831
Loss: 34.15090759513961
l2 norm of gradients: 0.23240459381612694
l2 norm of weights: 5.848021822592088

Iteration Number: 5832
Loss: 34.14819881521654
l2 norm of gradients: 0.2323775569254836
l2 norm of weights: 5.847940544768154

Iteration Number: 5833
Loss: 34.145490665678395
l2 norm of gradients: 0.23235052896359526
l2 norm of weights: 5.8478592743122615

Iteration Number: 5834
Loss: 34.14278314624643
l2 norm of gradients: 0.23232350992464373
l2 norm of weights: 5.847778011222939

Iteration Number: 5835
Loss: 34.14007625664735
l2 norm of gradients: 0.2322964998028161
l2 norm of weights: 5.847696755498716

Iteration Number: 5836
Loss: 34.13736999660701
l2 norm of gradients: 0.23226949859230508
l2 norm of weights: 5.847615507138123

Iteration Number: 5837
Loss: 34.13466436582745
l2 norm of gradients: 0.23224250628730866
l2 norm of weights: 5.84753426613969

Iteration Number: 5838
Loss: 34.131959364038806
l2 norm of gradients: 0.23221552288203023

```
12 norm of weights: 5.84745303250195
-----
Iteration Number: 5839
Loss: 34.12925499096491
12 norm of gradients: 0.23218854837067862
12 norm of weights: 5.8473718062234346
-----
Iteration Number: 5840
Loss: 34.12655124632011
12 norm of gradients: 0.2321615827474681
12 norm of weights: 5.847290587302677
-----
Iteration Number: 5841
Loss: 34.12384812982795
12 norm of gradients: 0.23213462600661822
12 norm of weights: 5.847209375738208
-----
Iteration Number: 5842
Loss: 34.12114564121964
12 norm of gradients: 0.23210767814235403
12 norm of weights: 5.847128171528564
-----
Iteration Number: 5843
Loss: 34.118443780202064
12 norm of gradients: 0.23208073914890587
12 norm of weights: 5.84704697467228
-----
Iteration Number: 5844
Loss: 34.11574254650479
12 norm of gradients: 0.2320538090205094
12 norm of weights: 5.84696578516789
-----
Iteration Number: 5845
Loss: 34.11304193984699
12 norm of gradients: 0.23202688775140587
12 norm of weights: 5.8468846030139305
-----
Iteration Number: 5846
Loss: 34.110341959955754
12 norm of gradients: 0.23199997533584177
12 norm of weights: 5.846803428208939
-----
Iteration Number: 5847
Loss: 34.107642606546925
12 norm of gradients: 0.2319730717680688
12 norm of weights: 5.84672226075145
-----
Iteration Number: 5848
```


Loss: 34.10494387934833
l2 norm of gradients: 0.23194617704234421
l2 norm of weights: 5.846641100640005

Iteration Number: 5849
Loss: 34.10224577807457
l2 norm of gradients: 0.23191929115293053
l2 norm of weights: 5.84655994787314

Iteration Number: 5850
Loss: 34.09954830246169
l2 norm of gradients: 0.23189241409409558
l2 norm of weights: 5.846478802449395

Iteration Number: 5851
Loss: 34.096851452226026
l2 norm of gradients: 0.23186554586011265
l2 norm of weights: 5.846397664367308

Iteration Number: 5852
Loss: 34.094155227086574
l2 norm of gradients: 0.23183868644526026
l2 norm of weights: 5.846316533625424

Iteration Number: 5853
Loss: 34.0914596267793
l2 norm of gradients: 0.23181183584382217
l2 norm of weights: 5.846235410222279

Iteration Number: 5854
Loss: 34.08876465101386
l2 norm of gradients: 0.23178499405008768
l2 norm of weights: 5.846154294156417

Iteration Number: 5855
Loss: 34.08607029952345
l2 norm of gradients: 0.2317581610583512
l2 norm of weights: 5.84607318542638

Iteration Number: 5856
Loss: 34.08337657202923
l2 norm of gradients: 0.23173133686291253
l2 norm of weights: 5.845992084030711

Iteration Number: 5857
Loss: 34.08068346825753
l2 norm of gradients: 0.23170452145807677
l2 norm of weights: 5.845910989967953

```
-----  
Iteration Number: 5858  
Loss: 34.07799098793022  
l2 norm of gradients: 0.23167771483815436  
l2 norm of weights: 5.845829903236651  
-----  
Iteration Number: 5859  
Loss: 34.07529913077142  
l2 norm of gradients: 0.2316509169974609  
l2 norm of weights: 5.845748823835349  
-----  
Iteration Number: 5860  
Loss: 34.07260789651453  
l2 norm of gradients: 0.23162412793031747  
l2 norm of weights: 5.845667751762592  
-----  
Iteration Number: 5861  
Loss: 34.06991728487269  
l2 norm of gradients: 0.23159734763105028  
l2 norm of weights: 5.845586687016929  
-----  
Iteration Number: 5862  
Loss: 34.06722729557771  
l2 norm of gradients: 0.2315705760939908  
l2 norm of weights: 5.845505629596903  
-----  
Iteration Number: 5863  
Loss: 34.064537928356266  
l2 norm of gradients: 0.2315438133134759  
l2 norm of weights: 5.845424579501063  
-----  
Iteration Number: 5864  
Loss: 34.061849182931994  
l2 norm of gradients: 0.23151705928384766  
l2 norm of weights: 5.845343536727956  
-----  
Iteration Number: 5865  
Loss: 34.05916105903139  
l2 norm of gradients: 0.23149031399945333  
l2 norm of weights: 5.845262501276133  
-----  
Iteration Number: 5866  
Loss: 34.05647355638292  
l2 norm of gradients: 0.2314635774546456  
l2 norm of weights: 5.845181473144139  
-----  
Iteration Number: 5867  
Loss: 34.05378667471226
```

```

12 norm of gradients: 0.23143684964378217
12 norm of weights: 5.845100452330528
-----
Iteration Number: 5868
Loss: 34.05110041373743
12 norm of gradients: 0.2314101305612262
12 norm of weights: 5.845019438833848
-----
Iteration Number: 5869
Loss: 34.04841477320125
12 norm of gradients: 0.23138342020134592
12 norm of weights: 5.8449384326526514
-----
Iteration Number: 5870
Loss: 34.04572975281847
12 norm of gradients: 0.23135671855851495
12 norm of weights: 5.8448574337854895
-----
Iteration Number: 5871
Loss: 34.04304535232031
12 norm of gradients: 0.23133002562711202
12 norm of weights: 5.844776442230914
-----
Iteration Number: 5872
Loss: 34.040361571433905
12 norm of gradients: 0.23130334140152103
12 norm of weights: 5.844695457987478
-----
Iteration Number: 5873
Loss: 34.03767840988693
12 norm of gradients: 0.23127666587613135
12 norm of weights: 5.844614481053735
-----
Iteration Number: 5874
Loss: 34.034995867407346
12 norm of gradients: 0.23124999904533733
12 norm of weights: 5.8445335114282395
-----
Iteration Number: 5875
Loss: 34.03231394372539
12 norm of gradients: 0.23122334090353852
12 norm of weights: 5.844452549109546
-----
Iteration Number: 5876
Loss: 34.02963263856462
12 norm of gradients: 0.23119669144513985
12 norm of weights: 5.844371594096209
-----

```

Iteration Number: 5877
Loss: 34.026951951652656
l2 norm of gradients: 0.23117005066455126
l2 norm of weights: 5.844290646386788

Iteration Number: 5878
Loss: 34.02427188272689
l2 norm of gradients: 0.23114341855618803
l2 norm of weights: 5.844209705979836

Iteration Number: 5879
Loss: 34.021592431508715
l2 norm of gradients: 0.23111679511447056
l2 norm of weights: 5.84412877287391

Iteration Number: 5880
Loss: 34.0189135977252
l2 norm of gradients: 0.23109018033382434
l2 norm of weights: 5.8440478470675705

Iteration Number: 5881
Loss: 34.01623538111243
l2 norm of gradients: 0.23106357420868007
l2 norm of weights: 5.843966928559375

Iteration Number: 5882
Loss: 34.01355778139773
l2 norm of gradients: 0.23103697673347384
l2 norm of weights: 5.843886017347883

Iteration Number: 5883
Loss: 34.01088079830859
l2 norm of gradients: 0.23101038790264664
l2 norm of weights: 5.843805113431651

Iteration Number: 5884
Loss: 34.008204431575734
l2 norm of gradients: 0.23098380771064467
l2 norm of weights: 5.843724216809244

Iteration Number: 5885
Loss: 34.00552868092994
l2 norm of gradients: 0.23095723615191935
l2 norm of weights: 5.843643327479221

Iteration Number: 5886
Loss: 34.002853546098265
l2 norm of gradients: 0.23093067322092728

```

12 norm of weights: 5.843562445440142
-----
Iteration Number: 5887
Loss: 34.000179026812546
12 norm of gradients: 0.23090411891213
12 norm of weights: 5.843481570690571
-----
Iteration Number: 5888
Loss: 33.997505122809265
12 norm of gradients: 0.2308775732199945
12 norm of weights: 5.843400703229071
-----
Iteration Number: 5889
Loss: 33.99483183380844
12 norm of gradients: 0.23085103613899258
12 norm of weights: 5.843319843054205
-----
Iteration Number: 5890
Loss: 33.992159159547974
12 norm of gradients: 0.2308245076636014
12 norm of weights: 5.843238990164536
-----
Iteration Number: 5891
Loss: 33.98948709976316
12 norm of gradients: 0.23079798778830318
12 norm of weights: 5.843158144558629
-----
Iteration Number: 5892
Loss: 33.986815654175714
12 norm of gradients: 0.23077147650758514
12 norm of weights: 5.84307730623505
-----
Iteration Number: 5893
Loss: 33.984144822517585
12 norm of gradients: 0.23074497381593984
12 norm of weights: 5.842996475192366
-----
Iteration Number: 5894
Loss: 33.98147460453166
12 norm of gradients: 0.2307184797078647
12 norm of weights: 5.84291565142914
-----
Iteration Number: 5895
Loss: 33.9788049999408
12 norm of gradients: 0.23069199417786235
12 norm of weights: 5.842834834943941
-----
Iteration Number: 5896

```

Loss: 33.976136008476516
l2 norm of gradients: 0.23066551722044057
l2 norm of weights: 5.842754025735338

Iteration Number: 5897
Loss: 33.97346762986948
l2 norm of gradients: 0.23063904883011221
l2 norm of weights: 5.842673223801897

Iteration Number: 5898
Loss: 33.9707998638583
l2 norm of gradients: 0.23061258900139506
l2 norm of weights: 5.842592429142187

Iteration Number: 5899
Loss: 33.968132710179255
l2 norm of gradients: 0.2305861377288122
l2 norm of weights: 5.842511641754779

Iteration Number: 5900
Loss: 33.965466168556226
l2 norm of gradients: 0.23055969500689164
l2 norm of weights: 5.842430861638243

Iteration Number: 5901
Loss: 33.96280023872508
l2 norm of gradients: 0.2305332608301665
l2 norm of weights: 5.8423500887911475

Iteration Number: 5902
Loss: 33.9601349204181
l2 norm of gradients: 0.23050683519317508
l2 norm of weights: 5.842269323212068

Iteration Number: 5903
Loss: 33.957470213366115
l2 norm of gradients: 0.23048041809046046
l2 norm of weights: 5.842188564899572

Iteration Number: 5904
Loss: 33.95480611731038
l2 norm of gradients: 0.23045400951657113
l2 norm of weights: 5.842107813852234

Iteration Number: 5905
Loss: 33.95214263198396
l2 norm of gradients: 0.23042760946606028
l2 norm of weights: 5.842027070068626

```

-----
Iteration Number: 5906
Loss: 33.94947975711411
l2 norm of gradients: 0.23040121793348642
l2 norm of weights: 5.841946333547325
-----
Iteration Number: 5907
Loss: 33.94681749243867
l2 norm of gradients: 0.2303748349134129
l2 norm of weights: 5.841865604286902
-----
Iteration Number: 5908
Loss: 33.94415583769755
l2 norm of gradients: 0.2303484604004082
l2 norm of weights: 5.841784882285933
-----
Iteration Number: 5909
Loss: 33.941494792611934
l2 norm of gradients: 0.2303220943890459
l2 norm of weights: 5.841704167542993
-----
Iteration Number: 5910
Loss: 33.93883435692619
l2 norm of gradients: 0.23029573687390448
l2 norm of weights: 5.841623460056659
-----
Iteration Number: 5911
Loss: 33.936174530375006
l2 norm of gradients: 0.23026938784956744
l2 norm of weights: 5.841542759825508
-----
Iteration Number: 5912
Loss: 33.93351531269217
l2 norm of gradients: 0.23024304731062334
l2 norm of weights: 5.841462066848116
-----
Iteration Number: 5913
Loss: 33.930856703616136
l2 norm of gradients: 0.23021671525166573
l2 norm of weights: 5.841381381123061
-----
Iteration Number: 5914
Loss: 33.928198702878646
l2 norm of gradients: 0.2301903916672933
l2 norm of weights: 5.841300702648924
-----
Iteration Number: 5915
Loss: 33.925541310212616

```

```

12 norm of gradients: 0.23016407655210938
12 norm of weights: 5.841220031424282
-----
Iteration Number: 5916
Loss: 33.9228845253599
12 norm of gradients: 0.23013776990072277
12 norm of weights: 5.841139367447714
-----
Iteration Number: 5917
Loss: 33.920228348054096
12 norm of gradients: 0.23011147170774682
12 norm of weights: 5.841058710717802
-----
Iteration Number: 5918
Loss: 33.91757277803126
12 norm of gradients: 0.23008518196780012
12 norm of weights: 5.840978061233127
-----
Iteration Number: 5919
Loss: 33.91491781502897
12 norm of gradients: 0.2300589006755061
12 norm of weights: 5.840897418992269
-----
Iteration Number: 5920
Loss: 33.91226345877811
12 norm of gradients: 0.23003262782549333
12 norm of weights: 5.840816783993812
-----
Iteration Number: 5921
Loss: 33.909609709028636
12 norm of gradients: 0.2300063634123953
12 norm of weights: 5.840736156236337
-----
Iteration Number: 5922
Loss: 33.90695656550218
12 norm of gradients: 0.22998010743085026
12 norm of weights: 5.840655535718429
-----
Iteration Number: 5923
Loss: 33.90430402794688
12 norm of gradients: 0.2299538598755016
12 norm of weights: 5.8405749224386705
-----
Iteration Number: 5924
Loss: 33.901652096093734
12 norm of gradients: 0.22992762074099765
12 norm of weights: 5.840494316395646
-----

```


Iteration Number: 5925
Loss: 33.899000769686545
l2 norm of gradients: 0.22990139002199167
l2 norm of weights: 5.840413717587943

Iteration Number: 5926
Loss: 33.89635004845666
l2 norm of gradients: 0.2298751677131418
l2 norm of weights: 5.840333126014143

Iteration Number: 5927
Loss: 33.89369993214525
l2 norm of gradients: 0.22984895380911127
l2 norm of weights: 5.840252541672837

Iteration Number: 5928
Loss: 33.8910504204924
l2 norm of gradients: 0.22982274830456806
l2 norm of weights: 5.840171964562608

Iteration Number: 5929
Loss: 33.88840151323358
l2 norm of gradients: 0.2297965511941852
l2 norm of weights: 5.840091394682045

Iteration Number: 5930
Loss: 33.8857532101057
l2 norm of gradients: 0.22977036247264054
l2 norm of weights: 5.840010832029738

Iteration Number: 5931
Loss: 33.883105510851074
l2 norm of gradients: 0.22974418213461695
l2 norm of weights: 5.839930276604272

Iteration Number: 5932
Loss: 33.88045841520704
l2 norm of gradients: 0.22971801017480223
l2 norm of weights: 5.839849728404238

Iteration Number: 5933
Loss: 33.877811922913764
l2 norm of gradients: 0.22969184658788885
l2 norm of weights: 5.839769187428227

Iteration Number: 5934
Loss: 33.875166033708425
l2 norm of gradients: 0.2296656913685746

```

12 norm of weights: 5.839688653674828
-----
Iteration Number: 5935
Loss: 33.87252074733479
12 norm of gradients: 0.22963954451156168
12 norm of weights: 5.839608127142632
-----
Iteration Number: 5936
Loss: 33.86987606352432
12 norm of gradients: 0.22961340601155755
12 norm of weights: 5.839527607830231
-----
Iteration Number: 5937
Loss: 33.86723198202549
12 norm of gradients: 0.2295872758632744
12 norm of weights: 5.839447095736217
-----
Iteration Number: 5938
Loss: 33.86458850257321
12 norm of gradients: 0.2295611540614293
12 norm of weights: 5.839366590859183
-----
Iteration Number: 5939
Loss: 33.86194562490884
12 norm of gradients: 0.22953504060074428
12 norm of weights: 5.839286093197723
-----
Iteration Number: 5940
Loss: 33.85930334877625
12 norm of gradients: 0.22950893547594617
12 norm of weights: 5.83920560275043
-----
Iteration Number: 5941
Loss: 33.85666167391058
12 norm of gradients: 0.22948283868176664
12 norm of weights: 5.839125119515899
-----
Iteration Number: 5942
Loss: 33.854020600057986
12 norm of gradients: 0.22945675021294232
12 norm of weights: 5.839044643492724
-----
Iteration Number: 5943
Loss: 33.851380126949365
12 norm of gradients: 0.22943067006421458
12 norm of weights: 5.838964174679503
-----
Iteration Number: 5944

```

Loss: 33.84874025433405
l2 norm of gradients: 0.22940459823032977
l2 norm of weights: 5.83888371307483

Iteration Number: 5945
Loss: 33.84610098195667
l2 norm of gradients: 0.22937853470603894
l2 norm of weights: 5.838803258677304

Iteration Number: 5946
Loss: 33.84346230955026
l2 norm of gradients: 0.22935247948609808
l2 norm of weights: 5.838722811485522

Iteration Number: 5947
Loss: 33.840824236862744
l2 norm of gradients: 0.22932643256526802
l2 norm of weights: 5.83864237149808

Iteration Number: 5948
Loss: 33.8381867636317
l2 norm of gradients: 0.22930039393831442
l2 norm of weights: 5.838561938713579

Iteration Number: 5949
Loss: 33.83554988960442
l2 norm of gradients: 0.22927436360000772
l2 norm of weights: 5.838481513130618

Iteration Number: 5950
Loss: 33.83291361451438
l2 norm of gradients: 0.22924834154512308
l2 norm of weights: 5.838401094747795

Iteration Number: 5951
Loss: 33.83027793811491
l2 norm of gradients: 0.22922232776844073
l2 norm of weights: 5.838320683563712

Iteration Number: 5952
Loss: 33.82764286014002
l2 norm of gradients: 0.2291963222647456
l2 norm of weights: 5.8382402795769694

Iteration Number: 5953
Loss: 33.8250083803325
l2 norm of gradients: 0.22917032502882728
l2 norm of weights: 5.83815988278617

```

-----
Iteration Number: 5954
Loss: 33.82237449843706
l2 norm of gradients: 0.22914433605548046
l2 norm of weights: 5.838079493189913
-----
Iteration Number: 5955
Loss: 33.81974121420459
l2 norm of gradients: 0.2291183553395043
l2 norm of weights: 5.837999110786805
-----
Iteration Number: 5956
Loss: 33.817108527366265
l2 norm of gradients: 0.22909238287570297
l2 norm of weights: 5.837918735575445
-----
Iteration Number: 5957
Loss: 33.81447643767241
l2 norm of gradients: 0.2290664186588854
l2 norm of weights: 5.837838367554441
-----
Iteration Number: 5958
Loss: 33.81184494486569
l2 norm of gradients: 0.22904046268386521
l2 norm of weights: 5.837758006722394
-----
Iteration Number: 5959
Loss: 33.80921404869178
l2 norm of gradients: 0.2290145149454609
l2 norm of weights: 5.837677653077913
-----
Iteration Number: 5960
Loss: 33.80658374888728
l2 norm of gradients: 0.22898857543849557
l2 norm of weights: 5.8375973066196
-----
Iteration Number: 5961
Loss: 33.80395404520396
l2 norm of gradients: 0.2289626441577974
l2 norm of weights: 5.837516967346062
-----
Iteration Number: 5962
Loss: 33.80132493738221
l2 norm of gradients: 0.228936721098199
l2 norm of weights: 5.837436635255908
-----
Iteration Number: 5963
Loss: 33.79869642516867

```

```

12 norm of gradients: 0.22891080625453789
12 norm of weights: 5.837356310347742
-----
Iteration Number: 5964
Loss: 33.79606850830741
12 norm of gradients: 0.22888489962165642
12 norm of weights: 5.8372759926201745
-----
Iteration Number: 5965
Loss: 33.793441186546985
12 norm of gradients: 0.22885900119440153
12 norm of weights: 5.837195682071812
-----
Iteration Number: 5966
Loss: 33.790814459627214
12 norm of gradients: 0.22883311096762507
12 norm of weights: 5.837115378701265
-----
Iteration Number: 5967
Loss: 33.788188327293206
12 norm of gradients: 0.22880722893618338
12 norm of weights: 5.837035082507142
-----
Iteration Number: 5968
Loss: 33.78556278929281
12 norm of gradients: 0.2287813550949378
12 norm of weights: 5.836954793488055
-----
Iteration Number: 5969
Loss: 33.78293784536945
12 norm of gradients: 0.22875548943875423
12 norm of weights: 5.8368745116426135
-----
Iteration Number: 5970
Loss: 33.78031349527038
12 norm of gradients: 0.22872963196250334
12 norm of weights: 5.836794236969429
-----
Iteration Number: 5971
Loss: 33.777689738746986
12 norm of gradients: 0.22870378266106062
12 norm of weights: 5.836713969467113
-----
Iteration Number: 5972
Loss: 33.7750665755375
12 norm of gradients: 0.22867794152930607
12 norm of weights: 5.836633709134279
-----

```

Iteration Number: 5973
Loss: 33.772444005389595
l2 norm of gradients: 0.2286521085621246
l2 norm of weights: 5.83655345596954

Iteration Number: 5974
Loss: 33.76982202805246
l2 norm of gradients: 0.22862628375440558
l2 norm of weights: 5.836473209971508

Iteration Number: 5975
Loss: 33.767200643270556
l2 norm of gradients: 0.2286004671010434
l2 norm of weights: 5.8363929711388

Iteration Number: 5976
Loss: 33.76457985079207
l2 norm of gradients: 0.22857465859693693
l2 norm of weights: 5.836312739470028

Iteration Number: 5977
Loss: 33.761959650363586
l2 norm of gradients: 0.2285488582369897
l2 norm of weights: 5.83623251496381

Iteration Number: 5978
Loss: 33.759340041733005
l2 norm of gradients: 0.22852306601611003
l2 norm of weights: 5.836152297618759

Iteration Number: 5979
Loss: 33.75672102464837
l2 norm of gradients: 0.22849728192921093
l2 norm of weights: 5.836072087433495

Iteration Number: 5980
Loss: 33.75410259885109
l2 norm of gradients: 0.22847150597121005
l2 norm of weights: 5.835991884406631

Iteration Number: 5981
Loss: 33.751484764099615
l2 norm of gradients: 0.2284457381370296
l2 norm of weights: 5.835911688536788

Iteration Number: 5982
Loss: 33.74886752013273
l2 norm of gradients: 0.22841997842159675

```

12 norm of weights: 5.835831499822583
-----
Iteration Number: 5983
Loss: 33.74625086670266
12 norm of gradients: 0.22839422681984287
12 norm of weights: 5.835751318262634
-----
Iteration Number: 5984
Loss: 33.74363480356122
12 norm of gradients: 0.22836848332670448
12 norm of weights: 5.835671143855562
-----
Iteration Number: 5985
Loss: 33.74101933044856
12 norm of gradients: 0.22834274793712242
12 norm of weights: 5.8355909765999865
-----
Iteration Number: 5986
Loss: 33.73840444711582
12 norm of gradients: 0.2283170206460424
12 norm of weights: 5.835510816494526
-----
Iteration Number: 5987
Loss: 33.73579015331757
12 norm of gradients: 0.2282913014484145
12 norm of weights: 5.835430663537804
-----
Iteration Number: 5988
Loss: 33.733176448796705
12 norm of gradients: 0.22826559033919375
12 norm of weights: 5.835350517728441
-----
Iteration Number: 5989
Loss: 33.73056333330672
12 norm of gradients: 0.2282398873133395
12 norm of weights: 5.835270379065058
-----
Iteration Number: 5990
Loss: 33.72795080659258
12 norm of gradients: 0.22821419236581592
12 norm of weights: 5.83519024754628
-----
Iteration Number: 5991
Loss: 33.72533886840628
12 norm of gradients: 0.22818850549159186
12 norm of weights: 5.83511012317073
-----
Iteration Number: 5992

```

```

Loss: 33.722727518493436
l2 norm of gradients: 0.22816282668564056
l2 norm of weights: 5.83503000593703
-----
Iteration Number: 5993
Loss: 33.72011675661193
l2 norm of gradients: 0.22813715594294007
l2 norm of weights: 5.834949895843806
-----
Iteration Number: 5994
Loss: 33.71750658250961
l2 norm of gradients: 0.228111493258473
l2 norm of weights: 5.834869792889682
-----
Iteration Number: 5995
Loss: 33.71489699593065
l2 norm of gradients: 0.2280858386272265
l2 norm of weights: 5.834789697073284
-----
Iteration Number: 5996
Loss: 33.71228799663006
l2 norm of gradients: 0.22806019204419237
l2 norm of weights: 5.834709608393239
-----
Iteration Number: 5997
Loss: 33.709679584357154
l2 norm of gradients: 0.22803455350436702
l2 norm of weights: 5.834629526848174
-----
Iteration Number: 5998
Loss: 33.707071758866626
l2 norm of gradients: 0.22800892300275138
l2 norm of weights: 5.834549452436713
-----
Iteration Number: 5999
Loss: 33.70446451990341
l2 norm of gradients: 0.22798330053435106
l2 norm of weights: 5.834469385157487
-----
Iteration Number: 6000
Loss: 33.70185786721767
l2 norm of gradients: 0.22795768609417613
l2 norm of weights: 5.8343893250091226
-----
Iteration Number: 6001
Loss: 33.6992518005699
l2 norm of gradients: 0.2279320796772414
l2 norm of weights: 5.83430927199025

```



```

-----
Iteration Number: 6002
Loss: 33.69664631970417
l2 norm of gradients: 0.22790648127856614
l2 norm of weights: 5.834229226099497
-----
Iteration Number: 6003
Loss: 33.694041424368514
l2 norm of gradients: 0.22788089089317415
l2 norm of weights: 5.834149187335495
-----
Iteration Number: 6004
Loss: 33.691437114327414
l2 norm of gradients: 0.22785530851609384
l2 norm of weights: 5.8340691556968745
-----
Iteration Number: 6005
Loss: 33.688833389320266
l2 norm of gradients: 0.2278297341423582
l2 norm of weights: 5.833989131182267
-----
Iteration Number: 6006
Loss: 33.686230249107666
l2 norm of gradients: 0.2278041677670047
l2 norm of weights: 5.833909113790301
-----
Iteration Number: 6007
Loss: 33.68362769343704
l2 norm of gradients: 0.22777860938507552
l2 norm of weights: 5.833829103519613
-----
Iteration Number: 6008
Loss: 33.68102572206345
l2 norm of gradients: 0.22775305899161724
l2 norm of weights: 5.833749100368834
-----
Iteration Number: 6009
Loss: 33.67842433473771
l2 norm of gradients: 0.2277275165816809
l2 norm of weights: 5.833669104336598
-----
Iteration Number: 6010
Loss: 33.675823531211364
l2 norm of gradients: 0.22770198215032225
l2 norm of weights: 5.833589115421537
-----
Iteration Number: 6011
Loss: 33.67322331124587

```

```

12 norm of gradients: 0.22767645569260156
12 norm of weights: 5.833509133622287
-----
Iteration Number: 6012
Loss: 33.670623674577776
12 norm of gradients: 0.22765093720358337
12 norm of weights: 5.833429158937483
-----
Iteration Number: 6013
Loss: 33.66802462097567
12 norm of gradients: 0.22762542667833716
12 norm of weights: 5.833349191365761
-----
Iteration Number: 6014
Loss: 33.66542615018784
12 norm of gradients: 0.22759992411193644
12 norm of weights: 5.833269230905755
-----
Iteration Number: 6015
Loss: 33.66282826196904
12 norm of gradients: 0.22757442949945975
12 norm of weights: 5.833189277556105
-----
Iteration Number: 6016
Loss: 33.66023095607039
12 norm of gradients: 0.2275489428359897
12 norm of weights: 5.833109331315446
-----
Iteration Number: 6017
Loss: 33.65763423225169
12 norm of gradients: 0.22752346411661364
12 norm of weights: 5.833029392182415
-----
Iteration Number: 6018
Loss: 33.655038090259225
12 norm of gradients: 0.2274979933364232
12 norm of weights: 5.832949460155653
-----
Iteration Number: 6019
Loss: 33.652442529851996
12 norm of gradients: 0.2274725304905148
12 norm of weights: 5.832869535233797
-----
Iteration Number: 6020
Loss: 33.649847550780244
12 norm of gradients: 0.22744707557398916
12 norm of weights: 5.832789617415487
-----

```

Iteration Number: 6021
Loss: 33.64725315280858
l2 norm of gradients: 0.22742162858195142
l2 norm of weights: 5.8327097066993625

Iteration Number: 6022
Loss: 33.644659335681176
l2 norm of gradients: 0.2273961895095114
l2 norm of weights: 5.832629803084064

Iteration Number: 6023
Loss: 33.64206609915424
l2 norm of gradients: 0.22737075835178325
l2 norm of weights: 5.832549906568235

Iteration Number: 6024
Loss: 33.63947344299257
l2 norm of gradients: 0.22734533510388558
l2 norm of weights: 5.832470017150512

Iteration Number: 6025
Loss: 33.63688136694192
l2 norm of gradients: 0.2273199197609415
l2 norm of weights: 5.832390134829542

Iteration Number: 6026
Loss: 33.63428987076044
l2 norm of gradients: 0.22729451231807862
l2 norm of weights: 5.832310259603966

Iteration Number: 6027
Loss: 33.63169895420259
l2 norm of gradients: 0.22726911277042888
l2 norm of weights: 5.832230391472428

Iteration Number: 6028
Loss: 33.62910861702953
l2 norm of gradients: 0.22724372111312888
l2 norm of weights: 5.83215053043357

Iteration Number: 6029
Loss: 33.62651885899562
l2 norm of gradients: 0.22721833734131944
l2 norm of weights: 5.832070676486039

Iteration Number: 6030
Loss: 33.62392967984818
l2 norm of gradients: 0.2271929614501459

```

12 norm of weights: 5.831990829628476
-----
Iteration Number: 6031
Loss: 33.6213410793551
12 norm of gradients: 0.22716759343475815
12 norm of weights: 5.831910989859532
-----
Iteration Number: 6032
Loss: 33.61875305726445
12 norm of gradients: 0.22714223329031033
12 norm of weights: 5.831831157177849
-----
Iteration Number: 6033
Loss: 33.61616561334437
12 norm of gradients: 0.22711688101196106
12 norm of weights: 5.8317513315820735
-----
Iteration Number: 6034
Loss: 33.6135787473395
12 norm of gradients: 0.22709153659487355
12 norm of weights: 5.831671513070854
-----
Iteration Number: 6035
Loss: 33.61099245901006
12 norm of gradients: 0.22706620003421518
12 norm of weights: 5.8315917016428385
-----
Iteration Number: 6036
Loss: 33.608406748116735
12 norm of gradients: 0.22704087132515785
12 norm of weights: 5.831511897296673
-----
Iteration Number: 6037
Loss: 33.60582161441396
12 norm of gradients: 0.22701555046287786
12 norm of weights: 5.83143210003101
-----
Iteration Number: 6038
Loss: 33.60323705766406
12 norm of gradients: 0.22699023744255603
12 norm of weights: 5.831352309844494
-----
Iteration Number: 6039
Loss: 33.60065307761456
12 norm of gradients: 0.22696493225937733
12 norm of weights: 5.831272526735778
-----
Iteration Number: 6040

```

```

Loss: 33.598069674036786
12 norm of gradients: 0.2269396349085313
12 norm of weights: 5.831192750703512
-----
Iteration Number: 6041
Loss: 33.595486846675485
12 norm of gradients: 0.2269143453852119
12 norm of weights: 5.831112981746347
-----
Iteration Number: 6042
Loss: 33.59290459530133
12 norm of gradients: 0.22688906368461736
12 norm of weights: 5.831033219862934
-----
Iteration Number: 6043
Loss: 33.59032291966269
12 norm of gradients: 0.2268637898019504
12 norm of weights: 5.830953465051924
-----
Iteration Number: 6044
Loss: 33.5877418195229
12 norm of gradients: 0.22683852373241795
12 norm of weights: 5.83087371731197
-----
Iteration Number: 6045
Loss: 33.585161294638766
12 norm of gradients: 0.22681326547123146
12 norm of weights: 5.830793976641728
-----
Iteration Number: 6046
Loss: 33.58258134477377
12 norm of gradients: 0.22678801501360682
12 norm of weights: 5.830714243039848
-----
Iteration Number: 6047
Loss: 33.580001969682435
12 norm of gradients: 0.22676277235476397
12 norm of weights: 5.830634516504984
-----
Iteration Number: 6048
Loss: 33.57742316912165
12 norm of gradients: 0.2267375374899275
12 norm of weights: 5.830554797035794
-----
Iteration Number: 6049
Loss: 33.57484494285412
12 norm of gradients: 0.22671231041432632
12 norm of weights: 5.830475084630931

```

```

-----
Iteration Number: 6050
Loss: 33.57226729064226
l2 norm of gradients: 0.2266870911231935
l2 norm of weights: 5.83039537928905
-----
Iteration Number: 6051
Loss: 33.569690212243955
l2 norm of gradients: 0.22666187961176665
l2 norm of weights: 5.830315681008808
-----
Iteration Number: 6052
Loss: 33.56711370741658
l2 norm of gradients: 0.2266366758752877
l2 norm of weights: 5.830235989788863
-----
Iteration Number: 6053
Loss: 33.564537775920364
l2 norm of gradients: 0.22661147990900288
l2 norm of weights: 5.830156305627871
-----
Iteration Number: 6054
Loss: 33.56196241751698
l2 norm of gradients: 0.2265862917081626
l2 norm of weights: 5.83007662852449
-----
Iteration Number: 6055
Loss: 33.55938763196513
l2 norm of gradients: 0.2265611112680219
l2 norm of weights: 5.8299969584773805
-----
Iteration Number: 6056
Loss: 33.55681341903017
l2 norm of gradients: 0.2265359385838398
l2 norm of weights: 5.829917295485199
-----
Iteration Number: 6057
Loss: 33.554239778469615
l2 norm of gradients: 0.22651077365088002
l2 norm of weights: 5.829837639546605
-----
Iteration Number: 6058
Loss: 33.551666710042
l2 norm of gradients: 0.22648561646441023
l2 norm of weights: 5.82975799066026
-----
Iteration Number: 6059
Loss: 33.54909421350999

```

```

12 norm of gradients: 0.22646046701970263
12 norm of weights: 5.829678348824825
-----
Iteration Number: 6060
Loss: 33.546522288633994
12 norm of gradients: 0.22643532531203373
12 norm of weights: 5.8295987140389585
-----
Iteration Number: 6061
Loss: 33.54395093518208
12 norm of gradients: 0.2264101913366841
12 norm of weights: 5.829519086301325
-----
Iteration Number: 6062
Loss: 33.54138015290965
12 norm of gradients: 0.22638506508893894
12 norm of weights: 5.829439465610585
-----
Iteration Number: 6063
Loss: 33.53880994157432
12 norm of gradients: 0.2263599465640875
12 norm of weights: 5.829359851965401
-----
Iteration Number: 6064
Loss: 33.53624030094525
12 norm of gradients: 0.22633483575742344
12 norm of weights: 5.829280245364437
-----
Iteration Number: 6065
Loss: 33.533671230781046
12 norm of gradients: 0.22630973266424462
12 norm of weights: 5.829200645806357
-----
Iteration Number: 6066
Loss: 33.531102730848545
12 norm of gradients: 0.22628463727985323
12 norm of weights: 5.829121053289825
-----
Iteration Number: 6067
Loss: 33.528534800902285
12 norm of gradients: 0.22625954959955566
12 norm of weights: 5.829041467813505
-----
Iteration Number: 6068
Loss: 33.52596744071053
12 norm of gradients: 0.2262344696186627
12 norm of weights: 5.828961889376064
-----

```

Iteration Number: 6069
Loss: 33.52340065003285
l2 norm of gradients: 0.22620939733248932
l2 norm of weights: 5.828882317976167

Iteration Number: 6070
Loss: 33.52083442863251
l2 norm of gradients: 0.22618433273635477
l2 norm of weights: 5.828802753612481

Iteration Number: 6071
Loss: 33.51826877627536
l2 norm of gradients: 0.22615927582558254
l2 norm of weights: 5.828723196283673

Iteration Number: 6072
Loss: 33.51570369271888
l2 norm of gradients: 0.2261342265955003
l2 norm of weights: 5.828643645988408

Iteration Number: 6073
Loss: 33.51313917773833
l2 norm of gradients: 0.22610918504144017
l2 norm of weights: 5.828564102725357

Iteration Number: 6074
Loss: 33.51057523108013
l2 norm of gradients: 0.22608415115873828
l2 norm of weights: 5.828484566493187

Iteration Number: 6075
Loss: 33.50801185252296
l2 norm of gradients: 0.22605912494273517
l2 norm of weights: 5.828405037290568

Iteration Number: 6076
Loss: 33.50544904182124
l2 norm of gradients: 0.22603410638877544
l2 norm of weights: 5.828325515116169

Iteration Number: 6077
Loss: 33.50288679874007
l2 norm of gradients: 0.22600909549220818
l2 norm of weights: 5.82824599996866

Iteration Number: 6078
Loss: 33.500325123045144
l2 norm of gradients: 0.22598409224838645


```

12 norm of weights: 5.828166491846712
-----
Iteration Number: 6079
Loss: 33.49776401450645
12 norm of gradients: 0.2259590966526677
12 norm of weights: 5.8280869907489965
-----
Iteration Number: 6080
Loss: 33.49520347287903
12 norm of gradients: 0.2259341087004134
12 norm of weights: 5.828007496674186
-----
Iteration Number: 6081
Loss: 33.49264349793316
12 norm of gradients: 0.22590912838698946
12 norm of weights: 5.82792800962095
-----
Iteration Number: 6082
Loss: 33.49008408943025
12 norm of gradients: 0.2258841557077659
12 norm of weights: 5.827848529587963
-----
Iteration Number: 6083
Loss: 33.4875252471359
12 norm of gradients: 0.22585919065811688
12 norm of weights: 5.827769056573898
-----
Iteration Number: 6084
Loss: 33.484966970818846
12 norm of gradients: 0.2258342332334209
12 norm of weights: 5.82768959057743
-----
Iteration Number: 6085
Loss: 33.48240926024018
12 norm of gradients: 0.22580928342906045
12 norm of weights: 5.827610131597231
-----
Iteration Number: 6086
Loss: 33.479852115164555
12 norm of gradients: 0.2257843412404224
12 norm of weights: 5.827530679631978
-----
Iteration Number: 6087
Loss: 33.47729553536109
12 norm of gradients: 0.22575940666289773
12 norm of weights: 5.827451234680344
-----
Iteration Number: 6088

```

```

Loss: 33.474739520598106
12 norm of gradients: 0.2257344796918816
12 norm of weights: 5.827371796741008
-----
Iteration Number: 6089
Loss: 33.4721840706319
12 norm of gradients: 0.22570956032277337
12 norm of weights: 5.827292365812643
-----
Iteration Number: 6090
Loss: 33.46962918523829
12 norm of gradients: 0.2256846485509765
12 norm of weights: 5.827212941893929
-----
Iteration Number: 6091
Loss: 33.4670748641763
12 norm of gradients: 0.2256597443718987
12 norm of weights: 5.827133524983543
-----
Iteration Number: 6092
Loss: 33.464521107217436
12 norm of gradients: 0.22563484778095186
12 norm of weights: 5.827054115080161
-----
Iteration Number: 6093
Loss: 33.46196791412249
12 norm of gradients: 0.22560995877355194
12 norm of weights: 5.826974712182461
-----
Iteration Number: 6094
Loss: 33.45941528466246
12 norm of gradients: 0.2255850773451191
12 norm of weights: 5.826895316289126
-----
Iteration Number: 6095
Loss: 33.45686321860649
12 norm of gradients: 0.22556020349107764
12 norm of weights: 5.826815927398832
-----
Iteration Number: 6096
Loss: 33.45431171571766
12 norm of gradients: 0.22553533720685606
12 norm of weights: 5.82673654551026
-----
Iteration Number: 6097
Loss: 33.45176077575803
12 norm of gradients: 0.22551047848788683
12 norm of weights: 5.8266571706220915

```

```

-----
Iteration Number: 6098
Loss: 33.44921039850937
l2 norm of gradients: 0.22548562732960686
l2 norm of weights: 5.8265778027330075
-----
Iteration Number: 6099
Loss: 33.44666058372488
l2 norm of gradients: 0.2254607837274569
l2 norm of weights: 5.826498441841688
-----
Iteration Number: 6100
Loss: 33.44411133117687
l2 norm of gradients: 0.22543594767688202
l2 norm of weights: 5.8264190879468165
-----
Iteration Number: 6101
Loss: 33.441562640634174
l2 norm of gradients: 0.22541111917333123
l2 norm of weights: 5.826339741047075
-----
Iteration Number: 6102
Loss: 33.43901451186831
l2 norm of gradients: 0.22538629821225792
l2 norm of weights: 5.8262604011411465
-----
Iteration Number: 6103
Loss: 33.43646694463969
l2 norm of gradients: 0.22536148478911935
l2 norm of weights: 5.826181068227716
-----
Iteration Number: 6104
Loss: 33.4339199387233
l2 norm of gradients: 0.22533667889937706
l2 norm of weights: 5.826101742305467
-----
Iteration Number: 6105
Loss: 33.43137349388068
l2 norm of gradients: 0.22531188053849652
l2 norm of weights: 5.826022423373084
-----
Iteration Number: 6106
Loss: 33.42882760988844
l2 norm of gradients: 0.22528708970194747
l2 norm of weights: 5.825943111429251
-----
Iteration Number: 6107
Loss: 33.42628228650601

```

```

12 norm of gradients: 0.22526230638520373
12 norm of weights: 5.825863806472657
-----
Iteration Number: 6108
Loss: 33.42373752351168
12 norm of gradients: 0.22523753058374316
12 norm of weights: 5.825784508501984
-----
Iteration Number: 6109
Loss: 33.42119332066653
12 norm of gradients: 0.2252127622930476
12 norm of weights: 5.825705217515923
-----
Iteration Number: 6110
Loss: 33.41864967774985
12 norm of gradients: 0.2251880015086033
12 norm of weights: 5.825625933513158
-----
Iteration Number: 6111
Loss: 33.4161065945161
12 norm of gradients: 0.22516324822590025
12 norm of weights: 5.825546656492379
-----
Iteration Number: 6112
Loss: 33.41356407074834
12 norm of gradients: 0.2251385024404327
12 norm of weights: 5.825467386452273
-----
Iteration Number: 6113
Loss: 33.41102210620922
12 norm of gradients: 0.22511376414769899
12 norm of weights: 5.825388123391528
-----
Iteration Number: 6114
Loss: 33.40848070067013
12 norm of gradients: 0.22508903334320135
12 norm of weights: 5.825308867308836
-----
Iteration Number: 6115
Loss: 33.40593985389935
12 norm of gradients: 0.2250643100224463
12 norm of weights: 5.8252296182028855
-----
Iteration Number: 6116
Loss: 33.403399565675244
12 norm of gradients: 0.22503959418094432
12 norm of weights: 5.825150376072366
-----

```

Iteration Number: 6117
Loss: 33.40085983575657
l2 norm of gradients: 0.22501488581420986
l2 norm of weights: 5.82507114091597

Iteration Number: 6118
Loss: 33.398320663920536
l2 norm of gradients: 0.22499018491776154
l2 norm of weights: 5.824991912732387

Iteration Number: 6119
Loss: 33.395782049935335
l2 norm of gradients: 0.22496549148712208
l2 norm of weights: 5.8249126915203115

Iteration Number: 6120
Loss: 33.3932439935698
l2 norm of gradients: 0.22494080551781806
l2 norm of weights: 5.824833477278433

Iteration Number: 6121
Loss: 33.39070649460134
l2 norm of gradients: 0.22491612700538022
l2 norm of weights: 5.824754270005446

Iteration Number: 6122
Loss: 33.388169552795596
l2 norm of gradients: 0.2248914559453433
l2 norm of weights: 5.824675069700044

Iteration Number: 6123
Loss: 33.38563316792156
l2 norm of gradients: 0.22486679233324613
l2 norm of weights: 5.82459587636092

Iteration Number: 6124
Loss: 33.383097339758685
l2 norm of gradients: 0.22484213616463153
l2 norm of weights: 5.824516689986769

Iteration Number: 6125
Loss: 33.38056206807161
l2 norm of gradients: 0.22481748743504626
l2 norm of weights: 5.824437510576287

Iteration Number: 6126
Loss: 33.378027352638156
l2 norm of gradients: 0.2247928461400412

```

12 norm of weights: 5.824358338128167
-----
Iteration Number: 6127
Loss: 33.37549319322181
12 norm of gradients: 0.22476821227517124
12 norm of weights: 5.824279172641107
-----
Iteration Number: 6128
Loss: 33.37295958960017
12 norm of gradients: 0.22474358583599524
12 norm of weights: 5.824200014113803
-----
Iteration Number: 6129
Loss: 33.37042654154429
12 norm of gradients: 0.22471896681807613
12 norm of weights: 5.824120862544951
-----
Iteration Number: 6130
Loss: 33.36789404882515
12 norm of gradients: 0.2246943552169807
12 norm of weights: 5.82404171793325
-----
Iteration Number: 6131
Loss: 33.365362111220264
12 norm of gradients: 0.2246697510282799
12 norm of weights: 5.823962580277397
-----
Iteration Number: 6132
Loss: 33.36283072849241
12 norm of gradients: 0.2246451542475486
12 norm of weights: 5.8238834495760905
-----
Iteration Number: 6133
Loss: 33.36029990042482
12 norm of gradients: 0.22462056487036566
12 norm of weights: 5.823804325828029
-----
Iteration Number: 6134
Loss: 33.35776962678386
12 norm of gradients: 0.22459598289231386
12 norm of weights: 5.823725209031914
-----
Iteration Number: 6135
Loss: 33.355239907343226
12 norm of gradients: 0.2245714083089801
12 norm of weights: 5.823646099186444
-----
Iteration Number: 6136

```

Loss: 33.35271074187972
l2 norm of gradients: 0.22454684111595521
l2 norm of weights: 5.82356699629032

Iteration Number: 6137
Loss: 33.350182130158124
l2 norm of gradients: 0.22452228130883395
l2 norm of weights: 5.823487900342241

Iteration Number: 6138
Loss: 33.34765407196514
l2 norm of gradients: 0.224497728883215
l2 norm of weights: 5.8234088113409115

Iteration Number: 6139
Loss: 33.345126567064625
l2 norm of gradients: 0.22447318383470113
l2 norm of weights: 5.823329729285033

Iteration Number: 6140
Loss: 33.34259961523133
l2 norm of gradients: 0.22444864615889903
l2 norm of weights: 5.823250654173306

Iteration Number: 6141
Loss: 33.34007321624279
l2 norm of gradients: 0.22442411585141928
l2 norm of weights: 5.823171586004435

Iteration Number: 6142
Loss: 33.33754736987165
l2 norm of gradients: 0.2243995929078765
l2 norm of weights: 5.823092524777123

Iteration Number: 6143
Loss: 33.33502207589018
l2 norm of gradients: 0.22437507732388917
l2 norm of weights: 5.823013470490075

Iteration Number: 6144
Loss: 33.33249733407438
l2 norm of gradients: 0.22435056909507978
l2 norm of weights: 5.822934423141994

Iteration Number: 6145
Loss: 33.32997314419639
l2 norm of gradients: 0.22432606821707476
l2 norm of weights: 5.8228553827315865

```

-----
Iteration Number: 6146
Loss: 33.327449506036785
l2 norm of gradients: 0.22430157468550443
l2 norm of weights: 5.822776349257557
-----
Iteration Number: 6147
Loss: 33.32492641936282
l2 norm of gradients: 0.22427708849600309
l2 norm of weights: 5.8226973227186125
-----
Iteration Number: 6148
Loss: 33.32240388395546
l2 norm of gradients: 0.22425260964420893
l2 norm of weights: 5.822618303113458
-----
Iteration Number: 6149
Loss: 33.31988189958956
l2 norm of gradients: 0.22422813812576406
l2 norm of weights: 5.822539290440802
-----
Iteration Number: 6150
Loss: 33.31736046603552
l2 norm of gradients: 0.22420367393631452
l2 norm of weights: 5.822460284699352
-----
Iteration Number: 6151
Loss: 33.31483958307428
l2 norm of gradients: 0.22417921707151028
l2 norm of weights: 5.822381285887814
-----
Iteration Number: 6152
Loss: 33.31231925047548
l2 norm of gradients: 0.22415476752700528
l2 norm of weights: 5.822302294004899
-----
Iteration Number: 6153
Loss: 33.30979946802225
l2 norm of gradients: 0.22413032529845717
l2 norm of weights: 5.822223309049314
-----
Iteration Number: 6154
Loss: 33.30728023548405
l2 norm of gradients: 0.22410589038152767
l2 norm of weights: 5.8221443310197705
-----
Iteration Number: 6155
Loss: 33.304761552639

```



```

12 norm of gradients: 0.22408146277188243
12 norm of weights: 5.822065359914976
-----
Iteration Number: 6156
Loss: 33.302243419264165
12 norm of gradients: 0.2240570424651908
12 norm of weights: 5.821986395733644
-----
Iteration Number: 6157
Loss: 33.29972583513178
12 norm of gradients: 0.22403262945712624
12 norm of weights: 5.821907438474483
-----
Iteration Number: 6158
Loss: 33.29720880002673
12 norm of gradients: 0.22400822374336601
12 norm of weights: 5.821828488136206
-----
Iteration Number: 6159
Loss: 33.29469231372228
12 norm of gradients: 0.22398382531959116
12 norm of weights: 5.821749544717524
-----
Iteration Number: 6160
Loss: 33.29217637599016
12 norm of gradients: 0.2239594341814867
12 norm of weights: 5.82167060821715
-----
Iteration Number: 6161
Loss: 33.28966098661296
12 norm of gradients: 0.22393505032474154
12 norm of weights: 5.821591678633798
-----
Iteration Number: 6162
Loss: 33.28714614536484
12 norm of gradients: 0.22391067374504842
12 norm of weights: 5.821512755966179
-----
Iteration Number: 6163
Loss: 33.28463185202263
12 norm of gradients: 0.22388630443810398
12 norm of weights: 5.821433840213008
-----
Iteration Number: 6164
Loss: 33.2821181063645
12 norm of gradients: 0.2238619423996087
12 norm of weights: 5.821354931373
-----

```

Iteration Number: 6165
Loss: 33.27960490817146
l2 norm of gradients: 0.22383758762526684
l2 norm of weights: 5.821276029444871

Iteration Number: 6166
Loss: 33.27709225721861
l2 norm of gradients: 0.22381324011078668
l2 norm of weights: 5.821197134427334

Iteration Number: 6167
Loss: 33.274580153276275
l2 norm of gradients: 0.22378889985188022
l2 norm of weights: 5.821118246319107

Iteration Number: 6168
Loss: 33.272068596136606
l2 norm of gradients: 0.22376456684426332
l2 norm of weights: 5.821039365118904

Iteration Number: 6169
Loss: 33.269557585568336
l2 norm of gradients: 0.2237402410836557
l2 norm of weights: 5.820960490825444

Iteration Number: 6170
Loss: 33.267047121354466
l2 norm of gradients: 0.22371592256578102
l2 norm of weights: 5.820881623437445

Iteration Number: 6171
Loss: 33.26453720326547
l2 norm of gradients: 0.22369161128636655
l2 norm of weights: 5.820802762953622

Iteration Number: 6172
Loss: 33.26202783108881
l2 norm of gradients: 0.2236673072411436
l2 norm of weights: 5.820723909372696

Iteration Number: 6173
Loss: 33.25951900459875
l2 norm of gradients: 0.22364301042584714
l2 norm of weights: 5.8206450626933846

Iteration Number: 6174
Loss: 33.257010723574375
l2 norm of gradients: 0.22361872083621617

```

12 norm of weights: 5.820566222914408
-----
Iteration Number: 6175
Loss: 33.25450298779788
12 norm of gradients: 0.22359443846799323
12 norm of weights: 5.820487390034486
-----
Iteration Number: 6176
Loss: 33.25199579703997
12 norm of gradients: 0.22357016331692486
12 norm of weights: 5.820408564052337
-----
Iteration Number: 6177
Loss: 33.24948915109366
12 norm of gradients: 0.22354589537876146
12 norm of weights: 5.820329744966685
-----
Iteration Number: 6178
Loss: 33.24698304972395
12 norm of gradients: 0.22352163464925706
12 norm of weights: 5.820250932776249
-----
Iteration Number: 6179
Loss: 33.244477492720144
12 norm of gradients: 0.22349738112416961
12 norm of weights: 5.8201721274797515
-----
Iteration Number: 6180
Loss: 33.24197247986126
12 norm of gradients: 0.22347313479926084
12 norm of weights: 5.820093329075915
-----
Iteration Number: 6181
Loss: 33.239468010921186
12 norm of gradients: 0.2234488956702962
12 norm of weights: 5.820014537563462
-----
Iteration Number: 6182
Loss: 33.2369640856815
12 norm of gradients: 0.22342466373304498
12 norm of weights: 5.819935752941116
-----
Iteration Number: 6183
Loss: 33.23446070392711
12 norm of gradients: 0.22340043898328038
12 norm of weights: 5.8198569752076
-----
Iteration Number: 6184

```

Loss: 33.2319578654337
l2 norm of gradients: 0.22337622141677912
l2 norm of weights: 5.81977820436164

Iteration Number: 6185
Loss: 33.22945556998677
l2 norm of gradients: 0.2233520110293219
l2 norm of weights: 5.819699440401959

Iteration Number: 6186
Loss: 33.22695381736121
l2 norm of gradients: 0.22332780781669317
l2 norm of weights: 5.819620683327283

Iteration Number: 6187
Loss: 33.22445260734392
l2 norm of gradients: 0.22330361177468108
l2 norm of weights: 5.819541933136338

Iteration Number: 6188
Loss: 33.22195193970727
l2 norm of gradients: 0.22327942289907754
l2 norm of weights: 5.819463189827848

Iteration Number: 6189
Loss: 33.219451814235775
l2 norm of gradients: 0.22325524118567827
l2 norm of weights: 5.819384453400544

Iteration Number: 6190
Loss: 33.21695223071852
l2 norm of gradients: 0.22323106663028283
l2 norm of weights: 5.819305723853148

Iteration Number: 6191
Loss: 33.214453188924
l2 norm of gradients: 0.22320689922869436
l2 norm of weights: 5.819227001184391

Iteration Number: 6192
Loss: 33.21195468864125
l2 norm of gradients: 0.22318273897671984
l2 norm of weights: 5.819148285393002

Iteration Number: 6193
Loss: 33.209456729652985
l2 norm of gradients: 0.22315858587016987
l2 norm of weights: 5.819069576477706

```

-----
Iteration Number: 6194
Loss: 33.206959311737144
l2 norm of gradients: 0.22313443990485912
l2 norm of weights: 5.818990874437235
-----
Iteration Number: 6195
Loss: 33.20446243467675
l2 norm of gradients: 0.22311030107660573
l2 norm of weights: 5.818912179270317
-----
Iteration Number: 6196
Loss: 33.20196609825445
l2 norm of gradients: 0.22308616938123152
l2 norm of weights: 5.818833490975683
-----
Iteration Number: 6197
Loss: 33.199470302249246
l2 norm of gradients: 0.22306204481456227
l2 norm of weights: 5.818754809552063
-----
Iteration Number: 6198
Loss: 33.19697504645048
l2 norm of gradients: 0.2230379273724272
l2 norm of weights: 5.818676134998188
-----
Iteration Number: 6199
Loss: 33.194480330634704
l2 norm of gradients: 0.22301381705065965
l2 norm of weights: 5.81859746731279
-----
Iteration Number: 6200
Loss: 33.19198615458707
l2 norm of gradients: 0.22298971384509625
l2 norm of weights: 5.818518806494601
-----
Iteration Number: 6201
Loss: 33.18949251808879
l2 norm of gradients: 0.22296561775157756
l2 norm of weights: 5.818440152542352
-----
Iteration Number: 6202
Loss: 33.18699942092256
l2 norm of gradients: 0.22294152876594792
l2 norm of weights: 5.818361505454777
-----
Iteration Number: 6203
Loss: 33.18450686287207

```

```

12 norm of gradients: 0.22291744688405524
12 norm of weights: 5.818282865230611
-----
Iteration Number: 6204
Loss: 33.182014843720324
12 norm of gradients: 0.22289337210175117
12 norm of weights: 5.8182042318685845
-----
Iteration Number: 6205
Loss: 33.17952336325454
12 norm of gradients: 0.22286930441489103
12 norm of weights: 5.818125605367435
-----
Iteration Number: 6206
Loss: 33.17703242124712
12 norm of gradients: 0.22284524381933396
12 norm of weights: 5.818046985725895
-----
Iteration Number: 6207
Loss: 33.17454201749767
12 norm of gradients: 0.2228211903109426
12 norm of weights: 5.817968372942702
-----
Iteration Number: 6208
Loss: 33.17205215177835
12 norm of gradients: 0.22279714388558342
12 norm of weights: 5.81788976701659
-----
Iteration Number: 6209
Loss: 33.169562823873434
12 norm of gradients: 0.22277310453912652
12 norm of weights: 5.817811167946296
-----
Iteration Number: 6210
Loss: 33.16707403357136
12 norm of gradients: 0.22274907226744575
12 norm of weights: 5.817732575730557
-----
Iteration Number: 6211
Loss: 33.16458578065697
12 norm of gradients: 0.22272504706641852
12 norm of weights: 5.817653990368109
-----
Iteration Number: 6212
Loss: 33.16209806491095
12 norm of gradients: 0.2227010289319259
12 norm of weights: 5.817575411857693
-----

```

Iteration Number: 6213
Loss: 33.159610886116184
l2 norm of gradients: 0.2226770178598528
l2 norm of weights: 5.817496840198043

Iteration Number: 6214
Loss: 33.15712424406187
l2 norm of gradients: 0.22265301384608763
l2 norm of weights: 5.8174182753879

Iteration Number: 6215
Loss: 33.15463813853219
l2 norm of gradients: 0.2226290168865225
l2 norm of weights: 5.817339717426003

Iteration Number: 6216
Loss: 33.15215256930853
l2 norm of gradients: 0.22260502697705328
l2 norm of weights: 5.817261166311091

Iteration Number: 6217
Loss: 33.14966753617788
l2 norm of gradients: 0.22258104411357932
l2 norm of weights: 5.817182622041904

Iteration Number: 6218
Loss: 33.14718303892969
l2 norm of gradients: 0.2225570682920037
l2 norm of weights: 5.817104084617184

Iteration Number: 6219
Loss: 33.144699077342246
l2 norm of gradients: 0.22253309950823313
l2 norm of weights: 5.81702555403567

Iteration Number: 6220
Loss: 33.14221565120234
l2 norm of gradients: 0.22250913775817804
l2 norm of weights: 5.816947030296105

Iteration Number: 6221
Loss: 33.13973276030224
l2 norm of gradients: 0.2224851830377524
l2 norm of weights: 5.81686851339723

Iteration Number: 6222
Loss: 33.13725040441867
l2 norm of gradients: 0.2224612353428738

```

12 norm of weights: 5.816790003337789
-----
Iteration Number: 6223
Loss: 33.1347685833395
12 norm of gradients: 0.22243729466946358
12 norm of weights: 5.816711500116523
-----
Iteration Number: 6224
Loss: 33.13228729685889
12 norm of gradients: 0.22241336101344655
12 norm of weights: 5.816633003732177
-----
Iteration Number: 6225
Loss: 33.129806544753876
12 norm of gradients: 0.22238943437075126
12 norm of weights: 5.816554514183493
-----
Iteration Number: 6226
Loss: 33.1273263268153
12 norm of gradients: 0.22236551473730987
12 norm of weights: 5.816476031469218
-----
Iteration Number: 6227
Loss: 33.12484664282477
12 norm of gradients: 0.222341602109058
12 norm of weights: 5.816397555588094
-----
Iteration Number: 6228
Loss: 33.12236749257351
12 norm of gradients: 0.2223176964819351
12 norm of weights: 5.816319086538868
-----
Iteration Number: 6229
Loss: 33.11988887584304
12 norm of gradients: 0.22229379785188416
12 norm of weights: 5.816240624320286
-----
Iteration Number: 6230
Loss: 33.117410792433105
12 norm of gradients: 0.22226990621485163
12 norm of weights: 5.8161621689310925
-----
Iteration Number: 6231
Loss: 33.114933242115455
12 norm of gradients: 0.2222460215667877
12 norm of weights: 5.816083720370036
-----
Iteration Number: 6232

```



```

Loss: 33.112456224680734
l2 norm of gradients: 0.22222214390364606
l2 norm of weights: 5.816005278635863
-----
Iteration Number: 6233
Loss: 33.10997973992045
l2 norm of gradients: 0.22219827322138416
l2 norm of weights: 5.815926843727321
-----
Iteration Number: 6234
Loss: 33.10750378762311
l2 norm of gradients: 0.22217440951596287
l2 norm of weights: 5.815848415643159
-----
Iteration Number: 6235
Loss: 33.105028367568316
l2 norm of gradients: 0.2221505527833467
l2 norm of weights: 5.815769994382124
-----
Iteration Number: 6236
Loss: 33.10255347955201
l2 norm of gradients: 0.22212670301950369
l2 norm of weights: 5.815691579942966
-----
Iteration Number: 6237
Loss: 33.10007912335757
l2 norm of gradients: 0.22210286022040557
l2 norm of weights: 5.815613172324435
-----
Iteration Number: 6238
Loss: 33.09760529877681
l2 norm of gradients: 0.2220790243820276
l2 norm of weights: 5.81553477152528
-----
Iteration Number: 6239
Loss: 33.09513200559326
l2 norm of gradients: 0.2220551955003484
l2 norm of weights: 5.81545637754425
-----
Iteration Number: 6240
Loss: 33.092659243597936
l2 norm of gradients: 0.2220313735713506
l2 norm of weights: 5.8153779903800995
-----
Iteration Number: 6241
Loss: 33.09018701257789
l2 norm of gradients: 0.22200755859101992
l2 norm of weights: 5.815299610031578

```

```
-----  
Iteration Number: 6242  
Loss: 33.08771531231681  
l2 norm of gradients: 0.22198375055534594  
l2 norm of weights: 5.815221236497437  
-----  
Iteration Number: 6243  
Loss: 33.08524414261497  
l2 norm of gradients: 0.2219599494603216  
l2 norm of weights: 5.815142869776429  
-----  
Iteration Number: 6244  
Loss: 33.08277350325364  
l2 norm of gradients: 0.2219361553019436  
l2 norm of weights: 5.815064509867309  
-----  
Iteration Number: 6245  
Loss: 33.08030339402296  
l2 norm of gradients: 0.22191236807621195  
l2 norm of weights: 5.814986156768826  
-----  
Iteration Number: 6246  
Loss: 33.07783381470903  
l2 norm of gradients: 0.22188858777913043  
l2 norm of weights: 5.814907810479736  
-----  
Iteration Number: 6247  
Loss: 33.07536476510638  
l2 norm of gradients: 0.2218648144067061  
l2 norm of weights: 5.814829470998794  
-----  
Iteration Number: 6248  
Loss: 33.07289624500153  
l2 norm of gradients: 0.22184104795494985  
l2 norm of weights: 5.814751138324754  
-----  
Iteration Number: 6249  
Loss: 33.07042825418039  
l2 norm of gradients: 0.22181728841987589  
l2 norm of weights: 5.8146728124563705  
-----  
Iteration Number: 6250  
Loss: 33.067960792438626  
l2 norm of gradients: 0.22179353579750197  
l2 norm of weights: 5.814594493392399  
-----  
Iteration Number: 6251  
Loss: 33.06549385956477
```

```

12 norm of gradients: 0.22176979008384948
12 norm of weights: 5.814516181131597
-----
Iteration Number: 6252
Loss: 33.06302745534912
12 norm of gradients: 0.22174605127494323
12 norm of weights: 5.814437875672719
-----
Iteration Number: 6253
Loss: 33.060561579574845
12 norm of gradients: 0.2217223193668115
12 norm of weights: 5.814359577014523
-----
Iteration Number: 6254
Loss: 33.05809623203894
12 norm of gradients: 0.22169859435548628
12 norm of weights: 5.814281285155767
-----
Iteration Number: 6255
Loss: 33.05563141253425
12 norm of gradients: 0.2216748762370028
12 norm of weights: 5.814203000095207
-----
Iteration Number: 6256
Loss: 33.05316712084306
12 norm of gradients: 0.2216511650074001
12 norm of weights: 5.814124721831604
-----
Iteration Number: 6257
Loss: 33.050703356759335
12 norm of gradients: 0.22162746066272043
12 norm of weights: 5.814046450363715
-----
Iteration Number: 6258
Loss: 33.04824012007761
12 norm of gradients: 0.22160376319900962
12 norm of weights: 5.813968185690299
-----
Iteration Number: 6259
Loss: 33.04577741058166
12 norm of gradients: 0.22158007261231716
12 norm of weights: 5.813889927810116
-----
Iteration Number: 6260
Loss: 33.04331522807064
12 norm of gradients: 0.2215563888986958
12 norm of weights: 5.813811676721928
-----

```

Iteration Number: 6261
Loss: 33.04085357232492
l2 norm of gradients: 0.22153271205420194
l2 norm of weights: 5.813733432424494

Iteration Number: 6262
Loss: 33.03839244314807
l2 norm of gradients: 0.2215090420748954
l2 norm of weights: 5.813655194916573

Iteration Number: 6263
Loss: 33.03593184032276
l2 norm of gradients: 0.2214853789568395
l2 norm of weights: 5.813576964196931

Iteration Number: 6264
Loss: 33.03347176364064
l2 norm of gradients: 0.22146172269610087
l2 norm of weights: 5.813498740264327

Iteration Number: 6265
Loss: 33.031012212899775
l2 norm of gradients: 0.22143807328874984
l2 norm of weights: 5.813420523117525

Iteration Number: 6266
Loss: 33.02855318788569
l2 norm of gradients: 0.2214144307308602
l2 norm of weights: 5.813342312755286

Iteration Number: 6267
Loss: 33.02609468839518
l2 norm of gradients: 0.22139079501850897
l2 norm of weights: 5.813264109176376

Iteration Number: 6268
Loss: 33.02363671421571
l2 norm of gradients: 0.22136716614777682
l2 norm of weights: 5.813185912379556

Iteration Number: 6269
Loss: 33.02117926514032
l2 norm of gradients: 0.2213435441147479
l2 norm of weights: 5.8131077223635925

Iteration Number: 6270
Loss: 33.01872234096165
l2 norm of gradients: 0.22131992891550967

```

12 norm of weights: 5.813029539127251
-----
Iteration Number: 6271
Loss: 33.01626594147461
12 norm of gradients: 0.2212963205461532
12 norm of weights: 5.812951362669294
-----
Iteration Number: 6272
Loss: 33.013810066469425
12 norm of gradients: 0.2212727190027728
12 norm of weights: 5.812873192988489
-----
Iteration Number: 6273
Loss: 33.01135471573852
12 norm of gradients: 0.22124912428146645
12 norm of weights: 5.8127950300836
-----
Iteration Number: 6274
Loss: 33.008899889076496
12 norm of gradients: 0.2212255363783354
12 norm of weights: 5.812716873953398
-----
Iteration Number: 6275
Loss: 33.00644558627198
12 norm of gradients: 0.22120195528948441
12 norm of weights: 5.812638724596645
-----
Iteration Number: 6276
Loss: 33.003991807125416
12 norm of gradients: 0.22117838101102155
12 norm of weights: 5.812560582012111
-----
Iteration Number: 6277
Loss: 33.00153855142279
12 norm of gradients: 0.2211548135390586
12 norm of weights: 5.812482446198565
-----
Iteration Number: 6278
Loss: 32.99908581896517
12 norm of gradients: 0.2211312528697105
12 norm of weights: 5.8124043171547735
-----
Iteration Number: 6279
Loss: 32.99663360953851
12 norm of gradients: 0.22110769899909563
12 norm of weights: 5.812326194879505
-----
Iteration Number: 6280

```

```

Loss: 32.9941819229402
l2 norm of gradients: 0.2210841519233359
l2 norm of weights: 5.812248079371531
-----
Iteration Number: 6281
Loss: 32.99173075895962
l2 norm of gradients: 0.2210606116385565
l2 norm of weights: 5.81216997062962
-----
Iteration Number: 6282
Loss: 32.989280117401776
l2 norm of gradients: 0.22103707814088625
l2 norm of weights: 5.8120918686525425
-----
Iteration Number: 6283
Loss: 32.98682999804502
l2 norm of gradients: 0.22101355142645707
l2 norm of weights: 5.812013773439068
-----
Iteration Number: 6284
Loss: 32.98438040069559
l2 norm of gradients: 0.2209900314914045
l2 norm of weights: 5.81193568498797
-----
Iteration Number: 6285
Loss: 32.98193132514554
l2 norm of gradients: 0.22096651833186748
l2 norm of weights: 5.811857603298017
-----
Iteration Number: 6286
Loss: 32.97948277118464
l2 norm of gradients: 0.22094301194398816
l2 norm of weights: 5.8117795283679845
-----
Iteration Number: 6287
Loss: 32.97703473861141
l2 norm of gradients: 0.2209195123239123
l2 norm of weights: 5.811701460196643
-----
Iteration Number: 6288
Loss: 32.974587227222514
l2 norm of gradients: 0.22089601946778883
l2 norm of weights: 5.811623398782765
-----
Iteration Number: 6289
Loss: 32.97214023680401
l2 norm of gradients: 0.2208725333717703
l2 norm of weights: 5.811545344125125

```

```

-----
Iteration Number: 6290
Loss: 32.96969376716234
l2 norm of gradients: 0.2208490540320124
l2 norm of weights: 5.811467296222497
-----
Iteration Number: 6291
Loss: 32.96724781808478
l2 norm of gradients: 0.2208255814446744
l2 norm of weights: 5.811389255073652
-----
Iteration Number: 6292
Loss: 32.964802389368636
l2 norm of gradients: 0.22080211560591884
l2 norm of weights: 5.8113112206773705
-----
Iteration Number: 6293
Loss: 32.9623574808113
l2 norm of gradients: 0.22077865651191153
l2 norm of weights: 5.811233193032424
-----
Iteration Number: 6294
Loss: 32.95991309220265
l2 norm of gradients: 0.22075520415882197
l2 norm of weights: 5.811155172137589
-----
Iteration Number: 6295
Loss: 32.95746922334294
l2 norm of gradients: 0.22073175854282256
l2 norm of weights: 5.81107715799164
-----
Iteration Number: 6296
Loss: 32.95502587402716
l2 norm of gradients: 0.2207083196600895
l2 norm of weights: 5.810999150593356
-----
Iteration Number: 6297
Loss: 32.95258304405136
l2 norm of gradients: 0.220684887506802
l2 norm of weights: 5.810921149941513
-----
Iteration Number: 6298
Loss: 32.95014073320878
l2 norm of gradients: 0.2206614620791429
l2 norm of weights: 5.810843156034888
-----
Iteration Number: 6299
Loss: 32.94769894129842

```

```

12 norm of gradients: 0.2206380433732981
12 norm of weights: 5.81076516887226
-----
Iteration Number: 6300
Loss: 32.94525766811773
12 norm of gradients: 0.22061463138545712
12 norm of weights: 5.810687188452406
-----
Iteration Number: 6301
Loss: 32.942816913461655
12 norm of gradients: 0.22059122611181264
12 norm of weights: 5.8106092147741055
-----
Iteration Number: 6302
Loss: 32.9403766771241
12 norm of gradients: 0.2205678275485608
12 norm of weights: 5.8105312478361375
-----
Iteration Number: 6303
Loss: 32.93793695890487
12 norm of gradients: 0.2205444356919009
12 norm of weights: 5.810453287637282
-----
Iteration Number: 6304
Loss: 32.93549775860023
12 norm of gradients: 0.2205210505380357
12 norm of weights: 5.810375334176319
-----
Iteration Number: 6305
Loss: 32.933059076007936
12 norm of gradients: 0.2204976720831713
12 norm of weights: 5.810297387452028
-----
Iteration Number: 6306
Loss: 32.9306209109221
12 norm of gradients: 0.2204743003235171
12 norm of weights: 5.810219447463192
-----
Iteration Number: 6307
Loss: 32.92818326314308
12 norm of gradients: 0.22045093525528572
12 norm of weights: 5.810141514208592
-----
Iteration Number: 6308
Loss: 32.92574613246603
12 norm of gradients: 0.22042757687469317
12 norm of weights: 5.810063587687008
-----

```


Iteration Number: 6309
Loss: 32.92330951868927
l2 norm of gradients: 0.22040422517795882
l2 norm of weights: 5.809985667897224

Iteration Number: 6310
Loss: 32.920873421606224
l2 norm of gradients: 0.22038088016130528
l2 norm of weights: 5.809907754838023

Iteration Number: 6311
Loss: 32.91843784102429
l2 norm of gradients: 0.22035754182095849
l2 norm of weights: 5.809829848508188

Iteration Number: 6312
Loss: 32.91600277673293
l2 norm of gradients: 0.22033421015314766
l2 norm of weights: 5.809751948906501

Iteration Number: 6313
Loss: 32.91356822853245
l2 norm of gradients: 0.22031088515410538
l2 norm of weights: 5.809674056031749

Iteration Number: 6314
Loss: 32.9111341962203
l2 norm of gradients: 0.22028756682006737
l2 norm of weights: 5.809596169882714

Iteration Number: 6315
Loss: 32.90870067959621
l2 norm of gradients: 0.22026425514727282
l2 norm of weights: 5.8095182904581835

Iteration Number: 6316
Loss: 32.906267678458185
l2 norm of gradients: 0.22024095013196412
l2 norm of weights: 5.8094404177569405

Iteration Number: 6317
Loss: 32.90383519260331
l2 norm of gradients: 0.22021765177038696
l2 norm of weights: 5.809362551777772

Iteration Number: 6318
Loss: 32.90140322183178
l2 norm of gradients: 0.2201943600587902

```
l2 norm of weights: 5.809284692519465
-----
Iteration Number: 6319
Loss: 32.89897176594282
l2 norm of gradients: 0.2201710749934262
l2 norm of weights: 5.809206839980805
-----
Iteration Number: 6320
Loss: 32.89654082473448
l2 norm of gradients: 0.22014779657055045
l2 norm of weights: 5.809128994160579
-----
Iteration Number: 6321
Loss: 32.89411039799986
l2 norm of gradients: 0.2201245247864216
l2 norm of weights: 5.809051155057577
-----
Iteration Number: 6322
Loss: 32.891680485549216
l2 norm of gradients: 0.22010125963730187
l2 norm of weights: 5.808973322670584
-----
Iteration Number: 6323
Loss: 32.889251087172646
l2 norm of gradients: 0.22007800111945644
l2 norm of weights: 5.808895496998392
-----
Iteration Number: 6324
Loss: 32.88682220267347
l2 norm of gradients: 0.2200547492291539
l2 norm of weights: 5.808817678039787
-----
Iteration Number: 6325
Loss: 32.884393831854865
l2 norm of gradients: 0.2200315039626661
l2 norm of weights: 5.8087398657935605
-----
Iteration Number: 6326
Loss: 32.88196597450903
l2 norm of gradients: 0.22000826531626805
l2 norm of weights: 5.8086620602584995
-----
Iteration Number: 6327
Loss: 32.87953863043461
l2 norm of gradients: 0.21998503328623809
l2 norm of weights: 5.808584261433398
-----
Iteration Number: 6328
```

Loss: 32.87711179943898
l2 norm of gradients: 0.21996180786885774
l2 norm of weights: 5.808506469317045

Iteration Number: 6329
Loss: 32.87468548132015
l2 norm of gradients: 0.21993858906041186
l2 norm of weights: 5.808428683908231

Iteration Number: 6330
Loss: 32.87225967588027
l2 norm of gradients: 0.21991537685718845
l2 norm of weights: 5.8083509052057485

Iteration Number: 6331
Loss: 32.869834382908024
l2 norm of gradients: 0.21989217125547872
l2 norm of weights: 5.808273133208388

Iteration Number: 6332
Loss: 32.867409602218956
l2 norm of gradients: 0.2198689722515773
l2 norm of weights: 5.808195367914945

Iteration Number: 6333
Loss: 32.86498533360452
l2 norm of gradients: 0.21984577984178183
l2 norm of weights: 5.80811760932421

Iteration Number: 6334
Loss: 32.86256157686472
l2 norm of gradients: 0.21982259402239324
l2 norm of weights: 5.808039857434977

Iteration Number: 6335
Loss: 32.86013833180663
l2 norm of gradients: 0.21979941478971574
l2 norm of weights: 5.807962112246041

Iteration Number: 6336
Loss: 32.8577155982251
l2 norm of gradients: 0.21977624214005664
l2 norm of weights: 5.8078843737561945

Iteration Number: 6337
Loss: 32.8552933759231
l2 norm of gradients: 0.21975307606972666
l2 norm of weights: 5.807806641964233

```

-----
Iteration Number: 6338
Loss: 32.85287166470763
l2 norm of gradients: 0.21972991657503949
l2 norm of weights: 5.807728916868951
-----
Iteration Number: 6339
Loss: 32.85045046436469
l2 norm of gradients: 0.2197067636523122
l2 norm of weights: 5.807651198469145
-----
Iteration Number: 6340
Loss: 32.84802977471459
l2 norm of gradients: 0.21968361729786498
l2 norm of weights: 5.80757348676361
-----
Iteration Number: 6341
Loss: 32.84560959554241
l2 norm of gradients: 0.2196604775080212
l2 norm of weights: 5.807495781751144
-----
Iteration Number: 6342
Loss: 32.84318992666341
l2 norm of gradients: 0.21963734427910753
l2 norm of weights: 5.807418083430542
-----
Iteration Number: 6343
Loss: 32.84077076787206
l2 norm of gradients: 0.21961421760745375
l2 norm of weights: 5.807340391800603
-----
Iteration Number: 6344
Loss: 32.838352118968345
l2 norm of gradients: 0.21959109748939284
l2 norm of weights: 5.807262706860123
-----
Iteration Number: 6345
Loss: 32.835933979759886
l2 norm of gradients: 0.21956798392126098
l2 norm of weights: 5.807185028607901
-----
Iteration Number: 6346
Loss: 32.83351635004091
l2 norm of gradients: 0.21954487689939753
l2 norm of weights: 5.807107357042736
-----
Iteration Number: 6347
Loss: 32.831099229622275

```

```

12 norm of gradients: 0.21952177642014498
12 norm of weights: 5.807029692163427
-----
Iteration Number: 6348
Loss: 32.82868261830279
12 norm of gradients: 0.21949868247984908
12 norm of weights: 5.806952033968773
-----
Iteration Number: 6349
Loss: 32.826266515886715
12 norm of gradients: 0.21947559507485867
12 norm of weights: 5.806874382457574
-----
Iteration Number: 6350
Loss: 32.82385092217376
12 norm of gradients: 0.2194525142015258
12 norm of weights: 5.80679673762863
-----
Iteration Number: 6351
Loss: 32.82143583696661
12 norm of gradients: 0.21942943985620578
12 norm of weights: 5.806719099480742
-----
Iteration Number: 6352
Loss: 32.819021260068894
12 norm of gradients: 0.2194063720352568
12 norm of weights: 5.80664146801271
-----
Iteration Number: 6353
Loss: 32.81660719128619
12 norm of gradients: 0.21938331073504055
12 norm of weights: 5.80656384322334
-----
Iteration Number: 6354
Loss: 32.814193630417456
12 norm of gradients: 0.21936025595192157
12 norm of weights: 5.806486225111429
-----
Iteration Number: 6355
Loss: 32.81178057727115
12 norm of gradients: 0.21933720768226786
12 norm of weights: 5.8064086136757815
-----
Iteration Number: 6356
Loss: 32.80936803164377
12 norm of gradients: 0.2193141659224503
12 norm of weights: 5.806331008915201
-----

```

Iteration Number: 6357
Loss: 32.80695599334556
l2 norm of gradients: 0.21929113066884293
l2 norm of weights: 5.80625341082849

Iteration Number: 6358
Loss: 32.80454446217249
l2 norm of gradients: 0.21926810191782317
l2 norm of weights: 5.806175819414453

Iteration Number: 6359
Loss: 32.80213343793943
l2 norm of gradients: 0.21924507966577142
l2 norm of weights: 5.806098234671893

Iteration Number: 6360
Loss: 32.799722920440004
l2 norm of gradients: 0.2192220639090711
l2 norm of weights: 5.806020656599617

Iteration Number: 6361
Loss: 32.79731290948145
l2 norm of gradients: 0.2191990546441089
l2 norm of weights: 5.805943085196427

Iteration Number: 6362
Loss: 32.794903404867796
l2 norm of gradients: 0.21917605186727473
l2 norm of weights: 5.80586552046113

Iteration Number: 6363
Loss: 32.79249440640993
l2 norm of gradients: 0.21915305557496137
l2 norm of weights: 5.8057879623925315

Iteration Number: 6364
Loss: 32.790085913900306
l2 norm of gradients: 0.21913006576356495
l2 norm of weights: 5.805710410989439

Iteration Number: 6365
Loss: 32.787677927148714
l2 norm of gradients: 0.2191070824294845
l2 norm of weights: 5.805632866250658

Iteration Number: 6366
Loss: 32.78527044596501
l2 norm of gradients: 0.2190841055691225

```

12 norm of weights: 5.805555328174997
-----
Iteration Number: 6367
Loss: 32.78286347014521
12 norm of gradients: 0.21906113517888415
12 norm of weights: 5.805477796761264
-----
Iteration Number: 6368
Loss: 32.780456999500814
12 norm of gradients: 0.21903817125517802
12 norm of weights: 5.805400272008264
-----
Iteration Number: 6369
Loss: 32.77805103382831
12 norm of gradients: 0.21901521379441563
12 norm of weights: 5.805322753914808
-----
Iteration Number: 6370
Loss: 32.775645572946175
12 norm of gradients: 0.21899226279301168
12 norm of weights: 5.805245242479704
-----
Iteration Number: 6371
Loss: 32.77324061664721
12 norm of gradients: 0.21896931824738408
12 norm of weights: 5.805167737701763
-----
Iteration Number: 6372
Loss: 32.77083616474056
12 norm of gradients: 0.21894638015395357
12 norm of weights: 5.805090239579792
-----
Iteration Number: 6373
Loss: 32.768432217036256
12 norm of gradients: 0.21892344850914416
12 norm of weights: 5.805012748112604
-----
Iteration Number: 6374
Loss: 32.76602877333666
12 norm of gradients: 0.2189005233093829
12 norm of weights: 5.804935263299008
-----
Iteration Number: 6375
Loss: 32.7636258334449
12 norm of gradients: 0.21887760455109995
12 norm of weights: 5.804857785137815
-----
Iteration Number: 6376

```

```

Loss: 32.761223397169644
l2 norm of gradients: 0.2188546922307285
l2 norm of weights: 5.8047803136278375
-----
Iteration Number: 6377
Loss: 32.75882146431631
l2 norm of gradients: 0.21883178634470488
l2 norm of weights: 5.804702848767886
-----
Iteration Number: 6378
Loss: 32.7564200346922
l2 norm of gradients: 0.21880888688946842
l2 norm of weights: 5.804625390556774
-----
Iteration Number: 6379
Loss: 32.75401910809996
l2 norm of gradients: 0.2187859938614616
l2 norm of weights: 5.804547938993315
-----
Iteration Number: 6380
Loss: 32.75161868435154
l2 norm of gradients: 0.2187631072571299
l2 norm of weights: 5.80447049407632
-----
Iteration Number: 6381
Loss: 32.74921876324473
l2 norm of gradients: 0.21874022707292187
l2 norm of weights: 5.804393055804603
-----
Iteration Number: 6382
Loss: 32.74681934459194
l2 norm of gradients: 0.21871735330528919
l2 norm of weights: 5.804315624176979
-----
Iteration Number: 6383
Loss: 32.74442042820645
l2 norm of gradients: 0.21869448595068658
l2 norm of weights: 5.8042381991922625
-----
Iteration Number: 6384
Loss: 32.74202201387599
l2 norm of gradients: 0.21867162500557163
l2 norm of weights: 5.804160780849269
-----
Iteration Number: 6385
Loss: 32.73962410142846
l2 norm of gradients: 0.21864877046640527
l2 norm of weights: 5.804083369146811

```



```

-----
Iteration Number: 6386
Loss: 32.73722669065781
l2 norm of gradients: 0.21862592232965125
l2 norm of weights: 5.804005964083707
-----
Iteration Number: 6387
Loss: 32.73482978137517
l2 norm of gradients: 0.2186030805917765
l2 norm of weights: 5.803928565658773
-----
Iteration Number: 6388
Loss: 32.732433373390556
l2 norm of gradients: 0.2185802452492509
l2 norm of weights: 5.803851173870823
-----
Iteration Number: 6389
Loss: 32.73003746650596
l2 norm of gradients: 0.21855741629854747
l2 norm of weights: 5.803773788718677
-----
Iteration Number: 6390
Loss: 32.727642060533206
l2 norm of gradients: 0.2185345937361421
l2 norm of weights: 5.80369641020115
-----
Iteration Number: 6391
Loss: 32.72524715527534
l2 norm of gradients: 0.2185117775585139
l2 norm of weights: 5.803619038317062
-----
Iteration Number: 6392
Loss: 32.72285275054417
l2 norm of gradients: 0.2184889677621449
l2 norm of weights: 5.803541673065229
-----
Iteration Number: 6393
Loss: 32.72045884614638
l2 norm of gradients: 0.21846616434352015
l2 norm of weights: 5.803464314444472
-----
Iteration Number: 6394
Loss: 32.71806544189181
l2 norm of gradients: 0.21844336729912772
l2 norm of weights: 5.803386962453608
-----
Iteration Number: 6395
Loss: 32.715672537584375

```

```
12 norm of gradients: 0.21842057662545886
12 norm of weights: 5.803309617091458
-----
Iteration Number: 6396
Loss: 32.71328013303764
12 norm of gradients: 0.21839779231900747
12 norm of weights: 5.803232278356842
-----
Iteration Number: 6397
Loss: 32.71088822805293
12 norm of gradients: 0.2183750143762708
12 norm of weights: 5.803154946248579
-----
Iteration Number: 6398
Loss: 32.70849682244664
12 norm of gradients: 0.218352242793749
12 norm of weights: 5.803077620765491
-----
Iteration Number: 6399
Loss: 32.706105916020284
12 norm of gradients: 0.21832947756794518
12 norm of weights: 5.803000301906397
-----
Iteration Number: 6400
Loss: 32.70371550858573
12 norm of gradients: 0.21830671869536553
12 norm of weights: 5.8029229896701215
-----
Iteration Number: 6401
Loss: 32.70132559995102
12 norm of gradients: 0.21828396617251916
12 norm of weights: 5.802845684055486
-----
Iteration Number: 6402
Loss: 32.698936189929135
12 norm of gradients: 0.21826121999591822
12 norm of weights: 5.80276838506131
-----
Iteration Number: 6403
Loss: 32.69654727832339
12 norm of gradients: 0.21823848016207775
12 norm of weights: 5.802691092686421
-----
Iteration Number: 6404
Loss: 32.69415886494681
12 norm of gradients: 0.218215746667516
12 norm of weights: 5.802613806929639
-----
```

Iteration Number: 6405
Loss: 32.69177094960949
l2 norm of gradients: 0.21819301950875405
l2 norm of weights: 5.802536527789788

Iteration Number: 6406
Loss: 32.689383532112195
l2 norm of gradients: 0.2181702986823159
l2 norm of weights: 5.802459255265693

Iteration Number: 6407
Loss: 32.68699661227558
l2 norm of gradients: 0.21814758418472863
l2 norm of weights: 5.802381989356178

Iteration Number: 6408
Loss: 32.684610189905364
l2 norm of gradients: 0.21812487601252234
l2 norm of weights: 5.802304730060068

Iteration Number: 6409
Loss: 32.682224264807886
l2 norm of gradients: 0.21810217416222993
l2 norm of weights: 5.802227477376189

Iteration Number: 6410
Loss: 32.679838836798446
l2 norm of gradients: 0.2180794786303874
l2 norm of weights: 5.802150231303366

Iteration Number: 6411
Loss: 32.67745390568146
l2 norm of gradients: 0.21805678941353385
l2 norm of weights: 5.802072991840425

Iteration Number: 6412
Loss: 32.675069471275755
l2 norm of gradients: 0.21803410650821095
l2 norm of weights: 5.801995758986193

Iteration Number: 6413
Loss: 32.67268553338036
l2 norm of gradients: 0.21801142991096364
l2 norm of weights: 5.8019185327394975

Iteration Number: 6414
Loss: 32.670302091810164
l2 norm of gradients: 0.21798875961833983

```

12 norm of weights: 5.801841313099165
-----
Iteration Number: 6415
Loss: 32.6679191463819
12 norm of gradients: 0.21796609562689015
12 norm of weights: 5.801764100064024
-----
Iteration Number: 6416
Loss: 32.66553669689467
12 norm of gradients: 0.2179434379331684
12 norm of weights: 5.801686893632903
-----
Iteration Number: 6417
Loss: 32.66315474316893
12 norm of gradients: 0.2179207865337312
12 norm of weights: 5.80160969380463
-----
Iteration Number: 6418
Loss: 32.66077328501174
12 norm of gradients: 0.21789814142513814
12 norm of weights: 5.801532500578034
-----
Iteration Number: 6419
Loss: 32.65839232223302
12 norm of gradients: 0.2178755026039518
12 norm of weights: 5.801455313951945
-----
Iteration Number: 6420
Loss: 32.65601185464807
12 norm of gradients: 0.21785287006673767
12 norm of weights: 5.801378133925193
-----
Iteration Number: 6421
Loss: 32.65363188206025
12 norm of gradients: 0.21783024381006408
12 norm of weights: 5.801300960496608
-----
Iteration Number: 6422
Loss: 32.65125240428997
12 norm of gradients: 0.21780762383050245
12 norm of weights: 5.801223793665021
-----
Iteration Number: 6423
Loss: 32.64887342113952
12 norm of gradients: 0.21778501012462703
12 norm of weights: 5.801146633429263
-----
Iteration Number: 6424

```

Loss: 32.64649493242765
l2 norm of gradients: 0.21776240268901498
l2 norm of weights: 5.801069479788165

Iteration Number: 6425
Loss: 32.644116937965926
l2 norm of gradients: 0.21773980152024644
l2 norm of weights: 5.80099233274056

Iteration Number: 6426
Loss: 32.641739437561725
l2 norm of gradients: 0.21771720661490443
l2 norm of weights: 5.80091519228528

Iteration Number: 6427
Loss: 32.639362431029795
l2 norm of gradients: 0.21769461796957487
l2 norm of weights: 5.800838058421157

Iteration Number: 6428
Loss: 32.63698591818013
l2 norm of gradients: 0.21767203558084666
l2 norm of weights: 5.800760931147026

Iteration Number: 6429
Loss: 32.63460989883058
l2 norm of gradients: 0.2176494594453115
l2 norm of weights: 5.800683810461718

Iteration Number: 6430
Loss: 32.63223437278054
l2 norm of gradients: 0.2176268895595641
l2 norm of weights: 5.800606696364071

Iteration Number: 6431
Loss: 32.62985933985764
l2 norm of gradients: 0.2176043259202019
l2 norm of weights: 5.800529588852914

Iteration Number: 6432
Loss: 32.627484799864526
l2 norm of gradients: 0.21758176852382557
l2 norm of weights: 5.800452487927087

Iteration Number: 6433
Loss: 32.6251107526163
l2 norm of gradients: 0.21755921736703832
l2 norm of weights: 5.800375393585423

```

-----
Iteration Number: 6434
Loss: 32.62273719792649
l2 norm of gradients: 0.21753667244644648
l2 norm of weights: 5.8002983058267565
-----
Iteration Number: 6435
Loss: 32.620364135605186
l2 norm of gradients: 0.21751413375865902
l2 norm of weights: 5.800221224649926
-----
Iteration Number: 6436
Loss: 32.61799156547269
l2 norm of gradients: 0.21749160130028822
l2 norm of weights: 5.800144150053766
-----
Iteration Number: 6437
Loss: 32.61561948733462
l2 norm of gradients: 0.21746907506794874
l2 norm of weights: 5.800067082037115
-----
Iteration Number: 6438
Loss: 32.613247901004414
l2 norm of gradients: 0.21744655505825855
l2 norm of weights: 5.79999002059881
-----
Iteration Number: 6439
Loss: 32.610876806296936
l2 norm of gradients: 0.21742404126783813
l2 norm of weights: 5.799912965737687
-----
Iteration Number: 6440
Loss: 32.60850620302844
l2 norm of gradients: 0.2174015336933111
l2 norm of weights: 5.799835917452587
-----
Iteration Number: 6441
Loss: 32.60613609100762
l2 norm of gradients: 0.21737903233130393
l2 norm of weights: 5.799758875742345
-----
Iteration Number: 6442
Loss: 32.60376647005267
l2 norm of gradients: 0.21735653717844577
l2 norm of weights: 5.799681840605803
-----
Iteration Number: 6443
Loss: 32.601397339975925

```

```

12 norm of gradients: 0.21733404823136876
12 norm of weights: 5.7996048120418
-----
Iteration Number: 6444
Loss: 32.59902870058152
12 norm of gradients: 0.21731156548670794
12 norm of weights: 5.799527790049174
-----
Iteration Number: 6445
Loss: 32.596660551700246
12 norm of gradients: 0.21728908894110113
12 norm of weights: 5.799450774626766
-----
Iteration Number: 6446
Loss: 32.59429289313718
12 norm of gradients: 0.217266618591189
12 norm of weights: 5.799373765773416
-----
Iteration Number: 6447
Loss: 32.59192572470804
12 norm of gradients: 0.21724415443361517
12 norm of weights: 5.799296763487967
-----
Iteration Number: 6448
Loss: 32.58955904622155
12 norm of gradients: 0.21722169646502598
12 norm of weights: 5.799219767769259
-----
Iteration Number: 6449
Loss: 32.587192857500455
12 norm of gradients: 0.21719924468207072
12 norm of weights: 5.799142778616133
-----
Iteration Number: 6450
Loss: 32.58482715835788
12 norm of gradients: 0.2171767990814014
12 norm of weights: 5.799065796027433
-----
Iteration Number: 6451
Loss: 32.582461948603
12 norm of gradients: 0.21715435965967295
12 norm of weights: 5.798988820002
-----
Iteration Number: 6452
Loss: 32.58009722805882
12 norm of gradients: 0.2171319264135432
12 norm of weights: 5.798911850538677
-----

```

Iteration Number: 6453
Loss: 32.577732996533484
l2 norm of gradients: 0.2171094993396727
l2 norm of weights: 5.798834887636309

Iteration Number: 6454
Loss: 32.575369253836186
l2 norm of gradients: 0.2170870784347248
l2 norm of weights: 5.798757931293738

Iteration Number: 6455
Loss: 32.573005999800486
l2 norm of gradients: 0.2170646636953658
l2 norm of weights: 5.7986809815098095

Iteration Number: 6456
Loss: 32.57064323422492
l2 norm of gradients: 0.21704225511826478
l2 norm of weights: 5.798604038283368

Iteration Number: 6457
Loss: 32.56828095693064
l2 norm of gradients: 0.2170198527000936
l2 norm of weights: 5.7985271016132565

Iteration Number: 6458
Loss: 32.5659191677372
l2 norm of gradients: 0.2169974564375269
l2 norm of weights: 5.798450171498324

Iteration Number: 6459
Loss: 32.56355786645091
l2 norm of gradients: 0.21697506632724234
l2 norm of weights: 5.7983732479374135

Iteration Number: 6460
Loss: 32.56119705289475
l2 norm of gradients: 0.2169526823659201
l2 norm of weights: 5.798296330929372

Iteration Number: 6461
Loss: 32.558836726884834
l2 norm of gradients: 0.2169303045502434
l2 norm of weights: 5.798219420473046

Iteration Number: 6462
Loss: 32.55647688823167
l2 norm of gradients: 0.21690793287689805


```

12 norm of weights: 5.7981425165672835
-----
Iteration Number: 6463
Loss: 32.55411753675537
12 norm of gradients: 0.21688556734257297
12 norm of weights: 5.798065619210931
-----
Iteration Number: 6464
Loss: 32.55175867226916
12 norm of gradients: 0.21686320794395958
12 norm of weights: 5.797988728402835
-----
Iteration Number: 6465
Loss: 32.5494002945926
12 norm of gradients: 0.2168408546777522
12 norm of weights: 5.797911844141847
-----
Iteration Number: 6466
Loss: 32.54704240353787
12 norm of gradients: 0.21681850754064794
12 norm of weights: 5.797834966426813
-----
Iteration Number: 6467
Loss: 32.54468499892608
12 norm of gradients: 0.21679616652934683
12 norm of weights: 5.797758095256583
-----
Iteration Number: 6468
Loss: 32.542328080572645
12 norm of gradients: 0.2167738316405514
12 norm of weights: 5.7976812306300065
-----
Iteration Number: 6469
Loss: 32.53997164829016
12 norm of gradients: 0.2167515028709672
12 norm of weights: 5.797604372545933
-----
Iteration Number: 6470
Loss: 32.53761570189839
12 norm of gradients: 0.21672918021730248
12 norm of weights: 5.797527521003214
-----
Iteration Number: 6471
Loss: 32.5352602412161
12 norm of gradients: 0.2167068636762683
12 norm of weights: 5.797450676000698
-----
Iteration Number: 6472

```

Loss: 32.53290526605651
l2 norm of gradients: 0.21668455324457847
l2 norm of weights: 5.797373837537237

Iteration Number: 6473
Loss: 32.53055077624127
l2 norm of gradients: 0.21666224891894945
l2 norm of weights: 5.797297005611682

Iteration Number: 6474
Loss: 32.52819677158641
l2 norm of gradients: 0.2166399506961007
l2 norm of weights: 5.797220180222886

Iteration Number: 6475
Loss: 32.52584325190133
l2 norm of gradients: 0.21661765857275433
l2 norm of weights: 5.7971433613697005

Iteration Number: 6476
Loss: 32.52349021701416
l2 norm of gradients: 0.21659537254563516
l2 norm of weights: 5.797066549050979

Iteration Number: 6477
Loss: 32.52113766673751
l2 norm of gradients: 0.21657309261147084
l2 norm of weights: 5.7969897432655735

Iteration Number: 6478
Loss: 32.51878560088956
l2 norm of gradients: 0.21655081876699173
l2 norm of weights: 5.796912944012336

Iteration Number: 6479
Loss: 32.516434019287615
l2 norm of gradients: 0.21652855100893095
l2 norm of weights: 5.796836151290124

Iteration Number: 6480
Loss: 32.51408292175561
l2 norm of gradients: 0.21650628933402438
l2 norm of weights: 5.796759365097788

Iteration Number: 6481
Loss: 32.511732308099205
l2 norm of gradients: 0.21648403373901073
l2 norm of weights: 5.796682585434186

```

-----
Iteration Number: 6482
Loss: 32.50938217814273
l2 norm of gradients: 0.21646178422063125
l2 norm of weights: 5.7966058122981705
-----
Iteration Number: 6483
Loss: 32.507032531710735
l2 norm of gradients: 0.2164395407756302
l2 norm of weights: 5.796529045688598
-----
Iteration Number: 6484
Loss: 32.50468336861464
l2 norm of gradients: 0.2164173034007543
l2 norm of weights: 5.796452285604324
-----
Iteration Number: 6485
Loss: 32.50233468867342
l2 norm of gradients: 0.21639507209275313
l2 norm of weights: 5.7963755320442045
-----
Iteration Number: 6486
Loss: 32.49998649170504
l2 norm of gradients: 0.2163728468483791
l2 norm of weights: 5.796298785007097
-----
Iteration Number: 6487
Loss: 32.497638777529964
l2 norm of gradients: 0.21635062766438717
l2 norm of weights: 5.796222044491857
-----
Iteration Number: 6488
Loss: 32.49529154597179
l2 norm of gradients: 0.21632841453753515
l2 norm of weights: 5.796145310497343
-----
Iteration Number: 6489
Loss: 32.492944796838074
l2 norm of gradients: 0.2163062074645835
l2 norm of weights: 5.796068583022413
-----
Iteration Number: 6490
Loss: 32.490598529957325
l2 norm of gradients: 0.21628400644229542
l2 norm of weights: 5.795991862065924
-----
Iteration Number: 6491
Loss: 32.488252745142105

```

```

12 norm of gradients: 0.2162618114674369
12 norm of weights: 5.7959151476267365
-----
Iteration Number: 6492
Loss: 32.485907442213744
12 norm of gradients: 0.21623962253677648
12 norm of weights: 5.795838439703707
-----
Iteration Number: 6493
Loss: 32.48356262099647
12 norm of gradients: 0.21621743964708556
12 norm of weights: 5.795761738295697
-----
Iteration Number: 6494
Loss: 32.4812182813062
12 norm of gradients: 0.21619526279513815
12 norm of weights: 5.795685043401566
-----
Iteration Number: 6495
Loss: 32.47887442295895
12 norm of gradients: 0.21617309197771106
12 norm of weights: 5.795608355020173
-----
Iteration Number: 6496
Loss: 32.47653104577719
12 norm of gradients: 0.21615092719158374
12 norm of weights: 5.795531673150379
-----
Iteration Number: 6497
Loss: 32.47418814958367
12 norm of gradients: 0.2161287684335383
12 norm of weights: 5.795454997791046
-----
Iteration Number: 6498
Loss: 32.47184573419428
12 norm of gradients: 0.2161066157003596
12 norm of weights: 5.795378328941033
-----
Iteration Number: 6499
Loss: 32.46950379943065
12 norm of gradients: 0.21608446898883524
12 norm of weights: 5.795301666599204
-----
Iteration Number: 6500
Loss: 32.46716234511428
12 norm of gradients: 0.21606232829575536
12 norm of weights: 5.79522501076442
-----

```

Iteration Number: 6501
Loss: 32.464821371059905
l2 norm of gradients: 0.216040193617913
l2 norm of weights: 5.795148361435543

Iteration Number: 6502
Loss: 32.46248087709086
l2 norm of gradients: 0.21601806495210366
l2 norm of weights: 5.795071718611438

Iteration Number: 6503
Loss: 32.460140863034155
l2 norm of gradients: 0.2159959422951256
l2 norm of weights: 5.794995082290967

Iteration Number: 6504
Loss: 32.45780132869728
l2 norm of gradients: 0.21597382564377984
l2 norm of weights: 5.794918452472992

Iteration Number: 6505
Loss: 32.45546227391202
l2 norm of gradients: 0.21595171499486998
l2 norm of weights: 5.794841829156381

Iteration Number: 6506
Loss: 32.45312369849567
l2 norm of gradients: 0.21592961034520233
l2 norm of weights: 5.794765212339994

Iteration Number: 6507
Loss: 32.45078560226752
l2 norm of gradients: 0.21590751169158587
l2 norm of weights: 5.7946886020227

Iteration Number: 6508
Loss: 32.448447985049235
l2 norm of gradients: 0.21588541903083222
l2 norm of weights: 5.794611998203362

Iteration Number: 6509
Loss: 32.4461108466622
l2 norm of gradients: 0.21586333235975574
l2 norm of weights: 5.794535400880846

Iteration Number: 6510
Loss: 32.443774186928536
l2 norm of gradients: 0.21584125167517326

```

12 norm of weights: 5.794458810054017
-----
Iteration Number: 6511
Loss: 32.441438005667166
12 norm of gradients: 0.21581917697390454
12 norm of weights: 5.794382225721743
-----
Iteration Number: 6512
Loss: 32.43910230269996
12 norm of gradients: 0.21579710825277176
12 norm of weights: 5.79430564788289
-----
Iteration Number: 6513
Loss: 32.43676707784646
12 norm of gradients: 0.21577504550859986
12 norm of weights: 5.794229076536327
-----
Iteration Number: 6514
Loss: 32.43443233093727
12 norm of gradients: 0.21575298873821638
12 norm of weights: 5.7941525116809185
-----
Iteration Number: 6515
Loss: 32.43209806178531
12 norm of gradients: 0.21573093793845155
12 norm of weights: 5.794075953315534
-----
Iteration Number: 6516
Loss: 32.42976427021346
12 norm of gradients: 0.2157088931061382
12 norm of weights: 5.793999401439042
-----
Iteration Number: 6517
Loss: 32.42743095604738
12 norm of gradients: 0.21568685423811193
12 norm of weights: 5.793922856050313
-----
Iteration Number: 6518
Loss: 32.42509811910558
12 norm of gradients: 0.2156648213312107
12 norm of weights: 5.793846317148212
-----
Iteration Number: 6519
Loss: 32.42276575921119
12 norm of gradients: 0.2156427943822754
12 norm of weights: 5.793769784731612
-----
Iteration Number: 6520

```

Loss: 32.420433876188795
l2 norm of gradients: 0.2156207733881494
l2 norm of weights: 5.793693258799382

Iteration Number: 6521
Loss: 32.41810246985226
l2 norm of gradients: 0.21559875834567868
l2 norm of weights: 5.793616739350392

Iteration Number: 6522
Loss: 32.41577154003978
l2 norm of gradients: 0.21557674925171189
l2 norm of weights: 5.793540226383512

Iteration Number: 6523
Loss: 32.413441086558166
l2 norm of gradients: 0.2155547461031003
l2 norm of weights: 5.793463719897615

Iteration Number: 6524
Loss: 32.4111111092387
l2 norm of gradients: 0.21553274889669782
l2 norm of weights: 5.793387219891571

Iteration Number: 6525
Loss: 32.40878160790123
l2 norm of gradients: 0.21551075762936087
l2 norm of weights: 5.793310726364252

Iteration Number: 6526
Loss: 32.40645258236943
l2 norm of gradients: 0.21548877229794866
l2 norm of weights: 5.79323423931453

Iteration Number: 6527
Loss: 32.40412403246922
l2 norm of gradients: 0.21546679289932283
l2 norm of weights: 5.793157758741279

Iteration Number: 6528
Loss: 32.40179595801607
l2 norm of gradients: 0.2154448194303478
l2 norm of weights: 5.79308128464337

Iteration Number: 6529
Loss: 32.39946835883886
l2 norm of gradients: 0.21542285188789045
l2 norm of weights: 5.793004817019678

```

-----
Iteration Number: 6530
Loss: 32.397141234759964
l2 norm of gradients: 0.21540089026882023
l2 norm of weights: 5.792928355869077
-----
Iteration Number: 6531
Loss: 32.39481458560466
l2 norm of gradients: 0.21537893457000937
l2 norm of weights: 5.79285190119044
-----
Iteration Number: 6532
Loss: 32.39248841119041
l2 norm of gradients: 0.21535698478833262
l2 norm of weights: 5.792775452982641
-----
Iteration Number: 6533
Loss: 32.39016271134852
l2 norm of gradients: 0.21533504092066721
l2 norm of weights: 5.792699011244558
-----
Iteration Number: 6534
Loss: 32.38783748589723
l2 norm of gradients: 0.21531310296389314
l2 norm of weights: 5.792622575975063
-----
Iteration Number: 6535
Loss: 32.38551273466513
l2 norm of gradients: 0.21529117091489286
l2 norm of weights: 5.7925461471730335
-----
Iteration Number: 6536
Loss: 32.38318845747109
l2 norm of gradients: 0.21526924477055137
l2 norm of weights: 5.7924697248373445
-----
Iteration Number: 6537
Loss: 32.38086465414079
l2 norm of gradients: 0.21524732452775652
l2 norm of weights: 5.792393308966873
-----
Iteration Number: 6538
Loss: 32.37854132450177
l2 norm of gradients: 0.21522541018339836
l2 norm of weights: 5.792316899560497
-----
Iteration Number: 6539
Loss: 32.376218468369956

```



```

12 norm of gradients: 0.21520350173436983
12 norm of weights: 5.792240496617092
-----
Iteration Number: 6540
Loss: 32.373896085582494
12 norm of gradients: 0.2151815991775662
12 norm of weights: 5.792164100135536
-----
Iteration Number: 6541
Loss: 32.3715741759494
12 norm of gradients: 0.2151597025098856
12 norm of weights: 5.792087710114708
-----
Iteration Number: 6542
Loss: 32.36925273930807
12 norm of gradients: 0.21513781172822838
12 norm of weights: 5.792011326553486
-----
Iteration Number: 6543
Loss: 32.36693177547411
12 norm of gradients: 0.21511592682949773
12 norm of weights: 5.791934949450748
-----
Iteration Number: 6544
Loss: 32.364611284276265
12 norm of gradients: 0.21509404781059924
12 norm of weights: 5.791858578805374
-----
Iteration Number: 6545
Loss: 32.36229126554124
12 norm of gradients: 0.21507217466844114
12 norm of weights: 5.7917822146162425
-----
Iteration Number: 6546
Loss: 32.35997171908803
12 norm of gradients: 0.21505030739993417
12 norm of weights: 5.791705856882236
-----
Iteration Number: 6547
Loss: 32.357652644748796
12 norm of gradients: 0.2150284460019917
12 norm of weights: 5.791629505602231
-----
Iteration Number: 6548
Loss: 32.35533404234001
12 norm of gradients: 0.21500659047152956
12 norm of weights: 5.79155316077511
-----

```

Iteration Number: 6549
Loss: 32.35301591169702
l2 norm of gradients: 0.21498474080546612
l2 norm of weights: 5.791476822399756

Iteration Number: 6550
Loss: 32.3506982526358
l2 norm of gradients: 0.21496289700072238
l2 norm of weights: 5.791400490475047

Iteration Number: 6551
Loss: 32.34838106499587
l2 norm of gradients: 0.21494105905422187
l2 norm of weights: 5.791324164999867

Iteration Number: 6552
Loss: 32.346064348586864
l2 norm of gradients: 0.21491922696289054
l2 norm of weights: 5.791247845973097

Iteration Number: 6553
Loss: 32.34374810324297
l2 norm of gradients: 0.21489740072365698
l2 norm of weights: 5.791171533393621

Iteration Number: 6554
Loss: 32.34143232878476
l2 norm of gradients: 0.2148755803334524
l2 norm of weights: 5.791095227260321

Iteration Number: 6555
Loss: 32.33911702504452
l2 norm of gradients: 0.21485376578921028
l2 norm of weights: 5.791018927572081

Iteration Number: 6556
Loss: 32.3368021918428
l2 norm of gradients: 0.21483195708786687
l2 norm of weights: 5.790942634327784

Iteration Number: 6557
Loss: 32.3344878290152
l2 norm of gradients: 0.21481015422636082
l2 norm of weights: 5.790866347526314

Iteration Number: 6558
Loss: 32.33217393637395
l2 norm of gradients: 0.21478835720163333

```
12 norm of weights: 5.790790067166557
-----
Iteration Number: 6559
Loss: 32.32986051375436
12 norm of gradients: 0.21476656601062818
12 norm of weights: 5.790713793247396
-----
Iteration Number: 6560
Loss: 32.32754756098284
12 norm of gradients: 0.2147447806502915
12 norm of weights: 5.790637525767717
-----
Iteration Number: 6561
Loss: 32.32523507788176
12 norm of gradients: 0.21472300111757217
12 norm of weights: 5.790561264726405
-----
Iteration Number: 6562
Loss: 32.32292306428192
12 norm of gradients: 0.21470122740942138
12 norm of weights: 5.790485010122348
-----
Iteration Number: 6563
Loss: 32.32061152000599
12 norm of gradients: 0.2146794595227929
12 norm of weights: 5.7904087619544296
-----
Iteration Number: 6564
Loss: 32.31830044488436
12 norm of gradients: 0.214657697454643
12 norm of weights: 5.79033252022154
-----
Iteration Number: 6565
Loss: 32.31598983874368
12 norm of gradients: 0.21463594120193055
12 norm of weights: 5.790256284922562
-----
Iteration Number: 6566
Loss: 32.313679701408766
12 norm of gradients: 0.2146141907616167
12 norm of weights: 5.7901800560563865
-----
Iteration Number: 6567
Loss: 32.31137003270814
12 norm of gradients: 0.21459244613066525
12 norm of weights: 5.7901038336219
-----
Iteration Number: 6568
```

Loss: 32.30906083247095
l2 norm of gradients: 0.21457070730604255
l2 norm of weights: 5.790027617617992

Iteration Number: 6569
Loss: 32.30675210052024
l2 norm of gradients: 0.21454897428471725
l2 norm of weights: 5.7899514080435495

Iteration Number: 6570
Loss: 32.3044438366875
l2 norm of gradients: 0.21452724706366066
l2 norm of weights: 5.789875204897461

Iteration Number: 6571
Loss: 32.302136040798814
l2 norm of gradients: 0.21450552563984646
l2 norm of weights: 5.789799008178618

Iteration Number: 6572
Loss: 32.29982871268206
l2 norm of gradients: 0.2144838100102509
l2 norm of weights: 5.7897228178859095

Iteration Number: 6573
Loss: 32.29752185216083
l2 norm of gradients: 0.2144621001718527
l2 norm of weights: 5.789646634018226

Iteration Number: 6574
Loss: 32.295215459071414
l2 norm of gradients: 0.214440396121633
l2 norm of weights: 5.789570456574457

Iteration Number: 6575
Loss: 32.29290953323482
l2 norm of gradients: 0.2144186978565754
l2 norm of weights: 5.789494285553492

Iteration Number: 6576
Loss: 32.29060407448265
l2 norm of gradients: 0.21439700537366607
l2 norm of weights: 5.7894181209542275

Iteration Number: 6577
Loss: 32.28829908264416
l2 norm of gradients: 0.21437531866989357
l2 norm of weights: 5.789341962775549

```

-----
Iteration Number: 6578
Loss: 32.28599455754142
l2 norm of gradients: 0.21435363774224894
l2 norm of weights: 5.789265811016354
-----
Iteration Number: 6579
Loss: 32.283690499009715
l2 norm of gradients: 0.21433196258772574
l2 norm of weights: 5.78918966567553
-----
Iteration Number: 6580
Loss: 32.281386906873635
l2 norm of gradients: 0.21431029320331985
l2 norm of weights: 5.789113526751972
-----
Iteration Number: 6581
Loss: 32.27908378096426
l2 norm of gradients: 0.2142886295860298
l2 norm of weights: 5.789037394244573
-----
Iteration Number: 6582
Loss: 32.276781121104975
l2 norm of gradients: 0.21426697173285641
l2 norm of weights: 5.788961268152228
-----
Iteration Number: 6583
Loss: 32.27447892713139
l2 norm of gradients: 0.2142453196408031
l2 norm of weights: 5.788885148473828
-----
Iteration Number: 6584
Loss: 32.272177198870004
l2 norm of gradients: 0.21422367330687553
l2 norm of weights: 5.788809035208271
-----
Iteration Number: 6585
Loss: 32.26987593615247
l2 norm of gradients: 0.21420203272808203
l2 norm of weights: 5.788732928354447
-----
Iteration Number: 6586
Loss: 32.267575138800375
l2 norm of gradients: 0.2141803979014332
l2 norm of weights: 5.788656827911255
-----
Iteration Number: 6587
Loss: 32.26527480664734

```

```

12 norm of gradients: 0.21415876882394227
12 norm of weights: 5.788580733877588
-----
Iteration Number: 6588
Loss: 32.262974939523126
12 norm of gradients: 0.21413714549262472
12 norm of weights: 5.788504646252342
-----
Iteration Number: 6589
Loss: 32.26067553725809
12 norm of gradients: 0.21411552790449856
12 norm of weights: 5.788428565034414
-----
Iteration Number: 6590
Loss: 32.25837659967565
12 norm of gradients: 0.2140939160565842
12 norm of weights: 5.7883524902227
-----
Iteration Number: 6591
Loss: 32.25607812661377
12 norm of gradients: 0.21407230994590448
12 norm of weights: 5.7882764218160965
-----
Iteration Number: 6592
Loss: 32.253780117896824
12 norm of gradients: 0.21405070956948474
12 norm of weights: 5.7882003598135014
-----
Iteration Number: 6593
Loss: 32.251482573362324
12 norm of gradients: 0.21402911492435264
12 norm of weights: 5.788124304213813
-----
Iteration Number: 6594
Loss: 32.24918549282541
12 norm of gradients: 0.2140075260075384
12 norm of weights: 5.788048255015928
-----
Iteration Number: 6595
Loss: 32.24688887612924
12 norm of gradients: 0.21398594281607441
12 norm of weights: 5.787972212218745
-----
Iteration Number: 6596
Loss: 32.24459272310006
12 norm of gradients: 0.2139643653469958
12 norm of weights: 5.787896175821163
-----

```

Iteration Number: 6597
Loss: 32.24229703356609
l2 norm of gradients: 0.2139427935973398
l2 norm of weights: 5.787820145822081

Iteration Number: 6598
Loss: 32.24000180735725
l2 norm of gradients: 0.21392122756414636
l2 norm of weights: 5.787744122220397

Iteration Number: 6599
Loss: 32.23770704431079
l2 norm of gradients: 0.2138996672444576
l2 norm of weights: 5.787668105015013

Iteration Number: 6600
Loss: 32.23541274424936
l2 norm of gradients: 0.21387811263531814
l2 norm of weights: 5.787592094204828

Iteration Number: 6601
Loss: 32.23311890700642
l2 norm of gradients: 0.21385656373377498
l2 norm of weights: 5.787516089788743

Iteration Number: 6602
Loss: 32.23082553240984
l2 norm of gradients: 0.21383502053687747
l2 norm of weights: 5.78744009176566

Iteration Number: 6603
Loss: 32.2285326202977
l2 norm of gradients: 0.21381348304167752
l2 norm of weights: 5.787364100134478

Iteration Number: 6604
Loss: 32.22624017049267
l2 norm of gradients: 0.2137919512452293
l2 norm of weights: 5.787288114894101

Iteration Number: 6605
Loss: 32.22394818283073
l2 norm of gradients: 0.21377042514458947
l2 norm of weights: 5.787212136043428

Iteration Number: 6606
Loss: 32.22165665714479
l2 norm of gradients: 0.21374890473681687

```

12 norm of weights: 5.787136163581365
-----
Iteration Number: 6607
Loss: 32.21936559325808
12 norm of gradients: 0.213727390018973
12 norm of weights: 5.787060197506812
-----
Iteration Number: 6608
Loss: 32.21707499101093
12 norm of gradients: 0.21370588098812152
12 norm of weights: 5.786984237818674
-----
Iteration Number: 6609
Loss: 32.21478485022797
12 norm of gradients: 0.2136843776413287
12 norm of weights: 5.786908284515853
-----
Iteration Number: 6610
Loss: 32.21249517074526
12 norm of gradients: 0.21366287997566294
12 norm of weights: 5.786832337597255
-----
Iteration Number: 6611
Loss: 32.21020595238994
12 norm of gradients: 0.21364138798819524
12 norm of weights: 5.786756397061781
-----
Iteration Number: 6612
Loss: 32.20791719499968
12 norm of gradients: 0.2136199016759988
12 norm of weights: 5.786680462908339
-----
Iteration Number: 6613
Loss: 32.20562889839986
12 norm of gradients: 0.21359842103614923
12 norm of weights: 5.786604535135831
-----
Iteration Number: 6614
Loss: 32.20334106242897
12 norm of gradients: 0.2135769460657246
12 norm of weights: 5.786528613743164
-----
Iteration Number: 6615
Loss: 32.201053686911344
12 norm of gradients: 0.21355547676180528
12 norm of weights: 5.786452698729244
-----
Iteration Number: 6616

```



```

Loss: 32.19876677168307
l2 norm of gradients: 0.21353401312147402
l2 norm of weights: 5.786376790092977
-----
Iteration Number: 6617
Loss: 32.196480316581614
l2 norm of gradients: 0.21351255514181589
l2 norm of weights: 5.786300887833267
-----
Iteration Number: 6618
Loss: 32.19419432142978
l2 norm of gradients: 0.21349110281991837
l2 norm of weights: 5.786224991949024
-----
Iteration Number: 6619
Loss: 32.191908786063635
l2 norm of gradients: 0.21346965615287125
l2 norm of weights: 5.786149102439154
-----
Iteration Number: 6620
Loss: 32.18962371031809
l2 norm of gradients: 0.21344821513776674
l2 norm of weights: 5.786073219302565
-----
Iteration Number: 6621
Loss: 32.187339094025845
l2 norm of gradients: 0.2134267797716993
l2 norm of weights: 5.785997342538164
-----
Iteration Number: 6622
Loss: 32.1850549370135
l2 norm of gradients: 0.21340535005176584
l2 norm of weights: 5.785921472144859
-----
Iteration Number: 6623
Loss: 32.18277123912154
l2 norm of gradients: 0.21338392597506556
l2 norm of weights: 5.7858456081215595
-----
Iteration Number: 6624
Loss: 32.180488000176894
l2 norm of gradients: 0.21336250753869998
l2 norm of weights: 5.785769750467175
-----
Iteration Number: 6625
Loss: 32.17820522001566
l2 norm of gradients: 0.213341094739773
l2 norm of weights: 5.785693899180614

```

```

-----
Iteration Number: 6626
Loss: 32.17592289846926
l2 norm of gradients: 0.2133196875753909
l2 norm of weights: 5.785618054260786
-----
Iteration Number: 6627
Loss: 32.173641035375
l2 norm of gradients: 0.2132982860426622
l2 norm of weights: 5.785542215706601
-----
Iteration Number: 6628
Loss: 32.17135963056066
l2 norm of gradients: 0.2132768901386978
l2 norm of weights: 5.785466383516971
-----
Iteration Number: 6629
Loss: 32.16907868386434
l2 norm of gradients: 0.2132554998606109
l2 norm of weights: 5.785390557690805
-----
Iteration Number: 6630
Loss: 32.16679819511309
l2 norm of gradients: 0.2132341152055171
l2 norm of weights: 5.785314738227015
-----
Iteration Number: 6631
Loss: 32.1645181641482
l2 norm of gradients: 0.21321273617053418
l2 norm of weights: 5.785238925124513
-----
Iteration Number: 6632
Loss: 32.16223859079684
l2 norm of gradients: 0.21319136275278244
l2 norm of weights: 5.78516311838221
-----
Iteration Number: 6633
Loss: 32.159959474895324
l2 norm of gradients: 0.21316999494938432
l2 norm of weights: 5.785087317999018
-----
Iteration Number: 6634
Loss: 32.157680816279466
l2 norm of gradients: 0.2131486327574647
l2 norm of weights: 5.785011523973851
-----
Iteration Number: 6635
Loss: 32.15540261477982

```

```

12 norm of gradients: 0.21312727617415073
12 norm of weights: 5.78493573630562
-----
Iteration Number: 6636
Loss: 32.15312487023229
12 norm of gradients: 0.21310592519657182
12 norm of weights: 5.784859954993241
-----
Iteration Number: 6637
Loss: 32.15084758247322
12 norm of gradients: 0.21308457982185977
12 norm of weights: 5.784784180035626
-----
Iteration Number: 6638
Loss: 32.14857075133057
12 norm of gradients: 0.21306324004714855
12 norm of weights: 5.784708411431689
-----
Iteration Number: 6639
Loss: 32.14629437664807
12 norm of gradients: 0.21304190586957467
12 norm of weights: 5.784632649180345
-----
Iteration Number: 6640
Loss: 32.144018458247004
12 norm of gradients: 0.21302057728627674
12 norm of weights: 5.7845568932805085
-----
Iteration Number: 6641
Loss: 32.14174299597745
12 norm of gradients: 0.2129992542943957
12 norm of weights: 5.784481143731095
-----
Iteration Number: 6642
Loss: 32.139467989662535
12 norm of gradients: 0.21297793689107486
12 norm of weights: 5.784405400531019
-----
Iteration Number: 6643
Loss: 32.137193439135885
12 norm of gradients: 0.2129566250734597
12 norm of weights: 5.784329663679197
-----
Iteration Number: 6644
Loss: 32.13491934424222
12 norm of gradients: 0.21293531883869815
12 norm of weights: 5.784253933174546
-----

```

Iteration Number: 6645
Loss: 32.13264570481226
l2 norm of gradients: 0.21291401818394035
l2 norm of weights: 5.784178209015982

Iteration Number: 6646
Loss: 32.130372520677824
l2 norm of gradients: 0.21289272310633867
l2 norm of weights: 5.784102491202422

Iteration Number: 6647
Loss: 32.12809979167607
l2 norm of gradients: 0.21287143360304786
l2 norm of weights: 5.784026779732784

Iteration Number: 6648
Loss: 32.125827517643685
l2 norm of gradients: 0.2128501496712249
l2 norm of weights: 5.783951074605983

Iteration Number: 6649
Loss: 32.123555698410954
l2 norm of gradients: 0.212828871308029
l2 norm of weights: 5.78387537582094

Iteration Number: 6650
Loss: 32.12128433382277
l2 norm of gradients: 0.2128075985106217
l2 norm of weights: 5.783799683376572

Iteration Number: 6651
Loss: 32.119013423702576
l2 norm of gradients: 0.21278633127616686
l2 norm of weights: 5.783723997271799

Iteration Number: 6652
Loss: 32.11674296789329
l2 norm of gradients: 0.2127650696018305
l2 norm of weights: 5.783648317505538

Iteration Number: 6653
Loss: 32.11447296623393
l2 norm of gradients: 0.212743813484781
l2 norm of weights: 5.783572644076711

Iteration Number: 6654
Loss: 32.11220341855099
l2 norm of gradients: 0.212722562922189

```
l2 norm of weights: 5.783496976984234
-----
Iteration Number: 6655
Loss: 32.10993432468782
l2 norm of gradients: 0.2127013179112273
l2 norm of weights: 5.783421316227032
-----
Iteration Number: 6656
Loss: 32.10766568447785
l2 norm of gradients: 0.21268007844907105
l2 norm of weights: 5.783345661804022
-----
Iteration Number: 6657
Loss: 32.105397497755575
l2 norm of gradients: 0.21265884453289768
l2 norm of weights: 5.7832700137141275
-----
Iteration Number: 6658
Loss: 32.10312976435879
l2 norm of gradients: 0.21263761615988674
l2 norm of weights: 5.783194371956267
-----
Iteration Number: 6659
Loss: 32.100862484121464
l2 norm of gradients: 0.21261639332722024
l2 norm of weights: 5.783118736529364
-----
Iteration Number: 6660
Loss: 32.0985956568882
l2 norm of gradients: 0.21259517603208228
l2 norm of weights: 5.78304310743234
-----
Iteration Number: 6661
Loss: 32.09632928248106
l2 norm of gradients: 0.21257396427165914
l2 norm of weights: 5.782967484664116
-----
Iteration Number: 6662
Loss: 32.09406336074903
l2 norm of gradients: 0.21255275804313967
l2 norm of weights: 5.782891868223618
-----
Iteration Number: 6663
Loss: 32.09179789152585
l2 norm of gradients: 0.2125315573437146
l2 norm of weights: 5.782816258109765
-----
Iteration Number: 6664
```

Loss: 32.08953287464493
l2 norm of gradients: 0.21251036217057712
l2 norm of weights: 5.782740654321484

Iteration Number: 6665
Loss: 32.08726830994347
l2 norm of gradients: 0.21248917252092245
l2 norm of weights: 5.782665056857697

Iteration Number: 6666
Loss: 32.085004197263
l2 norm of gradients: 0.21246798839194828
l2 norm of weights: 5.782589465717328

Iteration Number: 6667
Loss: 32.082740536436546
l2 norm of gradients: 0.21244680978085445
l2 norm of weights: 5.782513880899301

Iteration Number: 6668
Loss: 32.08047732730007
l2 norm of gradients: 0.21242563668484296
l2 norm of weights: 5.782438302402541

Iteration Number: 6669
Loss: 32.078214569692925
l2 norm of gradients: 0.2124044691011181
l2 norm of weights: 5.782362730225975

Iteration Number: 6670
Loss: 32.07595226345165
l2 norm of gradients: 0.21238330702688638
l2 norm of weights: 5.782287164368527

Iteration Number: 6671
Loss: 32.073690408418976
l2 norm of gradients: 0.2123621504593565
l2 norm of weights: 5.782211604829123

Iteration Number: 6672
Loss: 32.07142900442249
l2 norm of gradients: 0.21234099939573942
l2 norm of weights: 5.782136051606688

Iteration Number: 6673
Loss: 32.069168051308885
l2 norm of gradients: 0.21231985383324825
l2 norm of weights: 5.782060504700152

```

-----
Iteration Number: 6674
Loss: 32.06690754891072
l2 norm of gradients: 0.21229871376909842
l2 norm of weights: 5.781984964108439
-----
Iteration Number: 6675
Loss: 32.06464749706375
l2 norm of gradients: 0.21227757920050747
l2 norm of weights: 5.781909429830477
-----
Iteration Number: 6676
Loss: 32.06238789561561
l2 norm of gradients: 0.21225645012469518
l2 norm of weights: 5.781833901865194
-----
Iteration Number: 6677
Loss: 32.060128744392735
l2 norm of gradients: 0.2122353265388837
l2 norm of weights: 5.781758380211517
-----
Iteration Number: 6678
Loss: 32.057870043237784
l2 norm of gradients: 0.21221420844029698
l2 norm of weights: 5.781682864868376
-----
Iteration Number: 6679
Loss: 32.055611791989676
l2 norm of gradients: 0.21219309582616164
l2 norm of weights: 5.781607355834699
-----
Iteration Number: 6680
Loss: 32.05335399048596
l2 norm of gradients: 0.21217198869370615
l2 norm of weights: 5.781531853109415
-----
Iteration Number: 6681
Loss: 32.05109663856622
l2 norm of gradients: 0.21215088704016133
l2 norm of weights: 5.781456356691453
-----
Iteration Number: 6682
Loss: 32.04883973606788
l2 norm of gradients: 0.2121297908627603
l2 norm of weights: 5.781380866579743
-----
Iteration Number: 6683
Loss: 32.04658328282901

```

```

12 norm of gradients: 0.21210870015873812
12 norm of weights: 5.7813053827732155
-----
Iteration Number: 6684
Loss: 32.04432727869097
12 norm of gradients: 0.2120876149253322
12 norm of weights: 5.7812299052708
-----
Iteration Number: 6685
Loss: 32.04207172348587
12 norm of gradients: 0.21206653515978205
12 norm of weights: 5.7811544340714285
-----
Iteration Number: 6686
Loss: 32.03981661705909
12 norm of gradients: 0.21204546085932952
12 norm of weights: 5.781078969174032
-----
Iteration Number: 6687
Loss: 32.03756195924606
12 norm of gradients: 0.2120243920212185
12 norm of weights: 5.78100351057754
-----
Iteration Number: 6688
Loss: 32.03530774988509
12 norm of gradients: 0.21200332864269508
12 norm of weights: 5.780928058280888
-----
Iteration Number: 6689
Loss: 32.033053988821614
12 norm of gradients: 0.21198227072100762
12 norm of weights: 5.780852612283006
-----
Iteration Number: 6690
Loss: 32.030800675887214
12 norm of gradients: 0.21196121825340644
12 norm of weights: 5.780777172582826
-----
Iteration Number: 6691
Loss: 32.028547810923
12 norm of gradients: 0.21194017123714431
12 norm of weights: 5.780701739179282
-----
Iteration Number: 6692
Loss: 32.02629539376909
12 norm of gradients: 0.21191912966947601
12 norm of weights: 5.780626312071307
-----

```


Iteration Number: 6693
Loss: 32.02404342426775
l2 norm of gradients: 0.21189809354765848
l2 norm of weights: 5.780550891257835

Iteration Number: 6694
Loss: 32.021791902252964
l2 norm of gradients: 0.21187706286895083
l2 norm of weights: 5.7804754767378

Iteration Number: 6695
Loss: 32.01954082757212
l2 norm of gradients: 0.2118560376306144
l2 norm of weights: 5.7804000685101355

Iteration Number: 6696
Loss: 32.017290200059335
l2 norm of gradients: 0.21183501782991265
l2 norm of weights: 5.780324666573777

Iteration Number: 6697
Loss: 32.01504001955133
l2 norm of gradients: 0.21181400346411117
l2 norm of weights: 5.780249270927659

Iteration Number: 6698
Loss: 32.01279028589271
l2 norm of gradients: 0.21179299453047779
l2 norm of weights: 5.780173881570717

Iteration Number: 6699
Loss: 32.0105409989249
l2 norm of gradients: 0.21177199102628233
l2 norm of weights: 5.7800984985018875

Iteration Number: 6700
Loss: 32.008292158482945
l2 norm of gradients: 0.21175099294879696
l2 norm of weights: 5.780023121720105

Iteration Number: 6701
Loss: 32.00604376441095
l2 norm of gradients: 0.21173000029529587
l2 norm of weights: 5.779947751224307

Iteration Number: 6702
Loss: 32.00379581654638
l2 norm of gradients: 0.21170901306305548

```
12 norm of weights: 5.77987238701343
-----
Iteration Number: 6703
Loss: 32.00154831473293
12 norm of gradients: 0.21168803124935423
12 norm of weights: 5.779797029086412
-----
Iteration Number: 6704
Loss: 31.999301258812427
12 norm of gradients: 0.21166705485147275
12 norm of weights: 5.779721677442189
-----
Iteration Number: 6705
Loss: 31.997054648613243
12 norm of gradients: 0.21164608386669392
12 norm of weights: 5.779646332079699
-----
Iteration Number: 6706
Loss: 31.994808483994685
12 norm of gradients: 0.21162511829230266
12 norm of weights: 5.77957099299788
-----
Iteration Number: 6707
Loss: 31.99256276478306
12 norm of gradients: 0.21160415812558594
12 norm of weights: 5.779495660195673
-----
Iteration Number: 6708
Loss: 31.99031749082728
12 norm of gradients: 0.21158320336383302
12 norm of weights: 5.7794203336720145
-----
Iteration Number: 6709
Loss: 31.98807266195865
12 norm of gradients: 0.21156225400433518
12 norm of weights: 5.779345013425844
-----
Iteration Number: 6710
Loss: 31.98582827803178
12 norm of gradients: 0.2115413100443859
12 norm of weights: 5.7792696994561
-----
Iteration Number: 6711
Loss: 31.983584338873484
12 norm of gradients: 0.21152037148128072
12 norm of weights: 5.779194391761725
-----
Iteration Number: 6712
```

Loss: 31.981340844337648
l2 norm of gradients: 0.2114994383123174
l2 norm of weights: 5.779119090341657

Iteration Number: 6713
Loss: 31.979097794257953
l2 norm of gradients: 0.2114785105347957
l2 norm of weights: 5.779043795194838

Iteration Number: 6714
Loss: 31.976855188475852
l2 norm of gradients: 0.21145758814601753
l2 norm of weights: 5.778968506320209

Iteration Number: 6715
Loss: 31.974613026835858
l2 norm of gradients: 0.21143667114328696
l2 norm of weights: 5.778893223716709

Iteration Number: 6716
Loss: 31.972371309175863
l2 norm of gradients: 0.21141575952391017
l2 norm of weights: 5.778817947383282

Iteration Number: 6717
Loss: 31.970130035342976
l2 norm of gradients: 0.2113948532851953
l2 norm of weights: 5.77874267731887

Iteration Number: 6718
Loss: 31.96788920517227
l2 norm of gradients: 0.2113739524244529
l2 norm of weights: 5.778667413522413

Iteration Number: 6719
Loss: 31.965648818513717
l2 norm of gradients: 0.21135305693899525
l2 norm of weights: 5.778592155992857

Iteration Number: 6720
Loss: 31.9634088751995
l2 norm of gradients: 0.21133216682613712
l2 norm of weights: 5.778516904729142

Iteration Number: 6721
Loss: 31.961169375078946
l2 norm of gradients: 0.2113112820831951
l2 norm of weights: 5.778441659730213

```

-----
Iteration Number: 6722
Loss: 31.9589303179923
l2 norm of gradients: 0.21129040270748792
l2 norm of weights: 5.778366420995013
-----
Iteration Number: 6723
Loss: 31.956691703781384
l2 norm of gradients: 0.2112695286963365
l2 norm of weights: 5.778291188522487
-----
Iteration Number: 6724
Loss: 31.95445353228901
l2 norm of gradients: 0.21124866004706375
l2 norm of weights: 5.7782159623115765
-----
Iteration Number: 6725
Loss: 31.95221580335492
l2 norm of gradients: 0.2112277967569948
l2 norm of weights: 5.778140742361231
-----
Iteration Number: 6726
Loss: 31.94997851682482
l2 norm of gradients: 0.2112069388234568
l2 norm of weights: 5.778065528670392
-----
Iteration Number: 6727
Loss: 31.947741672540626
l2 norm of gradients: 0.2111860862437789
l2 norm of weights: 5.777990321238006
-----
Iteration Number: 6728
Loss: 31.94550527034341
l2 norm of gradients: 0.21116523901529247
l2 norm of weights: 5.777915120063018
-----
Iteration Number: 6729
Loss: 31.94326931007342
l2 norm of gradients: 0.2111443971353309
l2 norm of weights: 5.7778399251443755
-----
Iteration Number: 6730
Loss: 31.941033791578082
l2 norm of gradients: 0.21112356060122958
l2 norm of weights: 5.777764736481024
-----
Iteration Number: 6731
Loss: 31.938798714701797

```

l2 norm of gradients: 0.21110272941032618
l2 norm of weights: 5.77768955407191

Iteration Number: 6732
Loss: 31.93656407928098
l2 norm of gradients: 0.21108190355996023
l2 norm of weights: 5.777614377915982

Iteration Number: 6733
Loss: 31.93432988516738
l2 norm of gradients: 0.21106108304747345
l2 norm of weights: 5.777539208012186

Iteration Number: 6734
Loss: 31.93209613219306
l2 norm of gradients: 0.21104026787020966
l2 norm of weights: 5.77746404435947

Iteration Number: 6735
Loss: 31.92986282020944
l2 norm of gradients: 0.21101945802551464
l2 norm of weights: 5.777388886956783

Iteration Number: 6736
Loss: 31.927629949061085
l2 norm of gradients: 0.21099865351073618
l2 norm of weights: 5.777313735803072

Iteration Number: 6737
Loss: 31.92539751858467
l2 norm of gradients: 0.21097785432322447
l2 norm of weights: 5.777238590897288

Iteration Number: 6738
Loss: 31.92316552862779
l2 norm of gradients: 0.21095706046033139
l2 norm of weights: 5.777163452238379

Iteration Number: 6739
Loss: 31.92093397903432
l2 norm of gradients: 0.21093627191941094
l2 norm of weights: 5.777088319825294

Iteration Number: 6740
Loss: 31.91870286964461
l2 norm of gradients: 0.2109154886978194
l2 norm of weights: 5.777013193656984

Iteration Number: 6741
Loss: 31.916472200307503
l2 norm of gradients: 0.21089471079291486
l2 norm of weights: 5.776938073732397

Iteration Number: 6742
Loss: 31.914241970860967
l2 norm of gradients: 0.2108739382020576
l2 norm of weights: 5.776862960050487

Iteration Number: 6743
Loss: 31.912012181156026
l2 norm of gradients: 0.2108531709226099
l2 norm of weights: 5.776787852610203

Iteration Number: 6744
Loss: 31.90978283103018
l2 norm of gradients: 0.21083240895193608
l2 norm of weights: 5.7767127514104954

Iteration Number: 6745
Loss: 31.9075539203329
l2 norm of gradients: 0.21081165228740253
l2 norm of weights: 5.7766376564503155

Iteration Number: 6746
Loss: 31.905325448901426
l2 norm of gradients: 0.2107909009263776
l2 norm of weights: 5.776562567728618

Iteration Number: 6747
Loss: 31.90309741658852
l2 norm of gradients: 0.21077015486623185
l2 norm of weights: 5.7764874852443535

Iteration Number: 6748
Loss: 31.9008698232309
l2 norm of gradients: 0.21074941410433767
l2 norm of weights: 5.776412408996474

Iteration Number: 6749
Loss: 31.8986426686811
l2 norm of gradients: 0.21072867863806968
l2 norm of weights: 5.776337338983933

Iteration Number: 6750
Loss: 31.896415952776355
l2 norm of gradients: 0.21070794846480434

```

12 norm of weights: 5.776262275205683
-----
Iteration Number: 6751
Loss: 31.894189675365016
12 norm of gradients: 0.21068722358192032
12 norm of weights: 5.776187217660679
-----
Iteration Number: 6752
Loss: 31.89196383628982
12 norm of gradients: 0.21066650398679823
12 norm of weights: 5.7761121663478745
-----
Iteration Number: 6753
Loss: 31.88973843539481
12 norm of gradients: 0.21064578967682063
12 norm of weights: 5.7760371212662225
-----
Iteration Number: 6754
Loss: 31.887513472531666
12 norm of gradients: 0.21062508064937233
12 norm of weights: 5.7759620824146785
-----
Iteration Number: 6755
Loss: 31.885288947541454
12 norm of gradients: 0.21060437690183992
12 norm of weights: 5.775887049792199
-----
Iteration Number: 6756
Loss: 31.88306486026421
12 norm of gradients: 0.21058367843161202
12 norm of weights: 5.775812023397736
-----
Iteration Number: 6757
Loss: 31.88084121054977
12 norm of gradients: 0.21056298523607953
12 norm of weights: 5.775737003230249
-----
Iteration Number: 6758
Loss: 31.878617998244064
12 norm of gradients: 0.21054229731263507
12 norm of weights: 5.775661989288691
-----
Iteration Number: 6759
Loss: 31.876395223190052
12 norm of gradients: 0.21052161465867353
12 norm of weights: 5.775586981572019
-----
Iteration Number: 6760

```

Loss: 31.874172885238742
l2 norm of gradients: 0.2105009372715915
l2 norm of weights: 5.77551198007919

Iteration Number: 6761
Loss: 31.87195098422676
l2 norm of gradients: 0.21048026514878776
l2 norm of weights: 5.775436984809161

Iteration Number: 6762
Loss: 31.86972952000462
l2 norm of gradients: 0.21045959828766317
l2 norm of weights: 5.775361995760889

Iteration Number: 6763
Loss: 31.86750849241975
l2 norm of gradients: 0.21043893668562047
l2 norm of weights: 5.775287012933331

Iteration Number: 6764
Loss: 31.86528790131422
l2 norm of gradients: 0.2104182803400644
l2 norm of weights: 5.775212036325447

Iteration Number: 6765
Loss: 31.863067746541073
l2 norm of gradients: 0.21039762924840175
l2 norm of weights: 5.775137065936195

Iteration Number: 6766
Loss: 31.86084802793112
l2 norm of gradients: 0.2103769834080413
l2 norm of weights: 5.775062101764531

Iteration Number: 6767
Loss: 31.858628745348003
l2 norm of gradients: 0.21035634281639384
l2 norm of weights: 5.774987143809416

Iteration Number: 6768
Loss: 31.856409898626467
l2 norm of gradients: 0.21033570747087207
l2 norm of weights: 5.774912192069809

Iteration Number: 6769
Loss: 31.85419148761852
l2 norm of gradients: 0.21031507736889077
l2 norm of weights: 5.774837246544671


```

-----
Iteration Number: 6770
Loss: 31.85197351216924
l2 norm of gradients: 0.2102944525078666
l2 norm of weights: 5.77476230723296
-----
Iteration Number: 6771
Loss: 31.84975597212233
l2 norm of gradients: 0.2102738328852184
l2 norm of weights: 5.774687374133636
-----
Iteration Number: 6772
Loss: 31.847538867325753
l2 norm of gradients: 0.21025321849836667
l2 norm of weights: 5.774612447245662
-----
Iteration Number: 6773
Loss: 31.845322197626142
l2 norm of gradients: 0.21023260934473434
l2 norm of weights: 5.774537526567998
-----
Iteration Number: 6774
Loss: 31.84310596287065
l2 norm of gradients: 0.2102120054217459
l2 norm of weights: 5.774462612099604
-----
Iteration Number: 6775
Loss: 31.840890162905808
l2 norm of gradients: 0.21019140672682796
l2 norm of weights: 5.7743877038394436
-----
Iteration Number: 6776
Loss: 31.838674797577916
l2 norm of gradients: 0.21017081325740922
l2 norm of weights: 5.774312801786477
-----
Iteration Number: 6777
Loss: 31.836459866737783
l2 norm of gradients: 0.21015022501092023
l2 norm of weights: 5.7742379059396685
-----
Iteration Number: 6778
Loss: 31.834245370227904
l2 norm of gradients: 0.21012964198479347
l2 norm of weights: 5.77416301629798
-----
Iteration Number: 6779
Loss: 31.832031307896006

```

```

12 norm of gradients: 0.21010906417646358
12 norm of weights: 5.774088132860373
-----
Iteration Number: 6780
Loss: 31.82981767958725
12 norm of gradients: 0.21008849158336695
12 norm of weights: 5.774013255625813
-----
Iteration Number: 6781
Loss: 31.827604485155003
12 norm of gradients: 0.21006792420294199
12 norm of weights: 5.773938384593263
-----
Iteration Number: 6782
Loss: 31.825391724440557
12 norm of gradients: 0.21004736203262916
12 norm of weights: 5.7738635197616865
-----
Iteration Number: 6783
Loss: 31.823179397295405
12 norm of gradients: 0.21002680506987084
12 norm of weights: 5.773788661130047
-----
Iteration Number: 6784
Loss: 31.82096750356306
12 norm of gradients: 0.21000625331211126
12 norm of weights: 5.7737138086973125
-----
Iteration Number: 6785
Loss: 31.818756043097906
12 norm of gradients: 0.20998570675679676
12 norm of weights: 5.773638962462445
-----
Iteration Number: 6786
Loss: 31.816545015740353
12 norm of gradients: 0.2099651654013755
12 norm of weights: 5.7735641224244105
-----
Iteration Number: 6787
Loss: 31.81433442134114
12 norm of gradients: 0.20994462924329774
12 norm of weights: 5.773489288582176
-----
Iteration Number: 6788
Loss: 31.81212425975061
12 norm of gradients: 0.2099240982800155
12 norm of weights: 5.7734144609347045
-----

```

Iteration Number: 6789
Loss: 31.80991453081423
l2 norm of gradients: 0.20990357250898292
l2 norm of weights: 5.773339639480965

Iteration Number: 6790
Loss: 31.807705234378133
l2 norm of gradients: 0.2098830519276559
l2 norm of weights: 5.773264824219923

Iteration Number: 6791
Loss: 31.80549637029439
l2 norm of gradients: 0.20986253653349252
l2 norm of weights: 5.773190015150546

Iteration Number: 6792
Loss: 31.803287938406054
l2 norm of gradients: 0.20984202632395252
l2 norm of weights: 5.773115212271802

Iteration Number: 6793
Loss: 31.801079938566954
l2 norm of gradients: 0.20982152129649784
l2 norm of weights: 5.773040415582657

Iteration Number: 6794
Loss: 31.79887237061976
l2 norm of gradients: 0.20980102144859214
l2 norm of weights: 5.77296562508208

Iteration Number: 6795
Loss: 31.796665234418374
l2 norm of gradients: 0.20978052677770118
l2 norm of weights: 5.77289084076904

Iteration Number: 6796
Loss: 31.794458529812808
l2 norm of gradients: 0.20976003728129256
l2 norm of weights: 5.772816062642504

Iteration Number: 6797
Loss: 31.792252256646577
l2 norm of gradients: 0.20973955295683575
l2 norm of weights: 5.772741290701442

Iteration Number: 6798
Loss: 31.790046414765296
l2 norm of gradients: 0.20971907380180224

```

12 norm of weights: 5.772666524944823
-----
Iteration Number: 6799
Loss: 31.787841004028067
12 norm of gradients: 0.2096985998136655
12 norm of weights: 5.772591765371616
-----
Iteration Number: 6800
Loss: 31.78563602427545
12 norm of gradients: 0.20967813098990076
12 norm of weights: 5.772517011980794
-----
Iteration Number: 6801
Loss: 31.783431475357418
12 norm of gradients: 0.20965766732798535
12 norm of weights: 5.772442264771323
-----
Iteration Number: 6802
Loss: 31.78122735712941
12 norm of gradients: 0.20963720882539824
12 norm of weights: 5.772367523742176
-----
Iteration Number: 6803
Loss: 31.779023669433435
12 norm of gradients: 0.20961675547962066
12 norm of weights: 5.772292788892325
-----
Iteration Number: 6804
Loss: 31.776820412118724
12 norm of gradients: 0.20959630728813555
12 norm of weights: 5.772218060220737
-----
Iteration Number: 6805
Loss: 31.77461758503817
12 norm of gradients: 0.20957586424842767
12 norm of weights: 5.77214333772639
-----
Iteration Number: 6806
Loss: 31.772415188045862
12 norm of gradients: 0.20955542635798396
12 norm of weights: 5.77206862140825
-----
Iteration Number: 6807
Loss: 31.7702132209807
12 norm of gradients: 0.209534993614293
12 norm of weights: 5.771993911265294
-----
Iteration Number: 6808

```

Loss: 31.768011683693526
l2 norm of gradients: 0.20951456601484553
l2 norm of weights: 5.771919207296491

Iteration Number: 6809
Loss: 31.76581057604562
l2 norm of gradients: 0.2094941435571339
l2 norm of weights: 5.771844509500815

Iteration Number: 6810
Loss: 31.763609897871607
l2 norm of gradients: 0.20947372623865262
l2 norm of weights: 5.77176981787724

Iteration Number: 6811
Loss: 31.761409649030902
l2 norm of gradients: 0.20945331405689796
l2 norm of weights: 5.771695132424739

Iteration Number: 6812
Loss: 31.759209829372253
l2 norm of gradients: 0.2094329070093681
l2 norm of weights: 5.771620453142287

Iteration Number: 6813
Loss: 31.75701043874279
l2 norm of gradients: 0.2094125050935631
l2 norm of weights: 5.771545780028856

Iteration Number: 6814
Loss: 31.75481147699414
l2 norm of gradients: 0.20939210830698501
l2 norm of weights: 5.7714711130834235

Iteration Number: 6815
Loss: 31.752612943975276
l2 norm of gradients: 0.2093717166471376
l2 norm of weights: 5.771396452304962

Iteration Number: 6816
Loss: 31.750414839539403
l2 norm of gradients: 0.20935133011152668
l2 norm of weights: 5.771321797692447

Iteration Number: 6817
Loss: 31.748217163532416
l2 norm of gradients: 0.20933094869765992
l2 norm of weights: 5.7712471492448545

```

-----
Iteration Number: 6818
Loss: 31.746019915807224
l2 norm of gradients: 0.2093105724030468
l2 norm of weights: 5.771172506961162
-----
Iteration Number: 6819
Loss: 31.7438230962143
l2 norm of gradients: 0.2092902012251987
l2 norm of weights: 5.771097870840343
-----
Iteration Number: 6820
Loss: 31.741626704607082
l2 norm of gradients: 0.20926983516162898
l2 norm of weights: 5.771023240881375
-----
Iteration Number: 6821
Loss: 31.739430740830322
l2 norm of gradients: 0.20924947420985265
l2 norm of weights: 5.770948617083236
-----
Iteration Number: 6822
Loss: 31.73723520473896
l2 norm of gradients: 0.20922911836738686
l2 norm of weights: 5.770873999444901
-----
Iteration Number: 6823
Loss: 31.73504009618113
l2 norm of gradients: 0.2092087676317505
l2 norm of weights: 5.770799387965349
-----
Iteration Number: 6824
Loss: 31.732845415010715
l2 norm of gradients: 0.20918842200046425
l2 norm of weights: 5.770724782643558
-----
Iteration Number: 6825
Loss: 31.73065116107606
l2 norm of gradients: 0.20916808147105084
l2 norm of weights: 5.770650183478507
-----
Iteration Number: 6826
Loss: 31.728457334229446
l2 norm of gradients: 0.20914774604103473
l2 norm of weights: 5.770575590469172
-----
Iteration Number: 6827
Loss: 31.726263934322972

```

```

12 norm of gradients: 0.20912741570794233
12 norm of weights: 5.770501003614534
-----
Iteration Number: 6828
Loss: 31.724070961205836
12 norm of gradients: 0.20910709046930182
12 norm of weights: 5.77042642291357
-----
Iteration Number: 6829
Loss: 31.721878414731034
12 norm of gradients: 0.20908677032264328
12 norm of weights: 5.770351848365261
-----
Iteration Number: 6830
Loss: 31.719686294748588
12 norm of gradients: 0.2090664552654987
12 norm of weights: 5.770277279968587
-----
Iteration Number: 6831
Loss: 31.717494601111923
12 norm of gradients: 0.20904614529540183
12 norm of weights: 5.770202717722527
-----
Iteration Number: 6832
Loss: 31.715303333667244
12 norm of gradients: 0.20902584040988834
12 norm of weights: 5.770128161626063
-----
Iteration Number: 6833
Loss: 31.71311249227697
12 norm of gradients: 0.20900554060649573
12 norm of weights: 5.770053611678175
-----
Iteration Number: 6834
Loss: 31.710922076780683
12 norm of gradients: 0.20898524588276338
12 norm of weights: 5.769979067877842
-----
Iteration Number: 6835
Loss: 31.708732087037003
12 norm of gradients: 0.2089649562362324
12 norm of weights: 5.769904530224049
-----
Iteration Number: 6836
Loss: 31.70654252289718
12 norm of gradients: 0.20894467166444594
12 norm of weights: 5.7698299987157755
-----

```

Iteration Number: 6837
Loss: 31.704353384211345
l2 norm of gradients: 0.20892439216494882
l2 norm of weights: 5.769755473352005

Iteration Number: 6838
Loss: 31.70216467083475
l2 norm of gradients: 0.20890411773528775
l2 norm of weights: 5.769680954131717

Iteration Number: 6839
Loss: 31.699976382617976
l2 norm of gradients: 0.20888384837301135
l2 norm of weights: 5.769606441053897

Iteration Number: 6840
Loss: 31.69778851940828
l2 norm of gradients: 0.20886358407567
l2 norm of weights: 5.769531934117528

Iteration Number: 6841
Loss: 31.695601081067398
l2 norm of gradients: 0.20884332484081586
l2 norm of weights: 5.769457433321591

Iteration Number: 6842
Loss: 31.693414067435768
l2 norm of gradients: 0.20882307066600314
l2 norm of weights: 5.769382938665072

Iteration Number: 6843
Loss: 31.69122747837975
l2 norm of gradients: 0.20880282154878757
l2 norm of weights: 5.769308450146955

Iteration Number: 6844
Loss: 31.689041313739033
l2 norm of gradients: 0.208782577486727
l2 norm of weights: 5.7692339677662225

Iteration Number: 6845
Loss: 31.68685557337848
l2 norm of gradients: 0.2087623384773809
l2 norm of weights: 5.76915949152186

Iteration Number: 6846
Loss: 31.684670257140517
l2 norm of gradients: 0.2087421045183107


```
12 norm of weights: 5.769085021412853
-----
Iteration Number: 6847
Loss: 31.682485364883227
12 norm of gradients: 0.20872187560707967
12 norm of weights: 5.769010557438187
-----
Iteration Number: 6848
Loss: 31.6803008964566
12 norm of gradients: 0.2087016517412526
12 norm of weights: 5.768936099596846
-----
Iteration Number: 6849
Loss: 31.678116851713177
12 norm of gradients: 0.20868143291839644
12 norm of weights: 5.768861647887817
-----
Iteration Number: 6850
Loss: 31.67593323051417
12 norm of gradients: 0.20866121913607996
12 norm of weights: 5.768787202310087
-----
Iteration Number: 6851
Loss: 31.67375003270061
12 norm of gradients: 0.20864101039187344
12 norm of weights: 5.7687127628626405
-----
Iteration Number: 6852
Loss: 31.671567258131432
12 norm of gradients: 0.2086208066833492
12 norm of weights: 5.7686383295444665
-----
Iteration Number: 6853
Loss: 31.669384906665577
12 norm of gradients: 0.20860060800808142
12 norm of weights: 5.7685639023545505
-----
Iteration Number: 6854
Loss: 31.66720297814824
12 norm of gradients: 0.20858041436364588
12 norm of weights: 5.768489481291882
-----
Iteration Number: 6855
Loss: 31.665021472432763
12 norm of gradients: 0.2085602257476203
12 norm of weights: 5.7684150663554465
-----
Iteration Number: 6856
```

Loss: 31.6628403893802
l2 norm of gradients: 0.20854004215758418
l2 norm of weights: 5.7683406575442335

Iteration Number: 6857
Loss: 31.660659728836865
l2 norm of gradients: 0.20851986359111885
l2 norm of weights: 5.768266254857233

Iteration Number: 6858
Loss: 31.658479490658422
l2 norm of gradients: 0.20849969004580743
l2 norm of weights: 5.76819185829343

Iteration Number: 6859
Loss: 31.656299674697348
l2 norm of gradients: 0.20847952151923474
l2 norm of weights: 5.768117467851816

Iteration Number: 6860
Loss: 31.654120280811505
l2 norm of gradients: 0.20845935800898754
l2 norm of weights: 5.768043083531381

Iteration Number: 6861
Loss: 31.651941308853818
l2 norm of gradients: 0.2084391995126543
l2 norm of weights: 5.767968705331113

Iteration Number: 6862
Loss: 31.649762758671727
l2 norm of gradients: 0.20841904602782527
l2 norm of weights: 5.767894333250005

Iteration Number: 6863
Loss: 31.647584630132386
l2 norm of gradients: 0.20839889755209257
l2 norm of weights: 5.7678199672870445

Iteration Number: 6864
Loss: 31.645406923076038
l2 norm of gradients: 0.20837875408305
l2 norm of weights: 5.767745607441223

Iteration Number: 6865
Loss: 31.643229637364442
l2 norm of gradients: 0.20835861561829325
l2 norm of weights: 5.767671253711531

```

-----
Iteration Number: 6866
Loss: 31.641052772851264
l2 norm of gradients: 0.2083384821554197
l2 norm of weights: 5.767596906096962
-----
Iteration Number: 6867
Loss: 31.63887632939136
l2 norm of gradients: 0.20831835369202858
l2 norm of weights: 5.767522564596505
-----
Iteration Number: 6868
Loss: 31.636700306835902
l2 norm of gradients: 0.20829823022572083
l2 norm of weights: 5.767448229209154
-----
Iteration Number: 6869
Loss: 31.634524705041148
l2 norm of gradients: 0.20827811175409933
l2 norm of weights: 5.7673738999339
-----
Iteration Number: 6870
Loss: 31.632349523866004
l2 norm of gradients: 0.2082579982747685
l2 norm of weights: 5.767299576769737
-----
Iteration Number: 6871
Loss: 31.63017476315772
l2 norm of gradients: 0.20823788978533472
l2 norm of weights: 5.767225259715657
-----
Iteration Number: 6872
Loss: 31.628000422775763
l2 norm of gradients: 0.20821778628340598
l2 norm of weights: 5.767150948770652
-----
Iteration Number: 6873
Loss: 31.625826502571975
l2 norm of gradients: 0.20819768776659225
l2 norm of weights: 5.767076643933719
-----
Iteration Number: 6874
Loss: 31.62365300240887
l2 norm of gradients: 0.208177594232505
l2 norm of weights: 5.767002345203849
-----
Iteration Number: 6875
Loss: 31.6214799221316

```

```

12 norm of gradients: 0.20815750567875782
12 norm of weights: 5.766928052580037
-----
Iteration Number: 6876
Loss: 31.619307261599857
12 norm of gradients: 0.20813742210296568
12 norm of weights: 5.766853766061278
-----
Iteration Number: 6877
Loss: 31.61713502067152
12 norm of gradients: 0.20811734350274555
12 norm of weights: 5.766779485646566
-----
Iteration Number: 6878
Loss: 31.614963199197238
12 norm of gradients: 0.2080972698757161
12 norm of weights: 5.766705211334897
-----
Iteration Number: 6879
Loss: 31.612791797032635
12 norm of gradients: 0.2080772012194978
12 norm of weights: 5.766630943125266
-----
Iteration Number: 6880
Loss: 31.61062081403594
12 norm of gradients: 0.20805713753171276
12 norm of weights: 5.766556681016669
-----
Iteration Number: 6881
Loss: 31.60845025006178
12 norm of gradients: 0.20803707880998493
12 norm of weights: 5.766482425008102
-----
Iteration Number: 6882
Loss: 31.6062801049641
12 norm of gradients: 0.20801702505193997
12 norm of weights: 5.7664081750985625
-----
Iteration Number: 6883
Loss: 31.604110378604975
12 norm of gradients: 0.2079969762552054
12 norm of weights: 5.766333931287045
-----
Iteration Number: 6884
Loss: 31.60194107082844
12 norm of gradients: 0.2079769324174104
12 norm of weights: 5.7662596935725485
-----

```

Iteration Number: 6885
Loss: 31.599772181500544
l2 norm of gradients: 0.20795689353618582
l2 norm of weights: 5.766185461954069

Iteration Number: 6886
Loss: 31.59760371047466
l2 norm of gradients: 0.20793685960916436
l2 norm of weights: 5.766111236430605

Iteration Number: 6887
Loss: 31.595435657602955
l2 norm of gradients: 0.20791683063398045
l2 norm of weights: 5.766037017001156

Iteration Number: 6888
Loss: 31.593268022748326
l2 norm of gradients: 0.20789680660827028
l2 norm of weights: 5.765962803664717

Iteration Number: 6889
Loss: 31.59110080575757
l2 norm of gradients: 0.20787678752967167
l2 norm of weights: 5.765888596420288

Iteration Number: 6890
Loss: 31.588934006495073
l2 norm of gradients: 0.20785677339582428
l2 norm of weights: 5.76581439526687

Iteration Number: 6891
Loss: 31.58676762481373
l2 norm of gradients: 0.20783676420436953
l2 norm of weights: 5.765740200203459

Iteration Number: 6892
Loss: 31.584601660569405
l2 norm of gradients: 0.2078167599529504
l2 norm of weights: 5.765666011229056

Iteration Number: 6893
Loss: 31.582436113622098
l2 norm of gradients: 0.20779676063921187
l2 norm of weights: 5.765591828342663

Iteration Number: 6894
Loss: 31.58027098382533
l2 norm of gradients: 0.2077767662608003

```

12 norm of weights: 5.765517651543276
-----
Iteration Number: 6895
Loss: 31.578106271036127
12 norm of gradients: 0.2077567768153642
12 norm of weights: 5.7654434808299
-----
Iteration Number: 6896
Loss: 31.575941975111533
12 norm of gradients: 0.20773679230055334
12 norm of weights: 5.765369316201532
-----
Iteration Number: 6897
Loss: 31.573778095905446
12 norm of gradients: 0.20771681271401957
12 norm of weights: 5.7652951576571745
-----
Iteration Number: 6898
Loss: 31.571614633283076
12 norm of gradients: 0.20769683805341635
12 norm of weights: 5.76522100519583
-----
Iteration Number: 6899
Loss: 31.569451587090704
12 norm of gradients: 0.20767686831639884
12 norm of weights: 5.765146858816499
-----
Iteration Number: 6900
Loss: 31.567288957192723
12 norm of gradients: 0.20765690350062382
12 norm of weights: 5.765072718518185
-----
Iteration Number: 6901
Loss: 31.565126743445656
12 norm of gradients: 0.20763694360375004
12 norm of weights: 5.764998584299889
-----
Iteration Number: 6902
Loss: 31.56296494570219
12 norm of gradients: 0.20761698862343772
12 norm of weights: 5.764924456160615
-----
Iteration Number: 6903
Loss: 31.560803563825885
12 norm of gradients: 0.20759703855734896
12 norm of weights: 5.764850334099363
-----
Iteration Number: 6904

```

Loss: 31.558642597663802
l2 norm of gradients: 0.20757709340314734
l2 norm of weights: 5.76477621811514

Iteration Number: 6905
Loss: 31.556482047084312
l2 norm of gradients: 0.20755715315849838
l2 norm of weights: 5.764702108206947

Iteration Number: 6906
Loss: 31.554321911943198
l2 norm of gradients: 0.20753721782106935
l2 norm of weights: 5.76462800437379

Iteration Number: 6907
Loss: 31.552162192090886
l2 norm of gradients: 0.20751728738852884
l2 norm of weights: 5.764553906614673

Iteration Number: 6908
Loss: 31.550002887393994
l2 norm of gradients: 0.20749736185854756
l2 norm of weights: 5.7644798149286

Iteration Number: 6909
Loss: 31.547843997702802
l2 norm of gradients: 0.2074774412287978
l2 norm of weights: 5.764405729314575

Iteration Number: 6910
Loss: 31.545685522879403
l2 norm of gradients: 0.2074575254969534
l2 norm of weights: 5.764331649771604

Iteration Number: 6911
Loss: 31.54352746278093
l2 norm of gradients: 0.20743761466069002
l2 norm of weights: 5.764257576298694

Iteration Number: 6912
Loss: 31.541369817262566
l2 norm of gradients: 0.20741770871768497
l2 norm of weights: 5.764183508894847

Iteration Number: 6913
Loss: 31.539212586188444
l2 norm of gradients: 0.20739780766561736
l2 norm of weights: 5.764109447559074

```
-----  
Iteration Number: 6914  
Loss: 31.537055769409445  
l2 norm of gradients: 0.2073779115021678  
l2 norm of weights: 5.764035392290379  
-----  
Iteration Number: 6915  
Loss: 31.53489936678882  
l2 norm of gradients: 0.2073580202250188  
l2 norm of weights: 5.763961343087767  
-----  
Iteration Number: 6916  
Loss: 31.53274337817972  
l2 norm of gradients: 0.20733813383185437  
l2 norm of weights: 5.763887299950247  
-----  
Iteration Number: 6917  
Loss: 31.53058780345019  
l2 norm of gradients: 0.20731825232036039  
l2 norm of weights: 5.763813262876828  
-----  
Iteration Number: 6918  
Loss: 31.528432642449403  
l2 norm of gradients: 0.20729837568822412  
l2 norm of weights: 5.763739231866515  
-----  
Iteration Number: 6919  
Loss: 31.52627789503846  
l2 norm of gradients: 0.2072785039331349  
l2 norm of weights: 5.763665206918317  
-----  
Iteration Number: 6920  
Loss: 31.524123561075502  
l2 norm of gradients: 0.20725863705278347  
l2 norm of weights: 5.763591188031242  
-----  
Iteration Number: 6921  
Loss: 31.52196964041958  
l2 norm of gradients: 0.2072387750448622  
l2 norm of weights: 5.763517175204299  
-----  
Iteration Number: 6922  
Loss: 31.51981613293449  
l2 norm of gradients: 0.2072189179070655  
l2 norm of weights: 5.763443168436496  
-----  
Iteration Number: 6923  
Loss: 31.517663038469223
```



```

12 norm of gradients: 0.207199065637089
12 norm of weights: 5.763369167726844
-----
Iteration Number: 6924
Loss: 31.515510356892854
12 norm of gradients: 0.20717921823263033
12 norm of weights: 5.763295173074351
-----
Iteration Number: 6925
Loss: 31.513358088056066
12 norm of gradients: 0.20715937569138862
12 norm of weights: 5.763221184478027
-----
Iteration Number: 6926
Loss: 31.511206231820733
12 norm of gradients: 0.20713953801106472
12 norm of weights: 5.763147201936884
-----
Iteration Number: 6927
Loss: 31.509054788046623
12 norm of gradients: 0.2071197051893612
12 norm of weights: 5.76307322544993
-----
Iteration Number: 6928
Loss: 31.506903756592255
12 norm of gradients: 0.20709987722398215
12 norm of weights: 5.762999255016178
-----
Iteration Number: 6929
Loss: 31.504753137319042
12 norm of gradients: 0.20708005411263347
12 norm of weights: 5.762925290634637
-----
Iteration Number: 6930
Loss: 31.502602930083658
12 norm of gradients: 0.2070602358530227
12 norm of weights: 5.76285133230432
-----
Iteration Number: 6931
Loss: 31.500453134744575
12 norm of gradients: 0.20704042244285886
12 norm of weights: 5.762777380024236
-----
Iteration Number: 6932
Loss: 31.498303751166507
12 norm of gradients: 0.20702061387985288
12 norm of weights: 5.7627034337934
-----

```

Iteration Number: 6933
Loss: 31.496154779205316
l2 norm of gradients: 0.20700081016171717
l2 norm of weights: 5.762629493610825

Iteration Number: 6934
Loss: 31.49400621871918
l2 norm of gradients: 0.20698101128616594
l2 norm of weights: 5.762555559475519

Iteration Number: 6935
Loss: 31.491858069570338
l2 norm of gradients: 0.20696121725091488
l2 norm of weights: 5.7624816313865

Iteration Number: 6936
Loss: 31.489710331617292
l2 norm of gradients: 0.20694142805368146
l2 norm of weights: 5.762407709342779

Iteration Number: 6937
Loss: 31.487563004719558
l2 norm of gradients: 0.2069216436921847
l2 norm of weights: 5.762333793343369

Iteration Number: 6938
Loss: 31.485416088740234
l2 norm of gradients: 0.20690186416414533
l2 norm of weights: 5.762259883387285

Iteration Number: 6939
Loss: 31.483269583539325
l2 norm of gradients: 0.20688208946728573
l2 norm of weights: 5.762185979473541

Iteration Number: 6940
Loss: 31.481123488969956
l2 norm of gradients: 0.2068623195993299
l2 norm of weights: 5.76211208160115

Iteration Number: 6941
Loss: 31.478977804898566
l2 norm of gradients: 0.20684255455800346
l2 norm of weights: 5.762038189769129

Iteration Number: 6942
Loss: 31.476832531183494
l2 norm of gradients: 0.2068227943410337

```
l2 norm of weights: 5.761964303976493
-----
Iteration Number: 6943
Loss: 31.474687667686982
l2 norm of gradients: 0.20680303894614954
l2 norm of weights: 5.761890424222254
-----
Iteration Number: 6944
Loss: 31.47254321426575
l2 norm of gradients: 0.20678328837108154
l2 norm of weights: 5.761816550505431
-----
Iteration Number: 6945
Loss: 31.470399170784464
l2 norm of gradients: 0.2067635426135618
l2 norm of weights: 5.761742682825039
-----
Iteration Number: 6946
Loss: 31.46825553709952
l2 norm of gradients: 0.20674380167132422
l2 norm of weights: 5.761668821180095
-----
Iteration Number: 6947
Loss: 31.46611231307585
l2 norm of gradients: 0.20672406554210424
l2 norm of weights: 5.761594965569614
-----
Iteration Number: 6948
Loss: 31.46396949857446
l2 norm of gradients: 0.20670433422363887
l2 norm of weights: 5.761521115992616
-----
Iteration Number: 6949
Loss: 31.46182709344998
l2 norm of gradients: 0.2066846077136668
l2 norm of weights: 5.761447272448114
-----
Iteration Number: 6950
Loss: 31.4596850975656
l2 norm of gradients: 0.2066648860099285
l2 norm of weights: 5.761373434935128
-----
Iteration Number: 6951
Loss: 31.457543510784873
l2 norm of gradients: 0.20664516911016573
l2 norm of weights: 5.7612996034526756
-----
Iteration Number: 6952
```

Loss: 31.45540233296865
l2 norm of gradients: 0.20662545701212212
l2 norm of weights: 5.761225777999775

Iteration Number: 6953
Loss: 31.45326156397435
l2 norm of gradients: 0.20660574971354284
l2 norm of weights: 5.761151958575444

Iteration Number: 6954
Loss: 31.451121203668276
l2 norm of gradients: 0.2065860472121747
l2 norm of weights: 5.761078145178701

Iteration Number: 6955
Loss: 31.44898125190871
l2 norm of gradients: 0.20656634950576608
l2 norm of weights: 5.761004337808568

Iteration Number: 6956
Loss: 31.446841708556132
l2 norm of gradients: 0.206546656592067
l2 norm of weights: 5.76093053646406

Iteration Number: 6957
Loss: 31.44470257347551
l2 norm of gradients: 0.20652696846882915
l2 norm of weights: 5.760856741144201

Iteration Number: 6958
Loss: 31.442563846519946
l2 norm of gradients: 0.20650728513380567
l2 norm of weights: 5.7607829518480065

Iteration Number: 6959
Loss: 31.440425527562475
l2 norm of gradients: 0.20648760658475146
l2 norm of weights: 5.7607091685745

Iteration Number: 6960
Loss: 31.43828761645593
l2 norm of gradients: 0.20646793281942308
l2 norm of weights: 5.760635391322701

Iteration Number: 6961
Loss: 31.436150113069022
l2 norm of gradients: 0.20644826383557832
l2 norm of weights: 5.7605616200916305

```
-----  
Iteration Number: 6962  
Loss: 31.434013017257623  
l2 norm of gradients: 0.20642859963097715  
l2 norm of weights: 5.760487854880311  
-----  
Iteration Number: 6963  
Loss: 31.431876328884567  
l2 norm of gradients: 0.2064089402033807  
l2 norm of weights: 5.76041409568776  
-----  
Iteration Number: 6964  
Loss: 31.429740047813045  
l2 norm of gradients: 0.20638928555055178  
l2 norm of weights: 5.760340342513003  
-----  
Iteration Number: 6965  
Loss: 31.42760417390472  
l2 norm of gradients: 0.2063696356702549  
l2 norm of weights: 5.760266595355061  
-----  
Iteration Number: 6966  
Loss: 31.425468707020073  
l2 norm of gradients: 0.20634999056025607  
l2 norm of weights: 5.760192854212957  
-----  
Iteration Number: 6967  
Loss: 31.423333647025128  
l2 norm of gradients: 0.206330350218323  
l2 norm of weights: 5.760119119085712  
-----  
Iteration Number: 6968  
Loss: 31.42119899378164  
l2 norm of gradients: 0.2063107146422249  
l2 norm of weights: 5.76004538997235  
-----  
Iteration Number: 6969  
Loss: 31.419064747144958  
l2 norm of gradients: 0.20629108382973257  
l2 norm of weights: 5.759971666871895  
-----  
Iteration Number: 6970  
Loss: 31.41693090698388  
l2 norm of gradients: 0.2062714577786184  
l2 norm of weights: 5.759897949783369  
-----  
Iteration Number: 6971  
Loss: 31.414797473160778
```

```

12 norm of gradients: 0.20625183648665651
12 norm of weights: 5.7598242387057965
-----
Iteration Number: 6972
Loss: 31.412664445537487
12 norm of gradients: 0.20623221995162236
12 norm of weights: 5.759750533638202
-----
Iteration Number: 6973
Loss: 31.41053182397504
12 norm of gradients: 0.20621260817129317
12 norm of weights: 5.7596768345796105
-----
Iteration Number: 6974
Loss: 31.408399608337625
12 norm of gradients: 0.2061930011434477
12 norm of weights: 5.759603141529046
-----
Iteration Number: 6975
Loss: 31.40626779848608
12 norm of gradients: 0.20617339886586622
12 norm of weights: 5.759529454485533
-----
Iteration Number: 6976
Loss: 31.404136394280435
12 norm of gradients: 0.2061538013363307
12 norm of weights: 5.7594557734480984
-----
Iteration Number: 6977
Loss: 31.40200539559539
12 norm of gradients: 0.2061342085526246
12 norm of weights: 5.759382098415767
-----
Iteration Number: 6978
Loss: 31.399874802278564
12 norm of gradients: 0.20611462051253293
12 norm of weights: 5.759308429387565
-----
Iteration Number: 6979
Loss: 31.39774461420348
12 norm of gradients: 0.20609503721384245
12 norm of weights: 5.7592347663625185
-----
Iteration Number: 6980
Loss: 31.39561483123207
12 norm of gradients: 0.2060754586543412
12 norm of weights: 5.7591611093396535
-----

```

Iteration Number: 6981
Loss: 31.393485453224386
l2 norm of gradients: 0.20605588483181905
l2 norm of weights: 5.759087458317999

Iteration Number: 6982
Loss: 31.39135648004445
l2 norm of gradients: 0.20603631574406733
l2 norm of weights: 5.75901381329658

Iteration Number: 6983
Loss: 31.38922791155302
l2 norm of gradients: 0.2060167513888789
l2 norm of weights: 5.758940174274424

Iteration Number: 6984
Loss: 31.387099747620677
l2 norm of gradients: 0.20599719176404818
l2 norm of weights: 5.758866541250561

Iteration Number: 6985
Loss: 31.384971988103487
l2 norm of gradients: 0.20597763686737125
l2 norm of weights: 5.758792914224017

Iteration Number: 6986
Loss: 31.38284463286887
l2 norm of gradients: 0.20595808669664575
l2 norm of weights: 5.758719293193821

Iteration Number: 6987
Loss: 31.380717681778826
l2 norm of gradients: 0.20593854124967076
l2 norm of weights: 5.758645678159002

Iteration Number: 6988
Loss: 31.378591134699256
l2 norm of gradients: 0.20591900052424697
l2 norm of weights: 5.758572069118588

Iteration Number: 6989
Loss: 31.376464991489623
l2 norm of gradients: 0.20589946451817662
l2 norm of weights: 5.758498466071609

Iteration Number: 6990
Loss: 31.374339252019325
l2 norm of gradients: 0.2058799332292636

```
12 norm of weights: 5.758424869017096
-----
Iteration Number: 6991
Loss: 31.372213916151328
12 norm of gradients: 0.20586040665531322
12 norm of weights: 5.758351277954077
-----
Iteration Number: 6992
Loss: 31.370088983741944
12 norm of gradients: 0.20584088479413237
12 norm of weights: 5.758277692881583
-----
Iteration Number: 6993
Loss: 31.367964454664556
12 norm of gradients: 0.2058213676435296
12 norm of weights: 5.7582041137986435
-----
Iteration Number: 6994
Loss: 31.365840328780884
12 norm of gradients: 0.20580185520131478
12 norm of weights: 5.75813054070429
-----
Iteration Number: 6995
Loss: 31.363716605950923
12 norm of gradients: 0.20578234746529955
12 norm of weights: 5.7580569735975535
-----
Iteration Number: 6996
Loss: 31.361593286042403
12 norm of gradients: 0.205762844433297
12 norm of weights: 5.757983412477466
-----
Iteration Number: 6997
Loss: 31.359470368922082
12 norm of gradients: 0.2057433461031217
12 norm of weights: 5.757909857343058
-----
Iteration Number: 6998
Loss: 31.357347854447976
12 norm of gradients: 0.20572385247258992
12 norm of weights: 5.757836308193362
-----
Iteration Number: 6999
Loss: 31.355225742493225
12 norm of gradients: 0.20570436353951935
12 norm of weights: 5.75776276502741
-----
Iteration Number: 7000
```


Loss: 31.353104032910466
l2 norm of gradients: 0.20568487930172913
l2 norm of weights: 5.757689227844235

Iteration Number: 7001
Loss: 31.35098272557355
l2 norm of gradients: 0.20566539975704018
l2 norm of weights: 5.75761569664287

Iteration Number: 7002
Loss: 31.348861820344432
l2 norm of gradients: 0.20564592490327477
l2 norm of weights: 5.757542171422348

Iteration Number: 7003
Loss: 31.346741317087783
l2 norm of gradients: 0.20562645473825672
l2 norm of weights: 5.757468652181702

Iteration Number: 7004
Loss: 31.34462121567125
l2 norm of gradients: 0.20560698925981147
l2 norm of weights: 5.757395138919967

Iteration Number: 7005
Loss: 31.342501515952936
l2 norm of gradients: 0.20558752846576586
l2 norm of weights: 5.757321631636175

Iteration Number: 7006
Loss: 31.34038221780351
l2 norm of gradients: 0.20556807235394836
l2 norm of weights: 5.757248130329361

Iteration Number: 7007
Loss: 31.338263321088995
l2 norm of gradients: 0.2055486209221889
l2 norm of weights: 5.757174634998561

Iteration Number: 7008
Loss: 31.33614482566974
l2 norm of gradients: 0.20552917416831895
l2 norm of weights: 5.757101145642809

Iteration Number: 7009
Loss: 31.33402673141233
l2 norm of gradients: 0.20550973209017162
l2 norm of weights: 5.75702766226114

```

-----
Iteration Number: 7010
Loss: 31.331909038186197
l2 norm of gradients: 0.2054902946855813
l2 norm of weights: 5.7569541848525905
-----
Iteration Number: 7011
Loss: 31.32979174585185
l2 norm of gradients: 0.20547086195238404
l2 norm of weights: 5.756880713416195
-----
Iteration Number: 7012
Loss: 31.327674854276086
l2 norm of gradients: 0.20545143388841747
l2 norm of weights: 5.756807247950991
-----
Iteration Number: 7013
Loss: 31.325558363324106
l2 norm of gradients: 0.20543201049152063
l2 norm of weights: 5.756733788456014
-----
Iteration Number: 7014
Loss: 31.323442272860014
l2 norm of gradients: 0.20541259175953405
l2 norm of weights: 5.7566603349303005
-----
Iteration Number: 7015
Loss: 31.32132658275708
l2 norm of gradients: 0.20539317769029997
l2 norm of weights: 5.756586887372888
-----
Iteration Number: 7016
Loss: 31.319211292870193
l2 norm of gradients: 0.20537376828166176
l2 norm of weights: 5.756513445782813
-----
Iteration Number: 7017
Loss: 31.317096403071353
l2 norm of gradients: 0.20535436353146475
l2 norm of weights: 5.756440010159114
-----
Iteration Number: 7018
Loss: 31.314981913228074
l2 norm of gradients: 0.2053349634375554
l2 norm of weights: 5.7563665805008295
-----
Iteration Number: 7019
Loss: 31.312867823199028

```

```
12 norm of gradients: 0.20531556799778192
12 norm of weights: 5.756293156806996
-----
Iteration Number: 7020
Loss: 31.31075413285671
12 norm of gradients: 0.20529617720999394
12 norm of weights: 5.756219739076653
-----
Iteration Number: 7021
Loss: 31.308640842066392
12 norm of gradients: 0.20527679107204255
12 norm of weights: 5.75614632730884
-----
Iteration Number: 7022
Loss: 31.30652795069226
12 norm of gradients: 0.2052574095817804
12 norm of weights: 5.756072921502594
-----
Iteration Number: 7023
Loss: 31.304415458595745
12 norm of gradients: 0.20523803273706162
12 norm of weights: 5.755999521656955
-----
Iteration Number: 7024
Loss: 31.302303365654605
12 norm of gradients: 0.20521866053574184
12 norm of weights: 5.755926127770964
-----
Iteration Number: 7025
Loss: 31.3001916717269
12 norm of gradients: 0.2051992929756781
12 norm of weights: 5.75585273984366
-----
Iteration Number: 7026
Loss: 31.29808037667983
12 norm of gradients: 0.20517993005472906
12 norm of weights: 5.755779357874084
-----
Iteration Number: 7027
Loss: 31.29596948038557
12 norm of gradients: 0.20516057177075486
12 norm of weights: 5.755705981861276
-----
Iteration Number: 7028
Loss: 31.293858982702023
12 norm of gradients: 0.20514121812161706
12 norm of weights: 5.755632611804275
-----
```

Iteration Number: 7029
Loss: 31.291748883503885
l2 norm of gradients: 0.2051218691051787
l2 norm of weights: 5.7555592477021245

Iteration Number: 7030
Loss: 31.289639182650827
l2 norm of gradients: 0.20510252471930446
l2 norm of weights: 5.755485889553865

Iteration Number: 7031
Loss: 31.287529880014656
l2 norm of gradients: 0.20508318496186023
l2 norm of weights: 5.755412537358539

Iteration Number: 7032
Loss: 31.285420975458074
l2 norm of gradients: 0.20506384983071368
l2 norm of weights: 5.755339191115188

Iteration Number: 7033
Loss: 31.283312468849715
l2 norm of gradients: 0.20504451932373374
l2 norm of weights: 5.755265850822854

Iteration Number: 7034
Loss: 31.281204360060542
l2 norm of gradients: 0.20502519343879097
l2 norm of weights: 5.755192516480578

Iteration Number: 7035
Loss: 31.279096648951217
l2 norm of gradients: 0.20500587217375738
l2 norm of weights: 5.755119188087407

Iteration Number: 7036
Loss: 31.276989335395232
l2 norm of gradients: 0.20498655552650624
l2 norm of weights: 5.755045865642381

Iteration Number: 7037
Loss: 31.274882419252478
l2 norm of gradients: 0.2049672434949127
l2 norm of weights: 5.7549725491445445

Iteration Number: 7038
Loss: 31.272775900395033
l2 norm of gradients: 0.204947936076853

```

12 norm of weights: 5.75489923859294
-----
Iteration Number: 7039
Loss: 31.270669778690277
12 norm of gradients: 0.20492863327020513
12 norm of weights: 5.754825933986614
-----
Iteration Number: 7040
Loss: 31.268564054003978
12 norm of gradients: 0.2049093350728484
12 norm of weights: 5.754752635324608
-----
Iteration Number: 7041
Loss: 31.266458726205446
12 norm of gradients: 0.2048900414826636
12 norm of weights: 5.754679342605969
-----
Iteration Number: 7042
Loss: 31.264353795158698
12 norm of gradients: 0.20487075249753303
12 norm of weights: 5.754606055829739
-----
Iteration Number: 7043
Loss: 31.26224926073379
12 norm of gradients: 0.20485146811534044
12 norm of weights: 5.754532774994967
-----
Iteration Number: 7044
Loss: 31.260145122801507
12 norm of gradients: 0.20483218833397102
12 norm of weights: 5.754459500100697
-----
Iteration Number: 7045
Loss: 31.258041381225496
12 norm of gradients: 0.20481291315131148
12 norm of weights: 5.754386231145973
-----
Iteration Number: 7046
Loss: 31.25593803586951
12 norm of gradients: 0.20479364256524993
12 norm of weights: 5.754312968129843
-----
Iteration Number: 7047
Loss: 31.253835086611137
12 norm of gradients: 0.204774376573676
12 norm of weights: 5.754239711051352
-----
Iteration Number: 7048

```

```
Loss: 31.251732533312346
12 norm of gradients: 0.20475511517448067
12 norm of weights: 5.7541664599095474
-----
Iteration Number: 7049
Loss: 31.249630375841853
12 norm of gradients: 0.20473585836555655
12 norm of weights: 5.7540932147034765
-----
Iteration Number: 7050
Loss: 31.247528614065907
12 norm of gradients: 0.20471660614479753
12 norm of weights: 5.754019975432187
-----
Iteration Number: 7051
Loss: 31.245427247858515
12 norm of gradients: 0.2046973585100991
12 norm of weights: 5.753946742094724
-----
Iteration Number: 7052
Loss: 31.24332627708217
12 norm of gradients: 0.20467811545935802
12 norm of weights: 5.753873514690139
-----
Iteration Number: 7053
Loss: 31.241225701606766
12 norm of gradients: 0.20465887699047272
12 norm of weights: 5.753800293217476
-----
Iteration Number: 7054
Loss: 31.239125521303578
12 norm of gradients: 0.20463964310134292
12 norm of weights: 5.753727077675787
-----
Iteration Number: 7055
Loss: 31.2370257360379
12 norm of gradients: 0.20462041378986984
12 norm of weights: 5.753653868064118
-----
Iteration Number: 7056
Loss: 31.234926345678403
12 norm of gradients: 0.20460118905395616
12 norm of weights: 5.75358066438152
-----
Iteration Number: 7057
Loss: 31.232827350098017
12 norm of gradients: 0.20458196889150593
12 norm of weights: 5.753507466627041
```

```
-----
Iteration Number: 7058
Loss: 31.230728749156764
l2 norm of gradients: 0.20456275330042473
l2 norm of weights: 5.753434274799731
-----
Iteration Number: 7059
Loss: 31.228630542729846
l2 norm of gradients: 0.2045435422786196
l2 norm of weights: 5.753361088898639
-----
Iteration Number: 7060
Loss: 31.226532730683598
l2 norm of gradients: 0.2045243358239989
l2 norm of weights: 5.753287908922816
-----
Iteration Number: 7061
Loss: 31.224435312890474
l2 norm of gradients: 0.20450513393447253
l2 norm of weights: 5.753214734871312
-----
Iteration Number: 7062
Loss: 31.222338289213514
l2 norm of gradients: 0.2044859366079518
l2 norm of weights: 5.753141566743179
-----
Iteration Number: 7063
Loss: 31.220241659527403
l2 norm of gradients: 0.20446674384234936
l2 norm of weights: 5.753068404537466
-----
Iteration Number: 7064
Loss: 31.218145423696367
l2 norm of gradients: 0.20444755563557945
l2 norm of weights: 5.752995248253225
-----
Iteration Number: 7065
Loss: 31.216049581592184
l2 norm of gradients: 0.20442837198555774
l2 norm of weights: 5.752922097889509
-----
Iteration Number: 7066
Loss: 31.213954133086215
l2 norm of gradients: 0.2044091928902011
l2 norm of weights: 5.752848953445366
-----
Iteration Number: 7067
Loss: 31.211859078043673
```

```
12 norm of gradients: 0.20439001834742812
12 norm of weights: 5.752775814919853
-----
Iteration Number: 7068
Loss: 31.209764416333506
12 norm of gradients: 0.2043708483551586
12 norm of weights: 5.75270268231202
-----
Iteration Number: 7069
Loss: 31.207670147831763
12 norm of gradients: 0.20435168291131395
12 norm of weights: 5.75262955562092
-----
Iteration Number: 7070
Loss: 31.205576272398275
12 norm of gradients: 0.20433252201381683
12 norm of weights: 5.752556434845604
-----
Iteration Number: 7071
Loss: 31.203482789910616
12 norm of gradients: 0.2043133656605914
12 norm of weights: 5.752483319985129
-----
Iteration Number: 7072
Loss: 31.201389700236746
12 norm of gradients: 0.20429421384956326
12 norm of weights: 5.752410211038545
-----
Iteration Number: 7073
Loss: 31.199297003239696
12 norm of gradients: 0.20427506657865932
12 norm of weights: 5.75233710800491
-----
Iteration Number: 7074
Loss: 31.197204698800817
12 norm of gradients: 0.20425592384580818
12 norm of weights: 5.752264010883273
-----
Iteration Number: 7075
Loss: 31.19511278677999
12 norm of gradients: 0.20423678564893946
12 norm of weights: 5.752190919672692
-----
Iteration Number: 7076
Loss: 31.193021267050607
12 norm of gradients: 0.20421765198598457
12 norm of weights: 5.752117834372223
-----
```


Iteration Number: 7077
Loss: 31.19093013948302
l2 norm of gradients: 0.20419852285487602
l2 norm of weights: 5.752044754980916

Iteration Number: 7078
Loss: 31.18883940394685
l2 norm of gradients: 0.20417939825354806
l2 norm of weights: 5.751971681497831

Iteration Number: 7079
Loss: 31.18674906031431
l2 norm of gradients: 0.20416027817993607
l2 norm of weights: 5.7518986139220205

Iteration Number: 7080
Loss: 31.18465910845198
l2 norm of gradients: 0.20414116263197685
l2 norm of weights: 5.7518255522525426

Iteration Number: 7081
Loss: 31.182569548228578
l2 norm of gradients: 0.20412205160760888
l2 norm of weights: 5.75175249648845

Iteration Number: 7082
Loss: 31.18048037952165
l2 norm of gradients: 0.2041029451047717
l2 norm of weights: 5.7516794466288035

Iteration Number: 7083
Loss: 31.17839160219778
l2 norm of gradients: 0.2040838431214065
l2 norm of weights: 5.751606402672658

Iteration Number: 7084
Loss: 31.176303216126982
l2 norm of gradients: 0.20406474565545576
l2 norm of weights: 5.751533364619069

Iteration Number: 7085
Loss: 31.17421522117534
l2 norm of gradients: 0.2040456527048634
l2 norm of weights: 5.751460332467096

Iteration Number: 7086
Loss: 31.17212761722093
l2 norm of gradients: 0.20402656426757473

```

12 norm of weights: 5.751387306215796
-----
Iteration Number: 7087
Loss: 31.17004040413092
12 norm of gradients: 0.2040074803415365
12 norm of weights: 5.751314285864226
-----
Iteration Number: 7088
Loss: 31.167953581775116
12 norm of gradients: 0.20398840092469667
12 norm of weights: 5.751241271411446
-----
Iteration Number: 7089
Loss: 31.165867150028053
12 norm of gradients: 0.20396932601500492
12 norm of weights: 5.751168262856511
-----
Iteration Number: 7090
Loss: 31.16378110875714
12 norm of gradients: 0.203950255610412
12 norm of weights: 5.751095260198483
-----
Iteration Number: 7091
Loss: 31.161695457835215
12 norm of gradients: 0.20393118970887025
12 norm of weights: 5.751022263436419
-----
Iteration Number: 7092
Loss: 31.15961019712913
12 norm of gradients: 0.20391212830833338
12 norm of weights: 5.750949272569379
-----
Iteration Number: 7093
Loss: 31.157525326519206
12 norm of gradients: 0.20389307140675636
12 norm of weights: 5.750876287596424
-----
Iteration Number: 7094
Loss: 31.155440845867048
12 norm of gradients: 0.20387401900209573
12 norm of weights: 5.750803308516611
-----
Iteration Number: 7095
Loss: 31.15335675504592
12 norm of gradients: 0.20385497109230927
12 norm of weights: 5.750730335329002
-----
Iteration Number: 7096

```

Loss: 31.151273053929156
l2 norm of gradients: 0.20383592767535627
l2 norm of weights: 5.750657368032656

Iteration Number: 7097
Loss: 31.149189742387126
l2 norm of gradients: 0.20381688874919723
l2 norm of weights: 5.750584406626636

Iteration Number: 7098
Loss: 31.147106820292617
l2 norm of gradients: 0.2037978543117942
l2 norm of weights: 5.750511451110001

Iteration Number: 7099
Loss: 31.14502428751694
l2 norm of gradients: 0.20377882436111056
l2 norm of weights: 5.750438501481813

Iteration Number: 7100
Loss: 31.142942143925733
l2 norm of gradients: 0.20375979889511106
l2 norm of weights: 5.750365557741135

Iteration Number: 7101
Loss: 31.140860389398526
l2 norm of gradients: 0.2037407779117617
l2 norm of weights: 5.750292619887024

Iteration Number: 7102
Loss: 31.138779023801067
l2 norm of gradients: 0.2037217614090302
l2 norm of weights: 5.750219687918547

Iteration Number: 7103
Loss: 31.136698047012583
l2 norm of gradients: 0.2037027493848852
l2 norm of weights: 5.750146761834766

Iteration Number: 7104
Loss: 31.134617458899285
l2 norm of gradients: 0.2036837418372971
l2 norm of weights: 5.750073841634741

Iteration Number: 7105
Loss: 31.13253725933028
l2 norm of gradients: 0.2036647387642375
l2 norm of weights: 5.750000927317536

```

-----
Iteration Number: 7106
Loss: 31.130457448181797
l2 norm of gradients: 0.20364574016367934
l2 norm of weights: 5.7499280188822155
-----
Iteration Number: 7107
Loss: 31.128378025328626
l2 norm of gradients: 0.20362674603359696
l2 norm of weights: 5.749855116327842
-----
Iteration Number: 7108
Loss: 31.126298990636208
l2 norm of gradients: 0.20360775637196615
l2 norm of weights: 5.749782219653477
-----
Iteration Number: 7109
Loss: 31.124220343981598
l2 norm of gradients: 0.20358877117676394
l2 norm of weights: 5.74970932885819
-----
Iteration Number: 7110
Loss: 31.12214208523032
l2 norm of gradients: 0.20356979044596882
l2 norm of weights: 5.749636443941041
-----
Iteration Number: 7111
Loss: 31.120064214266435
l2 norm of gradients: 0.20355081417756063
l2 norm of weights: 5.749563564901096
-----
Iteration Number: 7112
Loss: 31.117986730951248
l2 norm of gradients: 0.2035318423695205
l2 norm of weights: 5.749490691737421
-----
Iteration Number: 7113
Loss: 31.115909635160925
l2 norm of gradients: 0.20351287501983092
l2 norm of weights: 5.749417824449078
-----
Iteration Number: 7114
Loss: 31.11383292676708
l2 norm of gradients: 0.20349391212647594
l2 norm of weights: 5.749344963035136
-----
Iteration Number: 7115
Loss: 31.11175660564428

```

```

12 norm of gradients: 0.20347495368744065
12 norm of weights: 5.749272107494657
-----
Iteration Number: 7116
Loss: 31.109680671666784
12 norm of gradients: 0.20345599970071174
12 norm of weights: 5.749199257826712
-----
Iteration Number: 7117
Loss: 31.10760512470314
12 norm of gradients: 0.20343705016427713
12 norm of weights: 5.749126414030363
-----
Iteration Number: 7118
Loss: 31.10552996462915
12 norm of gradients: 0.20341810507612623
12 norm of weights: 5.74905357610468
-----
Iteration Number: 7119
Loss: 31.103455191310772
12 norm of gradients: 0.20339916443424955
12 norm of weights: 5.748980744048726
-----
Iteration Number: 7120
Loss: 31.101380804626857
12 norm of gradients: 0.20338022823663923
12 norm of weights: 5.7489079178615725
-----
Iteration Number: 7121
Loss: 31.099306804454177
12 norm of gradients: 0.20336129648128862
12 norm of weights: 5.748835097542283
-----
Iteration Number: 7122
Loss: 31.097233190657207
12 norm of gradients: 0.2033423691661924
12 norm of weights: 5.748762283089928
-----
Iteration Number: 7123
Loss: 31.09515996311548
12 norm of gradients: 0.20332344628934662
12 norm of weights: 5.748689474503573
-----
Iteration Number: 7124
Loss: 31.093087121696655
12 norm of gradients: 0.2033045278487487
12 norm of weights: 5.748616671782289
-----

```

Iteration Number: 7125
Loss: 31.091014666278127
l2 norm of gradients: 0.20328561384239738
l2 norm of weights: 5.748543874925144

Iteration Number: 7126
Loss: 31.08894259672944
l2 norm of gradients: 0.20326670426829277
l2 norm of weights: 5.748471083931204

Iteration Number: 7127
Loss: 31.08687091292907
l2 norm of gradients: 0.20324779912443622
l2 norm of weights: 5.7483982987995415

Iteration Number: 7128
Loss: 31.084799614748842
l2 norm of gradients: 0.20322889840883054
l2 norm of weights: 5.748325519529224

Iteration Number: 7129
Loss: 31.082728702056073
l2 norm of gradients: 0.2032100021194799
l2 norm of weights: 5.748252746119323

Iteration Number: 7130
Loss: 31.080658174731646
l2 norm of gradients: 0.20319111025438957
l2 norm of weights: 5.748179978568905

Iteration Number: 7131
Loss: 31.078588032647264
l2 norm of gradients: 0.2031722228115665
l2 norm of weights: 5.748107216877045

Iteration Number: 7132
Loss: 31.0765182756733
l2 norm of gradients: 0.2031533397890187
l2 norm of weights: 5.748034461042809

Iteration Number: 7133
Loss: 31.07444890368911
l2 norm of gradients: 0.20313446118475556
l2 norm of weights: 5.747961711065271

Iteration Number: 7134
Loss: 31.072379916560923
l2 norm of gradients: 0.20311558699678797

```

12 norm of weights: 5.7478889669435
-----
Iteration Number: 7135
Loss: 31.070311314171292
12 norm of gradients: 0.20309671722312794
12 norm of weights: 5.7478162286765695
-----
Iteration Number: 7136
Loss: 31.068243096390173
12 norm of gradients: 0.2030778518617889
12 norm of weights: 5.747743496263548
-----
Iteration Number: 7137
Loss: 31.066175263088958
12 norm of gradients: 0.2030589909107856
12 norm of weights: 5.747670769703511
-----
Iteration Number: 7138
Loss: 31.064107814144062
12 norm of gradients: 0.2030401343681341
12 norm of weights: 5.7475980489955285
-----
Iteration Number: 7139
Loss: 31.06204074943152
12 norm of gradients: 0.20302128223185184
12 norm of weights: 5.747525334138672
-----
Iteration Number: 7140
Loss: 31.05997406882064
12 norm of gradients: 0.20300243449995753
12 norm of weights: 5.7474526251320155
-----
Iteration Number: 7141
Loss: 31.057907772188837
12 norm of gradients: 0.20298359117047107
12 norm of weights: 5.747379921974633
-----
Iteration Number: 7142
Loss: 31.05584185941357
12 norm of gradients: 0.20296475224141408
12 norm of weights: 5.747307224665597
-----
Iteration Number: 7143
Loss: 31.053776330359895
12 norm of gradients: 0.202945917710809
12 norm of weights: 5.747234533203981
-----
Iteration Number: 7144

```

Loss: 31.051711184913145
l2 norm of gradients: 0.2029270875766799
l2 norm of weights: 5.747161847588858

Iteration Number: 7145
Loss: 31.049646422939684
l2 norm of gradients: 0.20290826183705207
l2 norm of weights: 5.747089167819303

Iteration Number: 7146
Loss: 31.047582044317462
l2 norm of gradients: 0.20288944048995214
l2 norm of weights: 5.747016493894391

Iteration Number: 7147
Loss: 31.045518048920854
l2 norm of gradients: 0.20287062353340804
l2 norm of weights: 5.7469438258131955

Iteration Number: 7148
Loss: 31.043454436624963
l2 norm of gradients: 0.2028518109654491
l2 norm of weights: 5.7468711635747916

Iteration Number: 7149
Loss: 31.041391207303835
l2 norm of gradients: 0.20283300278410568
l2 norm of weights: 5.746798507178255

Iteration Number: 7150
Loss: 31.039328360831192
l2 norm of gradients: 0.20281419898740982
l2 norm of weights: 5.74672585662266

Iteration Number: 7151
Loss: 31.037265897084964
l2 norm of gradients: 0.20279539957339454
l2 norm of weights: 5.746653211907085

Iteration Number: 7152
Loss: 31.03520381593445
l2 norm of gradients: 0.20277660454009444
l2 norm of weights: 5.746580573030603

Iteration Number: 7153
Loss: 31.033142117262347
l2 norm of gradients: 0.20275781388554523
l2 norm of weights: 5.746507939992291


```

-----
Iteration Number: 7154
Loss: 31.031080800938383
l2 norm of gradients: 0.2027390276077839
l2 norm of weights: 5.746435312791229
-----
Iteration Number: 7155
Loss: 31.02901986684199
l2 norm of gradients: 0.20272024570484898
l2 norm of weights: 5.74636269142649
-----
Iteration Number: 7156
Loss: 31.02695931484071
l2 norm of gradients: 0.20270146817478005
l2 norm of weights: 5.746290075897151
-----
Iteration Number: 7157
Loss: 31.02489914481684
l2 norm of gradients: 0.20268269501561811
l2 norm of weights: 5.746217466202292
-----
Iteration Number: 7158
Loss: 31.022839356644326
l2 norm of gradients: 0.20266392622540547
l2 norm of weights: 5.746144862340989
-----
Iteration Number: 7159
Loss: 31.020779950198204
l2 norm of gradients: 0.20264516180218564
l2 norm of weights: 5.746072264312321
-----
Iteration Number: 7160
Loss: 31.01872092534989
l2 norm of gradients: 0.2026264017440035
l2 norm of weights: 5.7459996721153646
-----
Iteration Number: 7161
Loss: 31.01666228198172
l2 norm of gradients: 0.2026076460489053
l2 norm of weights: 5.7459270857492
-----
Iteration Number: 7162
Loss: 31.01460401996454
l2 norm of gradients: 0.20258889471493832
l2 norm of weights: 5.745854505212905
-----
Iteration Number: 7163
Loss: 31.012546139176553

```

l2 norm of gradients: 0.20257014774015145
l2 norm of weights: 5.74578193050556

Iteration Number: 7164
Loss: 31.01048863949134
l2 norm of gradients: 0.20255140512259462
l2 norm of weights: 5.745709361626242

Iteration Number: 7165
Loss: 31.008431520785575
l2 norm of gradients: 0.2025326668603192
l2 norm of weights: 5.745636798574033

Iteration Number: 7166
Loss: 31.006374782937154
l2 norm of gradients: 0.20251393295137776
l2 norm of weights: 5.745564241348011

Iteration Number: 7167
Loss: 31.004318425820152
l2 norm of gradients: 0.20249520339382415
l2 norm of weights: 5.745491689947258

Iteration Number: 7168
Loss: 31.0022624493078
l2 norm of gradients: 0.20247647818571363
l2 norm of weights: 5.745419144370852

Iteration Number: 7169
Loss: 31.000206853283395
l2 norm of gradients: 0.2024577573251026
l2 norm of weights: 5.745346604617876

Iteration Number: 7170
Loss: 30.99815163761356
l2 norm of gradients: 0.2024390408100488
l2 norm of weights: 5.74527407068741

Iteration Number: 7171
Loss: 30.996096802181007
l2 norm of gradients: 0.2024203286386113
l2 norm of weights: 5.745201542578536

Iteration Number: 7172
Loss: 30.9940423468628
l2 norm of gradients: 0.20240162080885035
l2 norm of weights: 5.745129020290334

Iteration Number: 7173
Loss: 30.991988271528513
l2 norm of gradients: 0.2023829173188274
l2 norm of weights: 5.745056503821886

Iteration Number: 7174
Loss: 30.98993457606341
l2 norm of gradients: 0.20236421816660552
l2 norm of weights: 5.744983993172277

Iteration Number: 7175
Loss: 30.987881260335314
l2 norm of gradients: 0.2023455233502487
l2 norm of weights: 5.744911488340586

Iteration Number: 7176
Loss: 30.985828324228397
l2 norm of gradients: 0.20232683286782235
l2 norm of weights: 5.744838989325896

Iteration Number: 7177
Loss: 30.983775767612826
l2 norm of gradients: 0.2023081467173931
l2 norm of weights: 5.744766496127292

Iteration Number: 7178
Loss: 30.98172359036894
l2 norm of gradients: 0.20228946489702898
l2 norm of weights: 5.744694008743855

Iteration Number: 7179
Loss: 30.979671792370116
l2 norm of gradients: 0.2022707874047991
l2 norm of weights: 5.744621527174671

Iteration Number: 7180
Loss: 30.977620373500496
l2 norm of gradients: 0.20225211423877396
l2 norm of weights: 5.7445490514188196

Iteration Number: 7181
Loss: 30.975569333628798
l2 norm of gradients: 0.20223344539702534
l2 norm of weights: 5.7444765814753875

Iteration Number: 7182
Loss: 30.97351867263289
l2 norm of gradients: 0.20221478087762623

```
12 norm of weights: 5.744404117343459
-----
Iteration Number: 7183
Loss: 30.971468390394776
12 norm of gradients: 0.20219612067865084
12 norm of weights: 5.7443316590221185
-----
Iteration Number: 7184
Loss: 30.969418486786775
12 norm of gradients: 0.2021774647981748
12 norm of weights: 5.744259206510449
-----
Iteration Number: 7185
Loss: 30.967368961685672
12 norm of gradients: 0.20215881323427484
12 norm of weights: 5.74418675980754
-----
Iteration Number: 7186
Loss: 30.96531981497195
12 norm of gradients: 0.20214016598502907
12 norm of weights: 5.744114318912471
-----
Iteration Number: 7187
Loss: 30.96327104652191
12 norm of gradients: 0.2021215230485167
12 norm of weights: 5.744041883824332
-----
Iteration Number: 7188
Loss: 30.961222656212918
12 norm of gradients: 0.2021028844228184
12 norm of weights: 5.743969454542206
-----
Iteration Number: 7189
Loss: 30.95917464391965
12 norm of gradients: 0.2020842501060159
12 norm of weights: 5.743897031065182
-----
Iteration Number: 7190
Loss: 30.95712700952279
12 norm of gradients: 0.2020656200961924
12 norm of weights: 5.743824613392344
-----
Iteration Number: 7191
Loss: 30.955079752895166
12 norm of gradients: 0.20204699439143214
12 norm of weights: 5.74375220152278
-----
Iteration Number: 7192
```

Loss: 30.953032873923373
l2 norm of gradients: 0.2020283729898208
l2 norm of weights: 5.743679795455577

Iteration Number: 7193
Loss: 30.950986372473043
l2 norm of gradients: 0.2020097558894451
l2 norm of weights: 5.743607395189821

Iteration Number: 7194
Loss: 30.94894024843039
l2 norm of gradients: 0.20199114308839322
l2 norm of weights: 5.743535000724601

Iteration Number: 7195
Loss: 30.94689450167008
l2 norm of gradients: 0.20197253458475445
l2 norm of weights: 5.7434626120590035

Iteration Number: 7196
Loss: 30.944849132068782
l2 norm of gradients: 0.20195393037661943
l2 norm of weights: 5.7433902291921175

Iteration Number: 7197
Loss: 30.942804139508137
l2 norm of gradients: 0.20193533046207987
l2 norm of weights: 5.74331785212303

Iteration Number: 7198
Loss: 30.940759523863253
l2 norm of gradients: 0.20191673483922903
l2 norm of weights: 5.743245480850831

Iteration Number: 7199
Loss: 30.938715285013664
l2 norm of gradients: 0.2018981435061611
l2 norm of weights: 5.743173115374609

Iteration Number: 7200
Loss: 30.93667142283371
l2 norm of gradients: 0.20187955646097167
l2 norm of weights: 5.743100755693453

Iteration Number: 7201
Loss: 30.934627937204148
l2 norm of gradients: 0.20186097370175757
l2 norm of weights: 5.743028401806452

```

-----
Iteration Number: 7202
Loss: 30.932584828003215
l2 norm of gradients: 0.2018423952266168
l2 norm of weights: 5.742956053712697
-----
Iteration Number: 7203
Loss: 30.93054209510792
l2 norm of gradients: 0.20182382103364863
l2 norm of weights: 5.742883711411276
-----
Iteration Number: 7204
Loss: 30.928499738399317
l2 norm of gradients: 0.20180525112095365
l2 norm of weights: 5.7428113749012795
-----
Iteration Number: 7205
Loss: 30.9264577577527
l2 norm of gradients: 0.20178668548663356
l2 norm of weights: 5.7427390441818
-----
Iteration Number: 7206
Loss: 30.92441615304439
l2 norm of gradients: 0.20176812412879136
l2 norm of weights: 5.742666719251925
-----
Iteration Number: 7207
Loss: 30.92237492415976
l2 norm of gradients: 0.2017495670455313
l2 norm of weights: 5.742594400110748
-----
Iteration Number: 7208
Loss: 30.920334070972284
l2 norm of gradients: 0.2017310142349588
l2 norm of weights: 5.74252208675736
-----
Iteration Number: 7209
Loss: 30.918293593366403
l2 norm of gradients: 0.20171246569518053
l2 norm of weights: 5.742449779190851
-----
Iteration Number: 7210
Loss: 30.916253491210373
l2 norm of gradients: 0.20169392142430445
l2 norm of weights: 5.742377477410314
-----
Iteration Number: 7211
Loss: 30.91421376438866

```

```

12 norm of gradients: 0.2016753814204397
12 norm of weights: 5.742305181414842
-----
Iteration Number: 7212
Loss: 30.912174412782687
12 norm of gradients: 0.20165684568169667
12 norm of weights: 5.742232891203526
-----
Iteration Number: 7213
Loss: 30.910135436266835
12 norm of gradients: 0.20163831420618683
12 norm of weights: 5.742160606775458
-----
Iteration Number: 7214
Loss: 30.908096834721974
12 norm of gradients: 0.20161978699202318
12 norm of weights: 5.742088328129732
-----
Iteration Number: 7215
Loss: 30.906058608023436
12 norm of gradients: 0.20160126403731968
12 norm of weights: 5.742016055265442
-----
Iteration Number: 7216
Loss: 30.904020756057268
12 norm of gradients: 0.20158274534019152
12 norm of weights: 5.741943788181678
-----
Iteration Number: 7217
Loss: 30.901983278698147
12 norm of gradients: 0.2015642308987554
12 norm of weights: 5.741871526877538
-----
Iteration Number: 7218
Loss: 30.899946175826617
12 norm of gradients: 0.2015457207111288
12 norm of weights: 5.741799271352113
-----
Iteration Number: 7219
Loss: 30.897909447316376
12 norm of gradients: 0.20152721477543084
12 norm of weights: 5.741727021604498
-----
Iteration Number: 7220
Loss: 30.895873093055798
12 norm of gradients: 0.2015087130897815
12 norm of weights: 5.741654777633787
-----

```

Iteration Number: 7221
Loss: 30.893837112919474
l2 norm of gradients: 0.20149021565230224
l2 norm of weights: 5.741582539439077

Iteration Number: 7222
Loss: 30.891801506786493
l2 norm of gradients: 0.20147172246111564
l2 norm of weights: 5.7415103070194595

Iteration Number: 7223
Loss: 30.889766274533777
l2 norm of gradients: 0.20145323351434544
l2 norm of weights: 5.741438080374032

Iteration Number: 7224
Loss: 30.887731416046584
l2 norm of gradients: 0.2014347488101166
l2 norm of weights: 5.74136585950189

Iteration Number: 7225
Loss: 30.88569693120108
l2 norm of gradients: 0.2014162683465555
l2 norm of weights: 5.741293644402129

Iteration Number: 7226
Loss: 30.88366281987745
l2 norm of gradients: 0.20139779212178938
l2 norm of weights: 5.741221435073845

Iteration Number: 7227
Loss: 30.881629081954685
l2 norm of gradients: 0.20137932013394702
l2 norm of weights: 5.741149231516133

Iteration Number: 7228
Loss: 30.879595717314494
l2 norm of gradients: 0.20136085238115817
l2 norm of weights: 5.741077033728092

Iteration Number: 7229
Loss: 30.877562725835208
l2 norm of gradients: 0.20134238886155387
l2 norm of weights: 5.7410048417088175

Iteration Number: 7230
Loss: 30.8755301073983
l2 norm of gradients: 0.20132392957326645


```

12 norm of weights: 5.740932655457407
-----
Iteration Number: 7231
Loss: 30.873497861880693
12 norm of gradients: 0.20130547451442926
12 norm of weights: 5.7408604749729575
-----
Iteration Number: 7232
Loss: 30.871465989161166
12 norm of gradients: 0.20128702368317705
12 norm of weights: 5.740788300254567
-----
Iteration Number: 7233
Loss: 30.869434489124085
12 norm of gradients: 0.20126857707764564
12 norm of weights: 5.740716131301333
-----
Iteration Number: 7234
Loss: 30.86740336164803
12 norm of gradients: 0.20125013469597208
12 norm of weights: 5.740643968112355
-----
Iteration Number: 7235
Loss: 30.86537260661517
12 norm of gradients: 0.2012316965362947
12 norm of weights: 5.74057181068673
-----
Iteration Number: 7236
Loss: 30.863342223899764
12 norm of gradients: 0.2012132625967529
12 norm of weights: 5.740499659023557
-----
Iteration Number: 7237
Loss: 30.86131221338856
12 norm of gradients: 0.20119483287548728
12 norm of weights: 5.7404275131219356
-----
Iteration Number: 7238
Loss: 30.859282574958538
12 norm of gradients: 0.20117640737063974
12 norm of weights: 5.740355372980964
-----
Iteration Number: 7239
Loss: 30.857253308493036
12 norm of gradients: 0.20115798608035337
12 norm of weights: 5.740283238599743
-----
Iteration Number: 7240

```

Loss: 30.855224413868534
l2 norm of gradients: 0.20113956900277236
l2 norm of weights: 5.740211109977372

Iteration Number: 7241
Loss: 30.853195890968603
l2 norm of gradients: 0.2011211561360422
l2 norm of weights: 5.7401389871129505

Iteration Number: 7242
Loss: 30.851167739671318
l2 norm of gradients: 0.20110274747830945
l2 norm of weights: 5.740066870005578

Iteration Number: 7243
Loss: 30.849139959859116
l2 norm of gradients: 0.20108434302772188
l2 norm of weights: 5.739994758654357

Iteration Number: 7244
Loss: 30.847112551412305
l2 norm of gradients: 0.20106594278242862
l2 norm of weights: 5.739922653058389

Iteration Number: 7245
Loss: 30.845085514210556
l2 norm of gradients: 0.2010475467405797
l2 norm of weights: 5.739850553216771

Iteration Number: 7246
Loss: 30.843058848134174
l2 norm of gradients: 0.2010291549003266
l2 norm of weights: 5.739778459128608

Iteration Number: 7247
Loss: 30.84103255306889
l2 norm of gradients: 0.20101076725982184
l2 norm of weights: 5.739706370793

Iteration Number: 7248
Loss: 30.839006628890715
l2 norm of gradients: 0.20099238381721918
l2 norm of weights: 5.73963428820905

Iteration Number: 7249
Loss: 30.836981075480338
l2 norm of gradients: 0.20097400457067355
l2 norm of weights: 5.739562211375858

```
-----
Iteration Number: 7250
Loss: 30.834955892726295
l2 norm of gradients: 0.200955629518341
l2 norm of weights: 5.739490140292531
-----
Iteration Number: 7251
Loss: 30.832931080499094
l2 norm of gradients: 0.2009372586583789
l2 norm of weights: 5.7394180749581665
-----
Iteration Number: 7252
Loss: 30.830906638687708
l2 norm of gradients: 0.2009188919889456
l2 norm of weights: 5.739346015371869
-----
Iteration Number: 7253
Loss: 30.828882567170627
l2 norm of gradients: 0.20090052950820086
l2 norm of weights: 5.739273961532744
-----
Iteration Number: 7254
Loss: 30.826858865828356
l2 norm of gradients: 0.2008821712143054
l2 norm of weights: 5.739201913439894
-----
Iteration Number: 7255
Loss: 30.82483553454265
l2 norm of gradients: 0.20086381710542134
l2 norm of weights: 5.73912987109242
-----
Iteration Number: 7256
Loss: 30.82281257319385
l2 norm of gradients: 0.20084546717971172
l2 norm of weights: 5.7390578344894285
-----
Iteration Number: 7257
Loss: 30.820789981669783
l2 norm of gradients: 0.20082712143534093
l2 norm of weights: 5.7389858036300225
-----
Iteration Number: 7258
Loss: 30.818767759844132
l2 norm of gradients: 0.2008087798704745
l2 norm of weights: 5.738913778513308
-----
Iteration Number: 7259
Loss: 30.816745907601703
```

l2 norm of gradients: 0.20079044248327912
l2 norm of weights: 5.7388417591383885

Iteration Number: 7260
Loss: 30.814724424824178
l2 norm of gradients: 0.20077210927192266
l2 norm of weights: 5.73876974550437

Iteration Number: 7261
Loss: 30.812703311395232
l2 norm of gradients: 0.20075378023457405
l2 norm of weights: 5.738697737610357

Iteration Number: 7262
Loss: 30.810682567191595
l2 norm of gradients: 0.2007354553694036
l2 norm of weights: 5.7386257354554555

Iteration Number: 7263
Loss: 30.808662192100634
l2 norm of gradients: 0.20071713467458255
l2 norm of weights: 5.738553739038772

Iteration Number: 7264
Loss: 30.806642185999138
l2 norm of gradients: 0.20069881814828353
l2 norm of weights: 5.73848174835941

Iteration Number: 7265
Loss: 30.80462254877275
l2 norm of gradients: 0.2006805057886802
l2 norm of weights: 5.738409763416478

Iteration Number: 7266
Loss: 30.802603280304034
l2 norm of gradients: 0.20066219759394735
l2 norm of weights: 5.738337784209083

Iteration Number: 7267
Loss: 30.800584380473374
l2 norm of gradients: 0.20064389356226106
l2 norm of weights: 5.738265810736331

Iteration Number: 7268
Loss: 30.798565849161477
l2 norm of gradients: 0.20062559369179847
l2 norm of weights: 5.738193842997328

Iteration Number: 7269
Loss: 30.796547686253426
l2 norm of gradients: 0.2006072979807379
l2 norm of weights: 5.738121880991184

Iteration Number: 7270
Loss: 30.794529891628045
l2 norm of gradients: 0.20058900642725894
l2 norm of weights: 5.738049924717004

Iteration Number: 7271
Loss: 30.79251246517129
l2 norm of gradients: 0.20057071902954207
l2 norm of weights: 5.737977974173897

Iteration Number: 7272
Loss: 30.79049540676275
l2 norm of gradients: 0.20055243578576923
l2 norm of weights: 5.737906029360971

Iteration Number: 7273
Loss: 30.788478716288964
l2 norm of gradients: 0.2005341566941233
l2 norm of weights: 5.737834090277335

Iteration Number: 7274
Loss: 30.786462393627726
l2 norm of gradients: 0.20051588175278845
l2 norm of weights: 5.737762156922098

Iteration Number: 7275
Loss: 30.78444643866363
l2 norm of gradients: 0.20049761095994992
l2 norm of weights: 5.737690229294367

Iteration Number: 7276
Loss: 30.782430851279784
l2 norm of gradients: 0.2004793443137941
l2 norm of weights: 5.737618307393253

Iteration Number: 7277
Loss: 30.780415631359674
l2 norm of gradients: 0.20046108181250857
l2 norm of weights: 5.737546391217865

Iteration Number: 7278
Loss: 30.778400778783038
l2 norm of gradients: 0.20044282345428205

```
l2 norm of weights: 5.737474480767313
-----
Iteration Number: 7279
Loss: 30.77638629343188
l2 norm of gradients: 0.20042456923730448
l2 norm of weights: 5.737402576040705
-----
Iteration Number: 7280
Loss: 30.774372175195488
l2 norm of gradients: 0.20040631915976678
l2 norm of weights: 5.737330677037154
-----
Iteration Number: 7281
Loss: 30.77235842395283
l2 norm of gradients: 0.2003880732198611
l2 norm of weights: 5.737258783755769
-----
Iteration Number: 7282
Loss: 30.770345039582594
l2 norm of gradients: 0.2003698314157808
l2 norm of weights: 5.737186896195662
-----
Iteration Number: 7283
Loss: 30.768332021976285
l2 norm of gradients: 0.20035159374572037
l2 norm of weights: 5.737115014355942
-----
Iteration Number: 7284
Loss: 30.76631937100881
l2 norm of gradients: 0.20033336020787526
l2 norm of weights: 5.737043138235721
-----
Iteration Number: 7285
Loss: 30.764307086573307
l2 norm of gradients: 0.2003151308004423
l2 norm of weights: 5.736971267834113
-----
Iteration Number: 7286
Loss: 30.762295168543005
l2 norm of gradients: 0.20029690552161936
l2 norm of weights: 5.736899403150225
-----
Iteration Number: 7287
Loss: 30.760283616805477
l2 norm of gradients: 0.20027868436960536
l2 norm of weights: 5.736827544183174
-----
Iteration Number: 7288
```

Loss: 30.75827243124281
l2 norm of gradients: 0.2002604673426006
l2 norm of weights: 5.736755690932068

Iteration Number: 7289
Loss: 30.75626161174433
l2 norm of gradients: 0.20024225443880628
l2 norm of weights: 5.7366838433960226

Iteration Number: 7290
Loss: 30.754251158182417
l2 norm of gradients: 0.20022404565642477
l2 norm of weights: 5.73661200157415

Iteration Number: 7291
Loss: 30.75224107045126
l2 norm of gradients: 0.20020584099365973
l2 norm of weights: 5.736540165465562

Iteration Number: 7292
Loss: 30.750231348426915
l2 norm of gradients: 0.2001876404487158
l2 norm of weights: 5.736468335069373

Iteration Number: 7293
Loss: 30.748221991994683
l2 norm of gradients: 0.20016944401979883
l2 norm of weights: 5.7363965103846954

Iteration Number: 7294
Loss: 30.746213001043454
l2 norm of gradients: 0.20015125170511575
l2 norm of weights: 5.736324691410645

Iteration Number: 7295
Loss: 30.744204375449886
l2 norm of gradients: 0.20013306350287463
l2 norm of weights: 5.7362528781463356

Iteration Number: 7296
Loss: 30.742196115102345
l2 norm of gradients: 0.20011487941128467
l2 norm of weights: 5.736181070590879

Iteration Number: 7297
Loss: 30.740188219882736
l2 norm of gradients: 0.2000966994285563
l2 norm of weights: 5.736109268743394

```
-----
Iteration Number: 7298
Loss: 30.738180689673086
l2 norm of gradients: 0.20007852355290096
l2 norm of weights: 5.736037472602992
-----
Iteration Number: 7299
Loss: 30.73617352436403
l2 norm of gradients: 0.2000603517825312
l2 norm of weights: 5.735965682168788
-----
Iteration Number: 7300
Loss: 30.734166723834964
l2 norm of gradients: 0.2000421841156608
l2 norm of weights: 5.735893897439901
-----
Iteration Number: 7301
Loss: 30.732160287968497
l2 norm of gradients: 0.20002402055050456
l2 norm of weights: 5.7358221184154425
-----
Iteration Number: 7302
Loss: 30.730154216649574
l2 norm of gradients: 0.20000586108527846
l2 norm of weights: 5.735750345094531
-----
Iteration Number: 7303
Loss: 30.7281485097677
l2 norm of gradients: 0.19998770571819965
l2 norm of weights: 5.735678577476282
-----
Iteration Number: 7304
Loss: 30.72614316720023
l2 norm of gradients: 0.19996955444748624
l2 norm of weights: 5.7356068155598114
-----
Iteration Number: 7305
Loss: 30.724138188833177
l2 norm of gradients: 0.19995140727135768
l2 norm of weights: 5.735535059344237
-----
Iteration Number: 7306
Loss: 30.722133574555695
l2 norm of gradients: 0.1999332641880343
l2 norm of weights: 5.735463308828675
-----
Iteration Number: 7307
Loss: 30.720129324243917
```



```

12 norm of gradients: 0.1999151251957378
12 norm of weights: 5.735391564012242
-----
Iteration Number: 7308
Loss: 30.718125437787503
12 norm of gradients: 0.1998969902926907
12 norm of weights: 5.735319824894056
-----
Iteration Number: 7309
Loss: 30.716121915072094
12 norm of gradients: 0.19987885947711698
12 norm of weights: 5.7352480914732356
-----
Iteration Number: 7310
Loss: 30.71411875598037
12 norm of gradients: 0.19986073274724145
12 norm of weights: 5.735176363748897
-----
Iteration Number: 7311
Loss: 30.712115960397945
12 norm of gradients: 0.19984261010129017
12 norm of weights: 5.735104641720159
-----
Iteration Number: 7312
Loss: 30.71011352820887
12 norm of gradients: 0.19982449153749024
12 norm of weights: 5.7350329253861405
-----
Iteration Number: 7313
Loss: 30.7081114592965
12 norm of gradients: 0.19980637705406998
12 norm of weights: 5.73496121474596
-----
Iteration Number: 7314
Loss: 30.70610975354912
12 norm of gradients: 0.19978826664925872
12 norm of weights: 5.734889509798738
-----
Iteration Number: 7315
Loss: 30.70410841084856
12 norm of gradients: 0.19977016032128686
12 norm of weights: 5.73481781054359
-----
Iteration Number: 7316
Loss: 30.702107431080563
12 norm of gradients: 0.1997520580683861
12 norm of weights: 5.73474611697964
-----

```

Iteration Number: 7317
Loss: 30.700106814131257
l2 norm of gradients: 0.19973395988878911
l2 norm of weights: 5.734674429106003

Iteration Number: 7318
Loss: 30.69810655988465
l2 norm of gradients: 0.19971586578072958
l2 norm of weights: 5.734602746921802

Iteration Number: 7319
Loss: 30.696106668227227
l2 norm of gradients: 0.19969777574244252
l2 norm of weights: 5.734531070426158

Iteration Number: 7320
Loss: 30.694107139041915
l2 norm of gradients: 0.19967968977216385
l2 norm of weights: 5.734459399618189

Iteration Number: 7321
Loss: 30.692107972214412
l2 norm of gradients: 0.19966160786813075
l2 norm of weights: 5.734387734497017

Iteration Number: 7322
Loss: 30.690109167629476
l2 norm of gradients: 0.19964353002858135
l2 norm of weights: 5.734316075061764

Iteration Number: 7323
Loss: 30.688110725178547
l2 norm of gradients: 0.199625456251755
l2 norm of weights: 5.7342444213115495

Iteration Number: 7324
Loss: 30.686112644736518
l2 norm of gradients: 0.1996073865358921
l2 norm of weights: 5.734172773245494

Iteration Number: 7325
Loss: 30.68411492620018
l2 norm of gradients: 0.19958932087923423
l2 norm of weights: 5.734101130862723

Iteration Number: 7326
Loss: 30.682117569446845
l2 norm of gradients: 0.1995712592800238

```

12 norm of weights: 5.734029494162357
-----
Iteration Number: 7327
Loss: 30.68012057436486
12 norm of gradients: 0.1995532017365047
12 norm of weights: 5.733957863143516
-----
Iteration Number: 7328
Loss: 30.678123940838603
12 norm of gradients: 0.19953514824692167
12 norm of weights: 5.733886237805325
-----
Iteration Number: 7329
Loss: 30.676127668759072
12 norm of gradients: 0.19951709880952057
12 norm of weights: 5.733814618146906
-----
Iteration Number: 7330
Loss: 30.674131758003295
12 norm of gradients: 0.19949905342254834
12 norm of weights: 5.733743004167381
-----
Iteration Number: 7331
Loss: 30.672136208465062
12 norm of gradients: 0.19948101208425317
12 norm of weights: 5.733671395865875
-----
Iteration Number: 7332
Loss: 30.670141020023674
12 norm of gradients: 0.1994629747928841
12 norm of weights: 5.7335997932415115
-----
Iteration Number: 7333
Loss: 30.66814619257094
12 norm of gradients: 0.1994449415466915
12 norm of weights: 5.733528196293412
-----
Iteration Number: 7334
Loss: 30.666151725989284
12 norm of gradients: 0.19942691234392673
12 norm of weights: 5.733456605020704
-----
Iteration Number: 7335
Loss: 30.66415762016637
12 norm of gradients: 0.1994088871828421
12 norm of weights: 5.733385019422509
-----
Iteration Number: 7336

```

Loss: 30.662163874987503
l2 norm of gradients: 0.19939086606169124
l2 norm of weights: 5.7333134394979535

Iteration Number: 7337
Loss: 30.660170490338373
l2 norm of gradients: 0.19937284897872867
l2 norm of weights: 5.7332418652461605

Iteration Number: 7338
Loss: 30.65817746610618
l2 norm of gradients: 0.19935483593221018
l2 norm of weights: 5.733170296666255

Iteration Number: 7339
Loss: 30.656184802178934
l2 norm of gradients: 0.19933682692039245
l2 norm of weights: 5.733098733757364

Iteration Number: 7340
Loss: 30.654192498437077
l2 norm of gradients: 0.19931882194153344
l2 norm of weights: 5.733027176518614

Iteration Number: 7341
Loss: 30.652200554774456
l2 norm of gradients: 0.19930082099389204
l2 norm of weights: 5.732955624949127

Iteration Number: 7342
Loss: 30.650208971070356
l2 norm of gradients: 0.1992828240757283
l2 norm of weights: 5.7328840790480315

Iteration Number: 7343
Loss: 30.648217747216968
l2 norm of gradients: 0.1992648311853033
l2 norm of weights: 5.7328125388144535

Iteration Number: 7344
Loss: 30.646226883098606
l2 norm of gradients: 0.19924684232087925
l2 norm of weights: 5.732741004247519

Iteration Number: 7345
Loss: 30.644236378601367
l2 norm of gradients: 0.19922885748071936
l2 norm of weights: 5.732669475346356

```

-----
Iteration Number: 7346
Loss: 30.642246233611253
l2 norm of gradients: 0.199210876663088
l2 norm of weights: 5.73259795211009
-----
Iteration Number: 7347
Loss: 30.64025644802036
l2 norm of gradients: 0.1991928998662506
l2 norm of weights: 5.732526434537849
-----
Iteration Number: 7348
Loss: 30.638267021708202
l2 norm of gradients: 0.19917492708847365
l2 norm of weights: 5.7324549226287616
-----
Iteration Number: 7349
Loss: 30.63627795456611
l2 norm of gradients: 0.19915695832802474
l2 norm of weights: 5.732383416381954
-----
Iteration Number: 7350
Loss: 30.63428924647976
l2 norm of gradients: 0.19913899358317247
l2 norm of weights: 5.732311915796555
-----
Iteration Number: 7351
Loss: 30.632300897334208
l2 norm of gradients: 0.19912103285218652
l2 norm of weights: 5.732240420871693
-----
Iteration Number: 7352
Loss: 30.63031290702035
l2 norm of gradients: 0.19910307613333775
l2 norm of weights: 5.7321689316064965
-----
Iteration Number: 7353
Loss: 30.62832527542458
l2 norm of gradients: 0.19908512342489798
l2 norm of weights: 5.732097448000094
-----
Iteration Number: 7354
Loss: 30.626338002428373
l2 norm of gradients: 0.1990671747251401
l2 norm of weights: 5.732025970051615
-----
Iteration Number: 7355
Loss: 30.6243510879274

```

```

12 norm of gradients: 0.19904923003233815
12 norm of weights: 5.731954497760189
-----
Iteration Number: 7356
Loss: 30.622364531803576
12 norm of gradients: 0.1990312893447672
12 norm of weights: 5.731883031124945
-----
Iteration Number: 7357
Loss: 30.62037833394365
12 norm of gradients: 0.1990133526607033
12 norm of weights: 5.7318115701450125
-----
Iteration Number: 7358
Loss: 30.618392494239522
12 norm of gradients: 0.19899541997842368
12 norm of weights: 5.731740114819522
-----
Iteration Number: 7359
Loss: 30.61640701257173
12 norm of gradients: 0.19897749129620662
12 norm of weights: 5.7316686651476045
-----
Iteration Number: 7360
Loss: 30.61442188883491
12 norm of gradients: 0.19895956661233138
12 norm of weights: 5.73159722112839
-----
Iteration Number: 7361
Loss: 30.612437122913644
12 norm of gradients: 0.1989416459250784
12 norm of weights: 5.731525782761008
-----
Iteration Number: 7362
Loss: 30.610452714694112
12 norm of gradients: 0.19892372923272905
12 norm of weights: 5.731454350044593
-----
Iteration Number: 7363
Loss: 30.608468664066333
12 norm of gradients: 0.19890581653356595
12 norm of weights: 5.731382922978272
-----
Iteration Number: 7364
Loss: 30.606484970912526
12 norm of gradients: 0.1988879078258725
12 norm of weights: 5.731311501561179
-----

```

Iteration Number: 7365
Loss: 30.60450163512955
l2 norm of gradients: 0.19887000310793346
l2 norm of weights: 5.731240085792447

Iteration Number: 7366
Loss: 30.602518656598555
l2 norm of gradients: 0.1988521023780344
l2 norm of weights: 5.731168675671205

Iteration Number: 7367
Loss: 30.600536035209352
l2 norm of gradients: 0.19883420563446214
l2 norm of weights: 5.731097271196587

Iteration Number: 7368
Loss: 30.59855377084897
l2 norm of gradients: 0.19881631287550444
l2 norm of weights: 5.731025872367725

Iteration Number: 7369
Loss: 30.596571863408283
l2 norm of gradients: 0.19879842409945014
l2 norm of weights: 5.730954479183753

Iteration Number: 7370
Loss: 30.594590312768904
l2 norm of gradients: 0.19878053930458903
l2 norm of weights: 5.730883091643803

Iteration Number: 7371
Loss: 30.592609118827546
l2 norm of gradients: 0.1987626584892123
l2 norm of weights: 5.730811709747009

Iteration Number: 7372
Loss: 30.590628281467584
l2 norm of gradients: 0.19874478165161175
l2 norm of weights: 5.730740333492504

Iteration Number: 7373
Loss: 30.588647800573455
l2 norm of gradients: 0.19872690879008048
l2 norm of weights: 5.730668962879421

Iteration Number: 7374
Loss: 30.58666767604385
l2 norm of gradients: 0.1987090399029126

```
l2 norm of weights: 5.730597597906896
-----
Iteration Number: 7375
Loss: 30.58468790775631
l2 norm of gradients: 0.19869117498840327
l2 norm of weights: 5.730526238574061
-----
Iteration Number: 7376
Loss: 30.58270849560335
l2 norm of gradients: 0.19867331404484867
l2 norm of weights: 5.730454884880053
-----
Iteration Number: 7377
Loss: 30.58072943947596
l2 norm of gradients: 0.1986554570705461
l2 norm of weights: 5.730383536824004
-----
Iteration Number: 7378
Loss: 30.57875073925749
l2 norm of gradients: 0.19863760406379383
l2 norm of weights: 5.730312194405051
-----
Iteration Number: 7379
Loss: 30.576772394841438
l2 norm of gradients: 0.19861975502289114
l2 norm of weights: 5.730240857622329
-----
Iteration Number: 7380
Loss: 30.574794406114247
l2 norm of gradients: 0.19860190994613844
l2 norm of weights: 5.730169526474973
-----
Iteration Number: 7381
Loss: 30.572816772962913
l2 norm of gradients: 0.19858406883183716
l2 norm of weights: 5.73009820096212
-----
Iteration Number: 7382
Loss: 30.57083949527753
l2 norm of gradients: 0.19856623167828974
l2 norm of weights: 5.730026881082905
-----
Iteration Number: 7383
Loss: 30.568862572948866
l2 norm of gradients: 0.19854839848379977
l2 norm of weights: 5.729955566836464
-----
Iteration Number: 7384
```


Loss: 30.566886005862678
l2 norm of gradients: 0.1985305692466717
l2 norm of weights: 5.729884258221934

Iteration Number: 7385
Loss: 30.56490979390863
l2 norm of gradients: 0.19851274396521118
l2 norm of weights: 5.729812955238451

Iteration Number: 7386
Loss: 30.56293393697775
l2 norm of gradients: 0.19849492263772478
l2 norm of weights: 5.729741657885153

Iteration Number: 7387
Loss: 30.5609584349548
l2 norm of gradients: 0.19847710526252021
l2 norm of weights: 5.729670366161178

Iteration Number: 7388
Loss: 30.558983287734577
l2 norm of gradients: 0.19845929183790614
l2 norm of weights: 5.72959908006566

Iteration Number: 7389
Loss: 30.557008495198993
l2 norm of gradients: 0.19844148236219233
l2 norm of weights: 5.729527799597741

Iteration Number: 7390
Loss: 30.55503405724369
l2 norm of gradients: 0.19842367683368956
l2 norm of weights: 5.729456524756557

Iteration Number: 7391
Loss: 30.553059973752173
l2 norm of gradients: 0.19840587525070955
l2 norm of weights: 5.729385255541247

Iteration Number: 7392
Loss: 30.55108624461862
l2 norm of gradients: 0.1983880776115652
l2 norm of weights: 5.7293139919509475

Iteration Number: 7393
Loss: 30.549112869732593
l2 norm of gradients: 0.19837028391457043
l2 norm of weights: 5.729242733984799

```

-----
Iteration Number: 7394
Loss: 30.547139848976926
l2 norm of gradients: 0.19835249415804002
l2 norm of weights: 5.729171481641941
-----
Iteration Number: 7395
Loss: 30.545167182249756
l2 norm of gradients: 0.19833470834029
l2 norm of weights: 5.72910023492151
-----
Iteration Number: 7396
Loss: 30.54319486943005
l2 norm of gradients: 0.1983169264596372
l2 norm of weights: 5.729028993822648
-----
Iteration Number: 7397
Loss: 30.541222910418085
l2 norm of gradients: 0.19829914851439978
l2 norm of weights: 5.728957758344494
-----
Iteration Number: 7398
Loss: 30.53925130509752
l2 norm of gradients: 0.1982813745028966
l2 norm of weights: 5.728886528486187
-----
Iteration Number: 7399
Loss: 30.53728005336146
l2 norm of gradients: 0.1982636044234478
l2 norm of weights: 5.7288153042468695
-----
Iteration Number: 7400
Loss: 30.535309155092328
l2 norm of gradients: 0.19824583827437436
l2 norm of weights: 5.728744085625681
-----
Iteration Number: 7401
Loss: 30.533338610187215
l2 norm of gradients: 0.19822807605399842
l2 norm of weights: 5.728672872621759
-----
Iteration Number: 7402
Loss: 30.531368418534196
l2 norm of gradients: 0.19821031776064307
l2 norm of weights: 5.728601665234249
-----
Iteration Number: 7403
Loss: 30.529398580021784

```

```

12 norm of gradients: 0.1981925633926325
12 norm of weights: 5.72853046346229
-----
Iteration Number: 7404
Loss: 30.527429094539897
12 norm of gradients: 0.19817481294829184
12 norm of weights: 5.728459267305023
-----
Iteration Number: 7405
Loss: 30.525459961981102
12 norm of gradients: 0.19815706642594727
12 norm of weights: 5.728388076761591
-----
Iteration Number: 7406
Loss: 30.523491182228582
12 norm of gradients: 0.19813932382392588
12 norm of weights: 5.728316891831135
-----
Iteration Number: 7407
Loss: 30.521522755180925
12 norm of gradients: 0.19812158514055606
12 norm of weights: 5.7282457125127975
-----
Iteration Number: 7408
Loss: 30.519554680724106
12 norm of gradients: 0.19810385037416695
12 norm of weights: 5.728174538805721
-----
Iteration Number: 7409
Loss: 30.517586958749625
12 norm of gradients: 0.19808611952308877
12 norm of weights: 5.728103370709048
-----
Iteration Number: 7410
Loss: 30.515619589144666
12 norm of gradients: 0.19806839258565287
12 norm of weights: 5.728032208221921
-----
Iteration Number: 7411
Loss: 30.513652571802886
12 norm of gradients: 0.1980506695601915
12 norm of weights: 5.727961051343484
-----
Iteration Number: 7412
Loss: 30.511685906609966
12 norm of gradients: 0.19803295044503796
12 norm of weights: 5.727889900072879
-----

```

Iteration Number: 7413
Loss: 30.509719593461877
l2 norm of gradients: 0.19801523523852654
l2 norm of weights: 5.727818754409251

Iteration Number: 7414
Loss: 30.507753632249024
l2 norm of gradients: 0.19799752393899256
l2 norm of weights: 5.727747614351743

Iteration Number: 7415
Loss: 30.505788022855924
l2 norm of gradients: 0.1979798165447724
l2 norm of weights: 5.727676479899499

Iteration Number: 7416
Loss: 30.503822765178903
l2 norm of gradients: 0.1979621130542034
l2 norm of weights: 5.727605351051664

Iteration Number: 7417
Loss: 30.501857859104863
l2 norm of gradients: 0.1979444134656239
l2 norm of weights: 5.727534227807382

Iteration Number: 7418
Loss: 30.499893304526747
l2 norm of gradients: 0.19792671777737322
l2 norm of weights: 5.7274631101657985

Iteration Number: 7419
Loss: 30.49792910133726
l2 norm of gradients: 0.19790902598779178
l2 norm of weights: 5.727391998126059

Iteration Number: 7420
Loss: 30.495965249418788
l2 norm of gradients: 0.19789133809522105
l2 norm of weights: 5.727320891687306

Iteration Number: 7421
Loss: 30.494001748669703
l2 norm of gradients: 0.19787365409800325
l2 norm of weights: 5.727249790848689

Iteration Number: 7422
Loss: 30.49203859898127
l2 norm of gradients: 0.19785597399448185

```

12 norm of weights: 5.727178695609351
-----
Iteration Number: 7423
Loss: 30.490075800240827
12 norm of gradients: 0.19783829778300124
12 norm of weights: 5.727107605968439
-----
Iteration Number: 7424
Loss: 30.48811335234049
12 norm of gradients: 0.1978206254619069
12 norm of weights: 5.727036521925099
-----
Iteration Number: 7425
Loss: 30.486151255172608
12 norm of gradients: 0.19780295702954515
12 norm of weights: 5.726965443478478
-----
Iteration Number: 7426
Loss: 30.484189508625413
12 norm of gradients: 0.19778529248426338
12 norm of weights: 5.726894370627722
-----
Iteration Number: 7427
Loss: 30.48222811259259
12 norm of gradients: 0.19776763182441004
12 norm of weights: 5.726823303371978
-----
Iteration Number: 7428
Loss: 30.480267066963865
12 norm of gradients: 0.19774997504833458
12 norm of weights: 5.726752241710394
-----
Iteration Number: 7429
Loss: 30.478306371631593
12 norm of gradients: 0.1977323221543873
12 norm of weights: 5.726681185642117
-----
Iteration Number: 7430
Loss: 30.47634602648495
12 norm of gradients: 0.1977146731409197
12 norm of weights: 5.726610135166295
-----
Iteration Number: 7431
Loss: 30.47438603141879
12 norm of gradients: 0.19769702800628414
12 norm of weights: 5.726539090282076
-----
Iteration Number: 7432

```

Loss: 30.47242638632052
l2 norm of gradients: 0.19767938674883404
l2 norm of weights: 5.726468050988609

Iteration Number: 7433
Loss: 30.47046709108455
l2 norm of gradients: 0.1976617493669238
l2 norm of weights: 5.72639701728504

Iteration Number: 7434
Loss: 30.468508145599976
l2 norm of gradients: 0.19764411585890876
l2 norm of weights: 5.72632598917052

Iteration Number: 7435
Loss: 30.466549549763545
l2 norm of gradients: 0.19762648622314538
l2 norm of weights: 5.726254966644196

Iteration Number: 7436
Loss: 30.464591303459706
l2 norm of gradients: 0.197608860457991
l2 norm of weights: 5.72618394970522

Iteration Number: 7437
Loss: 30.462633406583635
l2 norm of gradients: 0.19759123856180397
l2 norm of weights: 5.726112938352738

Iteration Number: 7438
Loss: 30.460675859029706
l2 norm of gradients: 0.19757362053294372
l2 norm of weights: 5.726041932585903

Iteration Number: 7439
Loss: 30.45871866068491
l2 norm of gradients: 0.19755600636977055
l2 norm of weights: 5.725970932403862

Iteration Number: 7440
Loss: 30.456761811441346
l2 norm of gradients: 0.1975383960706458
l2 norm of weights: 5.725899937805767

Iteration Number: 7441
Loss: 30.45480531119688
l2 norm of gradients: 0.19752078963393185
l2 norm of weights: 5.725828948790767

```

-----
Iteration Number: 7442
Loss: 30.452849159835818
l2 norm of gradients: 0.197503187057992
l2 norm of weights: 5.725757965358014
-----
Iteration Number: 7443
Loss: 30.450893357254078
l2 norm of gradients: 0.1974855883411905
l2 norm of weights: 5.725686987506657
-----
Iteration Number: 7444
Loss: 30.448937903342998
l2 norm of gradients: 0.19746799348189276
l2 norm of weights: 5.72561601523585
-----
Iteration Number: 7445
Loss: 30.44698279799298
l2 norm of gradients: 0.19745040247846496
l2 norm of weights: 5.725545048544743
-----
Iteration Number: 7446
Loss: 30.44502804110105
l2 norm of gradients: 0.1974328153292744
l2 norm of weights: 5.725474087432485
-----
Iteration Number: 7447
Loss: 30.44307363255732
l2 norm of gradients: 0.19741523203268938
l2 norm of weights: 5.725403131898232
-----
Iteration Number: 7448
Loss: 30.441119572250923
l2 norm of gradients: 0.19739765258707903
l2 norm of weights: 5.725332181941133
-----
Iteration Number: 7449
Loss: 30.43916586007425
l2 norm of gradients: 0.1973800769908136
l2 norm of weights: 5.725261237560342
-----
Iteration Number: 7450
Loss: 30.437212495925074
l2 norm of gradients: 0.19736250524226437
l2 norm of weights: 5.7251902987550105
-----
Iteration Number: 7451
Loss: 30.43525947969023

```

```
12 norm of gradients: 0.19734493733980343
12 norm of weights: 5.725119365524291
-----
Iteration Number: 7452
Loss: 30.433306811264856
12 norm of gradients: 0.19732737328180397
12 norm of weights: 5.725048437867338
-----
Iteration Number: 7453
Loss: 30.431354490541647
12 norm of gradients: 0.1973098130666401
12 norm of weights: 5.724977515783304
-----
Iteration Number: 7454
Loss: 30.42940251741063
12 norm of gradients: 0.19729225669268696
12 norm of weights: 5.724906599271342
-----
Iteration Number: 7455
Loss: 30.4274508917681
12 norm of gradients: 0.1972747041583206
12 norm of weights: 5.724835688330606
-----
Iteration Number: 7456
Loss: 30.425499613503266
12 norm of gradients: 0.1972571554619181
12 norm of weights: 5.72476478296025
-----
Iteration Number: 7457
Loss: 30.423548682510706
12 norm of gradients: 0.19723961060185755
12 norm of weights: 5.7246938831594285
-----
Iteration Number: 7458
Loss: 30.421598098685994
12 norm of gradients: 0.1972220695765179
12 norm of weights: 5.724622988927296
-----
Iteration Number: 7459
Loss: 30.419647861915617
12 norm of gradients: 0.1972045323842792
12 norm of weights: 5.724552100263006
-----
Iteration Number: 7460
Loss: 30.417697972096807
12 norm of gradients: 0.19718699902352235
12 norm of weights: 5.724481217165716
-----
```


Iteration Number: 7461
Loss: 30.4157484291212
l2 norm of gradients: 0.19716946949262934
l2 norm of weights: 5.724410339634578

Iteration Number: 7462
Loss: 30.413799232880613
l2 norm of gradients: 0.19715194378998308
l2 norm of weights: 5.724339467668748

Iteration Number: 7463
Loss: 30.411850383272075
l2 norm of gradients: 0.1971344219139674
l2 norm of weights: 5.724268601267383

Iteration Number: 7464
Loss: 30.409901880184876
l2 norm of gradients: 0.19711690386296718
l2 norm of weights: 5.72419774042964

Iteration Number: 7465
Loss: 30.407953723512858
l2 norm of gradients: 0.19709938963536824
l2 norm of weights: 5.7241268851546705

Iteration Number: 7466
Loss: 30.40600591314957
l2 norm of gradients: 0.1970818792295573
l2 norm of weights: 5.724056035441636

Iteration Number: 7467
Loss: 30.40405844898575
l2 norm of gradients: 0.19706437264392218
l2 norm of weights: 5.72398519128969

Iteration Number: 7468
Loss: 30.40211133092019
l2 norm of gradients: 0.1970468698768516
l2 norm of weights: 5.72391435269799

Iteration Number: 7469
Loss: 30.40016455884452
l2 norm of gradients: 0.19702937092673523
l2 norm of weights: 5.723843519665694

Iteration Number: 7470
Loss: 30.398218132650154
l2 norm of gradients: 0.19701187579196372

```
12 norm of weights: 5.723772692191957
-----
Iteration Number: 7471
Loss: 30.396272052229666
12 norm of gradients: 0.19699438447092865
12 norm of weights: 5.7237018702759395
-----
Iteration Number: 7472
Loss: 30.394326317477766
12 norm of gradients: 0.1969768969620226
12 norm of weights: 5.7236310539167965
-----
Iteration Number: 7473
Loss: 30.392380928291317
12 norm of gradients: 0.19695941326363922
12 norm of weights: 5.723560243113687
-----
Iteration Number: 7474
Loss: 30.390435884560198
12 norm of gradients: 0.1969419333741728
12 norm of weights: 5.723489437865771
-----
Iteration Number: 7475
Loss: 30.38849118617522
12 norm of gradients: 0.19692445729201896
12 norm of weights: 5.723418638172204
-----
Iteration Number: 7476
Loss: 30.38654683303822
12 norm of gradients: 0.1969069850155741
12 norm of weights: 5.723347844032147
-----
Iteration Number: 7477
Loss: 30.384602825035877
12 norm of gradients: 0.19688951654323558
12 norm of weights: 5.723277055444758
-----
Iteration Number: 7478
Loss: 30.382659162066265
12 norm of gradients: 0.19687205187340168
12 norm of weights: 5.723206272409195
-----
Iteration Number: 7479
Loss: 30.380715844021456
12 norm of gradients: 0.1968545910044718
12 norm of weights: 5.72313549492462
-----
Iteration Number: 7480
```

Loss: 30.37877287079309
l2 norm of gradients: 0.19683713393484614
l2 norm of weights: 5.72306472299019

Iteration Number: 7481
Loss: 30.376830242282615
l2 norm of gradients: 0.19681968066292585
l2 norm of weights: 5.722993956605067

Iteration Number: 7482
Loss: 30.374887958375602
l2 norm of gradients: 0.19680223118711315
l2 norm of weights: 5.72292319576841

Iteration Number: 7483
Loss: 30.372946018969067
l2 norm of gradients: 0.19678478550581116
l2 norm of weights: 5.72285244047938

Iteration Number: 7484
Loss: 30.37100442395522
l2 norm of gradients: 0.19676734361742396
l2 norm of weights: 5.722781690737137

Iteration Number: 7485
Loss: 30.369063173235357
l2 norm of gradients: 0.19674990552035657
l2 norm of weights: 5.7227109465408414

Iteration Number: 7486
Loss: 30.367122266697773
l2 norm of gradients: 0.1967324712130149
l2 norm of weights: 5.722640207889656

Iteration Number: 7487
Loss: 30.365181704236978
l2 norm of gradients: 0.1967150406938059
l2 norm of weights: 5.72256947478274

Iteration Number: 7488
Loss: 30.363241485748972
l2 norm of gradients: 0.19669761396113752
l2 norm of weights: 5.722498747219256

Iteration Number: 7489
Loss: 30.36130161112499
l2 norm of gradients: 0.19668019101341858
l2 norm of weights: 5.722428025198366

```

-----
Iteration Number: 7490
Loss: 30.359362080267136
l2 norm of gradients: 0.19666277184905867
l2 norm of weights: 5.722357308719231
-----
Iteration Number: 7491
Loss: 30.357422893059375
l2 norm of gradients: 0.19664535646646877
l2 norm of weights: 5.722286597781015
-----
Iteration Number: 7492
Loss: 30.355484049403696
l2 norm of gradients: 0.19662794486406043
l2 norm of weights: 5.722215892382879
-----
Iteration Number: 7493
Loss: 30.35354554919267
l2 norm of gradients: 0.19661053704024622
l2 norm of weights: 5.722145192523986
-----
Iteration Number: 7494
Loss: 30.351607392318794
l2 norm of gradients: 0.1965931329934398
l2 norm of weights: 5.722074498203499
-----
Iteration Number: 7495
Loss: 30.349669578679716
l2 norm of gradients: 0.19657573272205553
l2 norm of weights: 5.72200380942058
-----
Iteration Number: 7496
Loss: 30.347732108167996
l2 norm of gradients: 0.19655833622450905
l2 norm of weights: 5.721933126174394
-----
Iteration Number: 7497
Loss: 30.345794980681507
l2 norm of gradients: 0.19654094349921655
l2 norm of weights: 5.721862448464106
-----
Iteration Number: 7498
Loss: 30.343858196113317
l2 norm of gradients: 0.19652355454459552
l2 norm of weights: 5.7217917762888755
-----
Iteration Number: 7499
Loss: 30.34192175435689

```

```

12 norm of gradients: 0.19650616935906418
12 norm of weights: 5.721721109647871
-----
Iteration Number: 7500
Loss: 30.339985655308322
12 norm of gradients: 0.1964887879410418
12 norm of weights: 5.721650448540254
-----
Iteration Number: 7501
Loss: 30.338049898861804
12 norm of gradients: 0.1964714102889484
12 norm of weights: 5.72157979296519
-----
Iteration Number: 7502
Loss: 30.336114484916994
12 norm of gradients: 0.19645403640120526
12 norm of weights: 5.721509142921844
-----
Iteration Number: 7503
Loss: 30.33417941336186
12 norm of gradients: 0.19643666627623427
12 norm of weights: 5.721438498409381
-----
Iteration Number: 7504
Loss: 30.33224468409572
12 norm of gradients: 0.1964192999124585
12 norm of weights: 5.721367859426966
-----
Iteration Number: 7505
Loss: 30.33031029701315
12 norm of gradients: 0.1964019373083018
12 norm of weights: 5.721297225973764
-----
Iteration Number: 7506
Loss: 30.32837625200932
12 norm of gradients: 0.19638457846218904
12 norm of weights: 5.7212265980489425
-----
Iteration Number: 7507
Loss: 30.32644254897681
12 norm of gradients: 0.196367223372546
12 norm of weights: 5.721155975651667
-----
Iteration Number: 7508
Loss: 30.32450918781897
12 norm of gradients: 0.19634987203779944
12 norm of weights: 5.721085358781101
-----

```

Iteration Number: 7509
Loss: 30.322576168421374
l2 norm of gradients: 0.1963325244563769
l2 norm of weights: 5.721014747436414

Iteration Number: 7510
Loss: 30.320643490683057
l2 norm of gradients: 0.19631518062670708
l2 norm of weights: 5.720944141616772

Iteration Number: 7511
Loss: 30.318711154505085
l2 norm of gradients: 0.19629784054721944
l2 norm of weights: 5.720873541321341

Iteration Number: 7512
Loss: 30.31677915977501
l2 norm of gradients: 0.1962805042163444
l2 norm of weights: 5.720802946549289

Iteration Number: 7513
Loss: 30.314847506390645
l2 norm of gradients: 0.19626317163251342
l2 norm of weights: 5.720732357299783

Iteration Number: 7514
Loss: 30.312916194251564
l2 norm of gradients: 0.19624584279415874
l2 norm of weights: 5.720661773571991

Iteration Number: 7515
Loss: 30.310985223247265
l2 norm of gradients: 0.1962285176997137
l2 norm of weights: 5.7205911953650785

Iteration Number: 7516
Loss: 30.30905459327917
l2 norm of gradients: 0.19621119634761236
l2 norm of weights: 5.720520622678217

Iteration Number: 7517
Loss: 30.307124304242155
l2 norm of gradients: 0.1961938787362898
l2 norm of weights: 5.7204500555105735

Iteration Number: 7518
Loss: 30.305194356028355
l2 norm of gradients: 0.19617656486418214

```
12 norm of weights: 5.720379493861316
-----
Iteration Number: 7519
Loss: 30.303264748536236
12 norm of gradients: 0.1961592547297263
12 norm of weights: 5.720308937729614
-----
Iteration Number: 7520
Loss: 30.301335481662747
12 norm of gradients: 0.19614194833136012
12 norm of weights: 5.720238387114635
-----
Iteration Number: 7521
Loss: 30.299406555300507
12 norm of gradients: 0.1961246456675224
12 norm of weights: 5.72016784201555
-----
Iteration Number: 7522
Loss: 30.29747796935035
12 norm of gradients: 0.19610734673665292
12 norm of weights: 5.720097302431527
-----
Iteration Number: 7523
Loss: 30.295549723701807
12 norm of gradients: 0.19609005153719222
12 norm of weights: 5.720026768361736
-----
Iteration Number: 7524
Loss: 30.293621818259947
12 norm of gradients: 0.19607276006758198
12 norm of weights: 5.719956239805349
-----
Iteration Number: 7525
Loss: 30.291694252909853
12 norm of gradients: 0.1960554723262647
12 norm of weights: 5.7198857167615325
-----
Iteration Number: 7526
Loss: 30.28976702755704
12 norm of gradients: 0.19603818831168368
12 norm of weights: 5.719815199229459
-----
Iteration Number: 7527
Loss: 30.287840142094925
12 norm of gradients: 0.1960209080222833
12 norm of weights: 5.7197446872082995
-----
Iteration Number: 7528
```

Loss: 30.28591359641909
l2 norm of gradients: 0.1960036314565089
l2 norm of weights: 5.719674180697225

Iteration Number: 7529
Loss: 30.283987390429388
l2 norm of gradients: 0.19598635861280655
l2 norm of weights: 5.719603679695404

Iteration Number: 7530
Loss: 30.282061524014217
l2 norm of gradients: 0.19596908948962335
l2 norm of weights: 5.71953318420201

Iteration Number: 7531
Loss: 30.280135997079473
l2 norm of gradients: 0.19595182408540734
l2 norm of weights: 5.719462694216214

Iteration Number: 7532
Loss: 30.278210809513297
l2 norm of gradients: 0.1959345623986075
l2 norm of weights: 5.7193922097371885

Iteration Number: 7533
Loss: 30.276285961217653
l2 norm of gradients: 0.1959173044276735
l2 norm of weights: 5.719321730764105

Iteration Number: 7534
Loss: 30.274361452090144
l2 norm of gradients: 0.19590005017105624
l2 norm of weights: 5.7192512572961345

Iteration Number: 7535
Loss: 30.272437282023414
l2 norm of gradients: 0.19588279962720737
l2 norm of weights: 5.719180789332451

Iteration Number: 7536
Loss: 30.27051345091605
l2 norm of gradients: 0.1958655527945794
l2 norm of weights: 5.719110326872226

Iteration Number: 7537
Loss: 30.268589958665682
l2 norm of gradients: 0.19584830967162592
l2 norm of weights: 5.719039869914633


```
-----  
Iteration Number: 7538  
Loss: 30.266666805168768  
l2 norm of gradients: 0.19583107025680135  
l2 norm of weights: 5.718969418458844  
-----  
Iteration Number: 7539  
Loss: 30.26474399032218  
l2 norm of gradients: 0.1958138345485609  
l2 norm of weights: 5.718898972504036  
-----  
Iteration Number: 7540  
Loss: 30.26282151401981  
l2 norm of gradients: 0.1957966025453609  
l2 norm of weights: 5.718828532049378  
-----  
Iteration Number: 7541  
Loss: 30.260899376163373  
l2 norm of gradients: 0.19577937424565842  
l2 norm of weights: 5.718758097094047  
-----  
Iteration Number: 7542  
Loss: 30.258977576645886  
l2 norm of gradients: 0.19576214964791158  
l2 norm of weights: 5.718687667637215  
-----  
Iteration Number: 7543  
Loss: 30.257056115366723  
l2 norm of gradients: 0.1957449287505793  
l2 norm of weights: 5.718617243678058  
-----  
Iteration Number: 7544  
Loss: 30.255134992226044  
l2 norm of gradients: 0.19572771155212138  
l2 norm of weights: 5.718546825215749  
-----  
Iteration Number: 7545  
Loss: 30.253214207113494  
l2 norm of gradients: 0.19571049805099877  
l2 norm of weights: 5.718476412249463  
-----  
Iteration Number: 7546  
Loss: 30.25129375993407  
l2 norm of gradients: 0.195693288245673  
l2 norm of weights: 5.718406004778376  
-----  
Iteration Number: 7547  
Loss: 30.249373650579635
```

```

12 norm of gradients: 0.19567608213460672
12 norm of weights: 5.718335602801662
-----
Iteration Number: 7548
Loss: 30.247453878952083
12 norm of gradients: 0.19565887971626342
12 norm of weights: 5.718265206318497
-----
Iteration Number: 7549
Loss: 30.245534444941693
12 norm of gradients: 0.1956416809891075
12 norm of weights: 5.718194815328057
-----
Iteration Number: 7550
Loss: 30.24361534845324
12 norm of gradients: 0.1956244859516042
12 norm of weights: 5.718124429829517
-----
Iteration Number: 7551
Loss: 30.24169658937938
12 norm of gradients: 0.19560729460221987
12 norm of weights: 5.7180540498220545
-----
Iteration Number: 7552
Loss: 30.23977816762223
12 norm of gradients: 0.19559010693942144
12 norm of weights: 5.717983675304843
-----
Iteration Number: 7553
Loss: 30.237860083074843
12 norm of gradients: 0.19557292296167705
12 norm of weights: 5.717913306277063
-----
Iteration Number: 7554
Loss: 30.235942335638526
12 norm of gradients: 0.19555574266745554
12 norm of weights: 5.717842942737888
-----
Iteration Number: 7555
Loss: 30.234024925208075
12 norm of gradients: 0.19553856605522674
12 norm of weights: 5.717772584686497
-----
Iteration Number: 7556
Loss: 30.232107851682212
12 norm of gradients: 0.19552139312346134
12 norm of weights: 5.717702232122065
-----

```

Iteration Number: 7557
Loss: 30.23019111496074
l2 norm of gradients: 0.19550422387063093
l2 norm of weights: 5.7176318850437715

Iteration Number: 7558
Loss: 30.228274714938255
l2 norm of gradients: 0.1954870582952081
l2 norm of weights: 5.717561543450794

Iteration Number: 7559
Loss: 30.226358651514296
l2 norm of gradients: 0.19546989639566623
l2 norm of weights: 5.717491207342309

Iteration Number: 7560
Loss: 30.22444292458374
l2 norm of gradients: 0.19545273817047956
l2 norm of weights: 5.7174208767174965

Iteration Number: 7561
Loss: 30.222527534050737
l2 norm of gradients: 0.19543558361812327
l2 norm of weights: 5.7173505515755325

Iteration Number: 7562
Loss: 30.22061247981131
l2 norm of gradients: 0.19541843273707354
l2 norm of weights: 5.717280231915599

Iteration Number: 7563
Loss: 30.218697761759778
l2 norm of gradients: 0.19540128552580732
l2 norm of weights: 5.717209917736873

Iteration Number: 7564
Loss: 30.216783379798258
l2 norm of gradients: 0.19538414198280246
l2 norm of weights: 5.717139609038533

Iteration Number: 7565
Loss: 30.214869333822957
l2 norm of gradients: 0.19536700210653776
l2 norm of weights: 5.7170693058197575

Iteration Number: 7566
Loss: 30.21295562372987
l2 norm of gradients: 0.19534986589549286

```

12 norm of weights: 5.716999008079728
-----
Iteration Number: 7567
Loss: 30.211042249423826
12 norm of gradients: 0.19533273334814832
12 norm of weights: 5.716928715817625
-----
Iteration Number: 7568
Loss: 30.209129210794426
12 norm of gradients: 0.19531560446298565
12 norm of weights: 5.716858429032625
-----
Iteration Number: 7569
Loss: 30.207216507748463
12 norm of gradients: 0.19529847923848706
12 norm of weights: 5.71678814772391
-----
Iteration Number: 7570
Loss: 30.20530414018132
12 norm of gradients: 0.1952813576731359
12 norm of weights: 5.716717871890661
-----
Iteration Number: 7571
Loss: 30.203392107988318
12 norm of gradients: 0.1952642397654162
12 norm of weights: 5.716647601532058
-----
Iteration Number: 7572
Loss: 30.20148041107433
12 norm of gradients: 0.19524712551381304
12 norm of weights: 5.7165773366472825
-----
Iteration Number: 7573
Loss: 30.19956904933101
12 norm of gradients: 0.19523001491681224
12 norm of weights: 5.716507077235516
-----
Iteration Number: 7574
Loss: 30.197658022662203
12 norm of gradients: 0.19521290797290064
12 norm of weights: 5.716436823295936
-----
Iteration Number: 7575
Loss: 30.195747330962067
12 norm of gradients: 0.19519580468056585
12 norm of weights: 5.716366574827729
-----
Iteration Number: 7576

```

Loss: 30.193836974133447
l2 norm of gradients: 0.19517870503829648
l2 norm of weights: 5.716296331830074

Iteration Number: 7577
Loss: 30.191926952071384
l2 norm of gradients: 0.19516160904458194
l2 norm of weights: 5.716226094302154

Iteration Number: 7578
Loss: 30.190017264681437
l2 norm of gradients: 0.1951445166979125
l2 norm of weights: 5.716155862243152

Iteration Number: 7579
Loss: 30.188107911856385
l2 norm of gradients: 0.19512742799677946
l2 norm of weights: 5.716085635652248

Iteration Number: 7580
Loss: 30.186198893492648
l2 norm of gradients: 0.1951103429396748
l2 norm of weights: 5.716015414528626

Iteration Number: 7581
Loss: 30.184290209495558
l2 norm of gradients: 0.19509326152509157
l2 norm of weights: 5.715945198871469

Iteration Number: 7582
Loss: 30.182381859762806
l2 norm of gradients: 0.1950761837515236
l2 norm of weights: 5.7158749886799605

Iteration Number: 7583
Loss: 30.18047384418944
l2 norm of gradients: 0.19505910961746556
l2 norm of weights: 5.715804783953283

Iteration Number: 7584
Loss: 30.178566162678173
l2 norm of gradients: 0.19504203912141316
l2 norm of weights: 5.71573458469062

Iteration Number: 7585
Loss: 30.176658815129937
l2 norm of gradients: 0.19502497226186283
l2 norm of weights: 5.715664390891155

```
-----  
Iteration Number: 7586  
Loss: 30.174751801438116  
l2 norm of gradients: 0.1950079090373119  
l2 norm of weights: 5.715594202554073  
-----  
Iteration Number: 7587  
Loss: 30.172845121506207  
l2 norm of gradients: 0.1949908494462588  
l2 norm of weights: 5.7155240196785595  
-----  
Iteration Number: 7588  
Loss: 30.170938775233736  
l2 norm of gradients: 0.1949737934872024  
l2 norm of weights: 5.715453842263796  
-----  
Iteration Number: 7589  
Loss: 30.16903276251979  
l2 norm of gradients: 0.19495674115864287  
l2 norm of weights: 5.715383670308968  
-----  
Iteration Number: 7590  
Loss: 30.16712708325805  
l2 norm of gradients: 0.19493969245908102  
l2 norm of weights: 5.715313503813262  
-----  
Iteration Number: 7591  
Loss: 30.165221737357772  
l2 norm of gradients: 0.19492264738701862  
l2 norm of weights: 5.715243342775862  
-----  
Iteration Number: 7592  
Loss: 30.16331672471128  
l2 norm of gradients: 0.19490560594095832  
l2 norm of weights: 5.715173187195954  
-----  
Iteration Number: 7593  
Loss: 30.161412045220512  
l2 norm of gradients: 0.1948885681194036  
l2 norm of weights: 5.715103037072721  
-----  
Iteration Number: 7594  
Loss: 30.15950769878456  
l2 norm of gradients: 0.19487153392085882  
l2 norm of weights: 5.715032892405352  
-----  
Iteration Number: 7595  
Loss: 30.157603685303258
```

```

12 norm of gradients: 0.19485450334382928
12 norm of weights: 5.714962753193031
-----
Iteration Number: 7596
Loss: 30.15570000467694
12 norm of gradients: 0.194837476386821
12 norm of weights: 5.714892619434946
-----
Iteration Number: 7597
Loss: 30.153796656801312
12 norm of gradients: 0.19482045304834109
12 norm of weights: 5.714822491130282
-----
Iteration Number: 7598
Loss: 30.151893641584344
12 norm of gradients: 0.19480343332689734
12 norm of weights: 5.714752368278226
-----
Iteration Number: 7599
Loss: 30.14999095891541
12 norm of gradients: 0.19478641722099854
12 norm of weights: 5.714682250877965
-----
Iteration Number: 7600
Loss: 30.148088608704324
12 norm of gradients: 0.19476940472915424
12 norm of weights: 5.714612138928688
-----
Iteration Number: 7601
Loss: 30.14618659084569
12 norm of gradients: 0.19475239584987492
12 norm of weights: 5.714542032429579
-----
Iteration Number: 7602
Loss: 30.144284905238685
12 norm of gradients: 0.1947353905816719
12 norm of weights: 5.714471931379828
-----
Iteration Number: 7603
Loss: 30.14238355178896
12 norm of gradients: 0.1947183889230574
12 norm of weights: 5.714401835778621
-----
Iteration Number: 7604
Loss: 30.14048253038887
12 norm of gradients: 0.19470139087254457
12 norm of weights: 5.714331745625147
-----

```

Iteration Number: 7605
Loss: 30.13858184094425
l2 norm of gradients: 0.19468439642864727
l2 norm of weights: 5.714261660918595

Iteration Number: 7606
Loss: 30.13668148335124
l2 norm of gradients: 0.19466740558988024
l2 norm of weights: 5.714191581658152

Iteration Number: 7607
Loss: 30.134781457513817
l2 norm of gradients: 0.19465041835475932
l2 norm of weights: 5.714121507843008

Iteration Number: 7608
Loss: 30.132881763332044
l2 norm of gradients: 0.1946334347218009
l2 norm of weights: 5.7140514394723505

Iteration Number: 7609
Loss: 30.130982400700418
l2 norm of gradients: 0.19461645468952243
l2 norm of weights: 5.71398137654537

Iteration Number: 7610
Loss: 30.129083369528495
l2 norm of gradients: 0.19459947825644222
l2 norm of weights: 5.713911319061256

Iteration Number: 7611
Loss: 30.12718466970964
l2 norm of gradients: 0.19458250542107933
l2 norm of weights: 5.713841267019197

Iteration Number: 7612
Loss: 30.125286301145444
l2 norm of gradients: 0.1945655361819538
l2 norm of weights: 5.7137712204183835

Iteration Number: 7613
Loss: 30.12338826373948
l2 norm of gradients: 0.1945485705375864
l2 norm of weights: 5.713701179258004

Iteration Number: 7614
Loss: 30.1214905573878
l2 norm of gradients: 0.19453160848649884


```

12 norm of weights: 5.713631143537251
-----
Iteration Number: 7615
Loss: 30.11959318199436
12 norm of gradients: 0.19451465002721383
12 norm of weights: 5.713561113255313
-----
Iteration Number: 7616
Loss: 30.11769613745843
12 norm of gradients: 0.1944976951582546
12 norm of weights: 5.7134910884113825
-----
Iteration Number: 7617
Loss: 30.11579942368212
12 norm of gradients: 0.1944807438781455
12 norm of weights: 5.713421069004649
-----
Iteration Number: 7618
Loss: 30.113903040563223
12 norm of gradients: 0.19446379618541174
12 norm of weights: 5.713351055034305
-----
Iteration Number: 7619
Loss: 30.112006988006044
12 norm of gradients: 0.19444685207857929
12 norm of weights: 5.7132810464995405
-----
Iteration Number: 7620
Loss: 30.110111265909392
12 norm of gradients: 0.19442991155617498
12 norm of weights: 5.713211043399547
-----
Iteration Number: 7621
Loss: 30.108215874174167
12 norm of gradients: 0.19441297461672652
12 norm of weights: 5.713141045733518
-----
Iteration Number: 7622
Loss: 30.106320812701142
12 norm of gradients: 0.1943960412587625
12 norm of weights: 5.713071053500643
-----
Iteration Number: 7623
Loss: 30.104426081393065
12 norm of gradients: 0.19437911148081233
12 norm of weights: 5.713001066700117
-----
Iteration Number: 7624

```

```

Loss: 30.102531680148033
12 norm of gradients: 0.19436218528140634
12 norm of weights: 5.712931085331131
-----
Iteration Number: 7625
Loss: 30.100637608869143
12 norm of gradients: 0.1943452626590755
12 norm of weights: 5.712861109392876
-----
Iteration Number: 7626
Loss: 30.098743867459138
12 norm of gradients: 0.194328343612352
12 norm of weights: 5.7127911388845485
-----
Iteration Number: 7627
Loss: 30.09685045581266
12 norm of gradients: 0.19431142813976854
12 norm of weights: 5.712721173805339
-----
Iteration Number: 7628
Loss: 30.094957373838476
12 norm of gradients: 0.19429451623985883
12 norm of weights: 5.712651214154441
-----
Iteration Number: 7629
Loss: 30.09306462143288
12 norm of gradients: 0.19427760791115742
12 norm of weights: 5.712581259931049
-----
Iteration Number: 7630
Loss: 30.0911721984996
12 norm of gradients: 0.19426070315219973
12 norm of weights: 5.7125113111343575
-----
Iteration Number: 7631
Loss: 30.089280104940578
12 norm of gradients: 0.19424380196152194
12 norm of weights: 5.7124413677635575
-----
Iteration Number: 7632
Loss: 30.08738834065492
12 norm of gradients: 0.19422690433766115
12 norm of weights: 5.712371429817846
-----
Iteration Number: 7633
Loss: 30.085496905545707
12 norm of gradients: 0.19421001027915533
12 norm of weights: 5.712301497296417

```

```

-----
Iteration Number: 7634
Loss: 30.083605799511304
l2 norm of gradients: 0.1941931197845432
l2 norm of weights: 5.712231570198464
-----
Iteration Number: 7635
Loss: 30.081715022456706
l2 norm of gradients: 0.1941762328523645
l2 norm of weights: 5.712161648523182
-----
Iteration Number: 7636
Loss: 30.07982457428508
l2 norm of gradients: 0.19415934948115954
l2 norm of weights: 5.7120917322697675
-----
Iteration Number: 7637
Loss: 30.07793445489262
l2 norm of gradients: 0.1941424696694698
l2 norm of weights: 5.712021821437415
-----
Iteration Number: 7638
Loss: 30.07604466418213
l2 norm of gradients: 0.19412559341583738
l2 norm of weights: 5.71195191602532
-----
Iteration Number: 7639
Loss: 30.07415520206015
l2 norm of gradients: 0.19410872071880522
l2 norm of weights: 5.7118820160326775
-----
Iteration Number: 7640
Loss: 30.072266068423296
l2 norm of gradients: 0.19409185157691733
l2 norm of weights: 5.711812121458684
-----
Iteration Number: 7641
Loss: 30.070377263178464
l2 norm of gradients: 0.19407498598871833
l2 norm of weights: 5.711742232302537
-----
Iteration Number: 7642
Loss: 30.06848878622036
l2 norm of gradients: 0.19405812395275376
l2 norm of weights: 5.7116723485634315
-----
Iteration Number: 7643
Loss: 30.06660063745839

```

```

12 norm of gradients: 0.19404126546756997
12 norm of weights: 5.711602470240563
-----
Iteration Number: 7644
Loss: 30.064712816788113
12 norm of gradients: 0.19402441053171432
12 norm of weights: 5.711532597333131
-----
Iteration Number: 7645
Loss: 30.062825324114712
12 norm of gradients: 0.19400755914373471
12 norm of weights: 5.711462729840331
-----
Iteration Number: 7646
Loss: 30.060938159341895
12 norm of gradients: 0.19399071130218012
12 norm of weights: 5.711392867761361
-----
Iteration Number: 7647
Loss: 30.059051322369296
12 norm of gradients: 0.19397386700560038
12 norm of weights: 5.711323011095416
-----
Iteration Number: 7648
Loss: 30.057164813098105
12 norm of gradients: 0.1939570262525459
12 norm of weights: 5.711253159841697
-----
Iteration Number: 7649
Loss: 30.055278631429484
12 norm of gradients: 0.1939401890415683
12 norm of weights: 5.711183313999401
-----
Iteration Number: 7650
Loss: 30.05339277727102
12 norm of gradients: 0.1939233553712197
12 norm of weights: 5.711113473567726
-----
Iteration Number: 7651
Loss: 30.051507250522683
12 norm of gradients: 0.19390652524005328
12 norm of weights: 5.71104363854587
-----
Iteration Number: 7652
Loss: 30.049622051083773
12 norm of gradients: 0.19388969864662292
12 norm of weights: 5.71097380893303
-----

```

Iteration Number: 7653
Loss: 30.047737178859276
l2 norm of gradients: 0.1938728755894835
l2 norm of weights: 5.710903984728409

Iteration Number: 7654
Loss: 30.04585263375335
l2 norm of gradients: 0.19385605606719047
l2 norm of weights: 5.710834165931202

Iteration Number: 7655
Loss: 30.04396841566572
l2 norm of gradients: 0.19383924007830042
l2 norm of weights: 5.71076435254061

Iteration Number: 7656
Loss: 30.04208452449962
l2 norm of gradients: 0.19382242762137056
l2 norm of weights: 5.710694544555833

Iteration Number: 7657
Loss: 30.04020096015798
l2 norm of gradients: 0.19380561869495905
l2 norm of weights: 5.710624741976068

Iteration Number: 7658
Loss: 30.038317722540935
l2 norm of gradients: 0.1937888132976248
l2 norm of weights: 5.710554944800519

Iteration Number: 7659
Loss: 30.03643481155256
l2 norm of gradients: 0.19377201142792752
l2 norm of weights: 5.710485153028382

Iteration Number: 7660
Loss: 30.0345522270991
l2 norm of gradients: 0.19375521308442792
l2 norm of weights: 5.71041536665886

Iteration Number: 7661
Loss: 30.032669969076835
l2 norm of gradients: 0.1937384182656875
l2 norm of weights: 5.710345585691154

Iteration Number: 7662
Loss: 30.03078803739121
l2 norm of gradients: 0.19372162697026843

```

12 norm of weights: 5.710275810124461
-----
Iteration Number: 7663
Loss: 30.028906431948577
12 norm of gradients: 0.19370483919673379
12 norm of weights: 5.710206039957985
-----
Iteration Number: 7664
Loss: 30.02702515264742
12 norm of gradients: 0.1936880549436476
12 norm of weights: 5.710136275190927
-----
Iteration Number: 7665
Loss: 30.02514419939362
12 norm of gradients: 0.19367127420957458
12 norm of weights: 5.710066515822489
-----
Iteration Number: 7666
Loss: 30.023263572090382
12 norm of gradients: 0.19365449699308035
12 norm of weights: 5.70999676185187
-----
Iteration Number: 7667
Loss: 30.02138327063658
12 norm of gradients: 0.1936377232927313
12 norm of weights: 5.709927013278274
-----
Iteration Number: 7668
Loss: 30.019503294937326
12 norm of gradients: 0.19362095310709468
12 norm of weights: 5.709857270100903
-----
Iteration Number: 7669
Loss: 30.017623644895387
12 norm of gradients: 0.19360418643473856
12 norm of weights: 5.709787532318958
-----
Iteration Number: 7670
Loss: 30.015744320416772
12 norm of gradients: 0.1935874232742319
12 norm of weights: 5.709717799931642
-----
Iteration Number: 7671
Loss: 30.013865321401287
12 norm of gradients: 0.19357066362414438
12 norm of weights: 5.709648072938159
-----
Iteration Number: 7672

```

Loss: 30.011986647753496
l2 norm of gradients: 0.19355390748304646
l2 norm of weights: 5.7095783513377105

Iteration Number: 7673
Loss: 30.01010829937743
l2 norm of gradients: 0.19353715484950965
l2 norm of weights: 5.709508635129499

Iteration Number: 7674
Loss: 30.00823027617309
l2 norm of gradients: 0.19352040572210613
l2 norm of weights: 5.70943892431273

Iteration Number: 7675
Loss: 30.006352578048705
l2 norm of gradients: 0.19350366009940886
l2 norm of weights: 5.7093692188866045

Iteration Number: 7676
Loss: 30.004475204903354
l2 norm of gradients: 0.19348691797999176
l2 norm of weights: 5.709299518850329

Iteration Number: 7677
Loss: 30.00259815664411
l2 norm of gradients: 0.1934701793624294
l2 norm of weights: 5.709229824203105

Iteration Number: 7678
Loss: 30.00072143317189
l2 norm of gradients: 0.19345344424529737
l2 norm of weights: 5.709160134944138

Iteration Number: 7679
Loss: 29.998845034391053
l2 norm of gradients: 0.1934367126271719
l2 norm of weights: 5.709090451072632

Iteration Number: 7680
Loss: 29.996968960205315
l2 norm of gradients: 0.19341998450663023
l2 norm of weights: 5.709020772587792

Iteration Number: 7681
Loss: 29.99509321051649
l2 norm of gradients: 0.19340325988225018
l2 norm of weights: 5.708951099488822

```
-----  
Iteration Number: 7682  
Loss: 29.993217785233476  
l2 norm of gradients: 0.19338653875261055  
l2 norm of weights: 5.708881431774929  
-----  
Iteration Number: 7683  
Loss: 29.991342684255766  
l2 norm of gradients: 0.19336982111629103  
l2 norm of weights: 5.708811769445316  
-----  
Iteration Number: 7684  
Loss: 29.98946790748713  
l2 norm of gradients: 0.1933531069718719  
l2 norm of weights: 5.7087421124991895  
-----  
Iteration Number: 7685  
Loss: 29.98759345482928  
l2 norm of gradients: 0.19333639631793448  
l2 norm of weights: 5.708672460935755  
-----  
Iteration Number: 7686  
Loss: 29.985719326192598  
l2 norm of gradients: 0.1933196891530607  
l2 norm of weights: 5.70860281475422  
-----  
Iteration Number: 7687  
Loss: 29.983845521475814  
l2 norm of gradients: 0.19330298547583352  
l2 norm of weights: 5.708533173953788  
-----  
Iteration Number: 7688  
Loss: 29.981972040585934  
l2 norm of gradients: 0.19328628528483652  
l2 norm of weights: 5.708463538533666  
-----  
Iteration Number: 7689  
Loss: 29.98009888342296  
l2 norm of gradients: 0.1932695885786543  
l2 norm of weights: 5.708393908493062  
-----  
Iteration Number: 7690  
Loss: 29.978226049896286  
l2 norm of gradients: 0.1932528953558721  
l2 norm of weights: 5.708324283831182  
-----  
Iteration Number: 7691  
Loss: 29.97635353990299
```



```

12 norm of gradients: 0.193236205615076
12 norm of weights: 5.7082546645472325
-----
Iteration Number: 7692
Loss: 29.97448135335511
12 norm of gradients: 0.193219519354853
12 norm of weights: 5.70818505064042
-----
Iteration Number: 7693
Loss: 29.97260949015215
12 norm of gradients: 0.19320283657379078
12 norm of weights: 5.708115442109954
-----
Iteration Number: 7694
Loss: 29.970737950197012
12 norm of gradients: 0.19318615727047794
12 norm of weights: 5.708045838955041
-----
Iteration Number: 7695
Loss: 29.968866733398244
12 norm of gradients: 0.19316948144350385
12 norm of weights: 5.707976241174889
-----
Iteration Number: 7696
Loss: 29.966995839657656
12 norm of gradients: 0.1931528090914587
12 norm of weights: 5.707906648768706
-----
Iteration Number: 7697
Loss: 29.96512526888046
12 norm of gradients: 0.1931361402129334
12 norm of weights: 5.7078370617357
-----
Iteration Number: 7698
Loss: 29.963255020970337
12 norm of gradients: 0.19311947480651986
12 norm of weights: 5.707767480075079
-----
Iteration Number: 7699
Loss: 29.96138509583054
12 norm of gradients: 0.19310281287081058
12 norm of weights: 5.707697903786053
-----
Iteration Number: 7700
Loss: 29.959515493368155
12 norm of gradients: 0.19308615440439908
12 norm of weights: 5.70762833286783
-----

```

Iteration Number: 7701
Loss: 29.957646213485884
l2 norm of gradients: 0.1930694994058795
l2 norm of weights: 5.707558767319619

Iteration Number: 7702
Loss: 29.95577725608802
l2 norm of gradients: 0.1930528478738469
l2 norm of weights: 5.707489207140631

Iteration Number: 7703
Loss: 29.953908621084405
l2 norm of gradients: 0.19303619980689718
l2 norm of weights: 5.707419652330073

Iteration Number: 7704
Loss: 29.9520403083723
l2 norm of gradients: 0.19301955520362699
l2 norm of weights: 5.707350102887156

Iteration Number: 7705
Loss: 29.9501723178595
l2 norm of gradients: 0.19300291406263365
l2 norm of weights: 5.707280558811091

Iteration Number: 7706
Loss: 29.94830464945252
l2 norm of gradients: 0.19298627638251553
l2 norm of weights: 5.707211020101085

Iteration Number: 7707
Loss: 29.946437303053195
l2 norm of gradients: 0.19296964216187168
l2 norm of weights: 5.707141486756353

Iteration Number: 7708
Loss: 29.94457027856722
l2 norm of gradients: 0.19295301139930202
l2 norm of weights: 5.707071958776102

Iteration Number: 7709
Loss: 29.94270357590182
l2 norm of gradients: 0.19293638409340708
l2 norm of weights: 5.707002436159543

Iteration Number: 7710
Loss: 29.9408371949578
l2 norm of gradients: 0.19291976024278853

```
12 norm of weights: 5.706932918905889
-----
Iteration Number: 7711
Loss: 29.93897113564102
12 norm of gradients: 0.19290313984604854
12 norm of weights: 5.70686340701435
-----
Iteration Number: 7712
Loss: 29.937105397863565
12 norm of gradients: 0.19288652290179015
12 norm of weights: 5.706793900484137
-----
Iteration Number: 7713
Loss: 29.935239981519036
12 norm of gradients: 0.19286990940861737
12 norm of weights: 5.706724399314463
-----
Iteration Number: 7714
Loss: 29.933374886522877
12 norm of gradients: 0.19285329936513476
12 norm of weights: 5.706654903504539
-----
Iteration Number: 7715
Loss: 29.931510112772127
12 norm of gradients: 0.1928366927699479
12 norm of weights: 5.706585413053577
-----
Iteration Number: 7716
Loss: 29.929645660176252
12 norm of gradients: 0.19282008962166308
12 norm of weights: 5.7065159279607895
-----
Iteration Number: 7717
Loss: 29.927781528642374
12 norm of gradients: 0.19280348991888738
12 norm of weights: 5.706446448225388
-----
Iteration Number: 7718
Loss: 29.925917718071766
12 norm of gradients: 0.1927868936602286
12 norm of weights: 5.706376973846588
-----
Iteration Number: 7719
Loss: 29.92405422837124
12 norm of gradients: 0.19277030084429553
12 norm of weights: 5.706307504823599
-----
Iteration Number: 7720
```

Loss: 29.92219105944401
l2 norm of gradients: 0.19275371146969758
l2 norm of weights: 5.706238041155636

Iteration Number: 7721
Loss: 29.92032821120003
l2 norm of gradients: 0.19273712553504516
l2 norm of weights: 5.706168582841913

Iteration Number: 7722
Loss: 29.918465683541193
l2 norm of gradients: 0.19272054303894917
l2 norm of weights: 5.706099129881642

Iteration Number: 7723
Loss: 29.91660347637684
l2 norm of gradients: 0.19270396398002157
l2 norm of weights: 5.706029682274037

Iteration Number: 7724
Loss: 29.914741589607022
l2 norm of gradients: 0.19268738835687513
l2 norm of weights: 5.705960240018315

Iteration Number: 7725
Loss: 29.91288002314192
l2 norm of gradients: 0.1926708161681231
l2 norm of weights: 5.705890803113684

Iteration Number: 7726
Loss: 29.911018776883914
l2 norm of gradients: 0.19265424741238
l2 norm of weights: 5.705821371559364

Iteration Number: 7727
Loss: 29.909157850740158
l2 norm of gradients: 0.19263768208826063
l2 norm of weights: 5.705751945354566

Iteration Number: 7728
Loss: 29.907297244618317
l2 norm of gradients: 0.19262112019438105
l2 norm of weights: 5.705682524498508

Iteration Number: 7729
Loss: 29.905436958421035
l2 norm of gradients: 0.19260456172935775
l2 norm of weights: 5.705613108990402

```

-----
Iteration Number: 7730
Loss: 29.903576992057012
l2 norm of gradients: 0.19258800669180826
l2 norm of weights: 5.705543698829466
-----
Iteration Number: 7731
Loss: 29.901717345425315
l2 norm of gradients: 0.19257145508035078
l2 norm of weights: 5.705474294014913
-----
Iteration Number: 7732
Loss: 29.899858018443464
l2 norm of gradients: 0.19255490689360433
l2 norm of weights: 5.70540489454596
-----
Iteration Number: 7733
Loss: 29.897999011007855
l2 norm of gradients: 0.19253836213018877
l2 norm of weights: 5.705335500421823
-----
Iteration Number: 7734
Loss: 29.896140323029687
l2 norm of gradients: 0.19252182078872454
l2 norm of weights: 5.705266111641717
-----
Iteration Number: 7735
Loss: 29.89428195440907
l2 norm of gradients: 0.19250528286783322
l2 norm of weights: 5.705196728204859
-----
Iteration Number: 7736
Loss: 29.892423905058713
l2 norm of gradients: 0.1924887483661369
l2 norm of weights: 5.705127350110465
-----
Iteration Number: 7737
Loss: 29.890566174881794
l2 norm of gradients: 0.19247221728225858
l2 norm of weights: 5.705057977357753
-----
Iteration Number: 7738
Loss: 29.88870876378195
l2 norm of gradients: 0.19245568961482198
l2 norm of weights: 5.704988609945938
-----
Iteration Number: 7739
Loss: 29.886851671672595

```

```

12 norm of gradients: 0.1924391653624517
12 norm of weights: 5.70491924787424
-----
Iteration Number: 7740
Loss: 29.88499489845313
12 norm of gradients: 0.19242264452377306
12 norm of weights: 5.704849891141873
-----
Iteration Number: 7741
Loss: 29.883138444031502
12 norm of gradients: 0.19240612709741212
12 norm of weights: 5.7047805397480555
-----
Iteration Number: 7742
Loss: 29.881282308316067
12 norm of gradients: 0.19238961308199584
12 norm of weights: 5.704711193692007
-----
Iteration Number: 7743
Loss: 29.87942649121337
12 norm of gradients: 0.19237310247615197
12 norm of weights: 5.7046418529729435
-----
Iteration Number: 7744
Loss: 29.877570992625767
12 norm of gradients: 0.19235659527850893
12 norm of weights: 5.704572517590084
-----
Iteration Number: 7745
Loss: 29.87571581246152
12 norm of gradients: 0.1923400914876959
12 norm of weights: 5.704503187542646
-----
Iteration Number: 7746
Loss: 29.873860950631435
12 norm of gradients: 0.19232359110234312
12 norm of weights: 5.704433862829849
-----
Iteration Number: 7747
Loss: 29.87200640703674
12 norm of gradients: 0.1923070941210813
12 norm of weights: 5.704364543450913
-----
Iteration Number: 7748
Loss: 29.870152181585343
12 norm of gradients: 0.19229060054254202
12 norm of weights: 5.704295229405055
-----

```

Iteration Number: 7749
Loss: 29.868298274184674
l2 norm of gradients: 0.19227411036535777
l2 norm of weights: 5.704225920691494

Iteration Number: 7750
Loss: 29.866444684743126
l2 norm of gradients: 0.1922576235881617
l2 norm of weights: 5.70415661730945

Iteration Number: 7751
Loss: 29.864591413164923
l2 norm of gradients: 0.19224114020958777
l2 norm of weights: 5.704087319258144

Iteration Number: 7752
Loss: 29.86273845935605
l2 norm of gradients: 0.19222466022827073
l2 norm of weights: 5.704018026536793

Iteration Number: 7753
Loss: 29.860885823226848
l2 norm of gradients: 0.19220818364284606
l2 norm of weights: 5.70394873914462

Iteration Number: 7754
Loss: 29.85903350467901
l2 norm of gradients: 0.1921917104519501
l2 norm of weights: 5.703879457080844

Iteration Number: 7755
Loss: 29.857181503624865
l2 norm of gradients: 0.19217524065422004
l2 norm of weights: 5.703810180344685

Iteration Number: 7756
Loss: 29.855329819969576
l2 norm of gradients: 0.19215877424829353
l2 norm of weights: 5.703740908935362

Iteration Number: 7757
Loss: 29.85347845361882
l2 norm of gradients: 0.19214231123280934
l2 norm of weights: 5.703671642852101

Iteration Number: 7758
Loss: 29.851627404477323
l2 norm of gradients: 0.19212585160640697

```

12 norm of weights: 5.703602382094118
-----
Iteration Number: 7759
Loss: 29.849776672459715
12 norm of gradients: 0.19210939536772648
12 norm of weights: 5.703533126660637
-----
Iteration Number: 7760
Loss: 29.847926257467858
12 norm of gradients: 0.19209294251540882
12 norm of weights: 5.703463876550878
-----
Iteration Number: 7761
Loss: 29.84607615940597
12 norm of gradients: 0.19207649304809593
12 norm of weights: 5.703394631764064
-----
Iteration Number: 7762
Loss: 29.844226378190314
12 norm of gradients: 0.19206004696443021
12 norm of weights: 5.703325392299416
-----
Iteration Number: 7763
Loss: 29.842376913720184
12 norm of gradients: 0.19204360426305492
12 norm of weights: 5.703256158156156
-----
Iteration Number: 7764
Loss: 29.840527765903968
12 norm of gradients: 0.1920271649426142
12 norm of weights: 5.703186929333508
-----
Iteration Number: 7765
Loss: 29.838678934652208
12 norm of gradients: 0.19201072900175298
12 norm of weights: 5.703117705830691
-----
Iteration Number: 7766
Loss: 29.836830419873227
12 norm of gradients: 0.19199429643911678
12 norm of weights: 5.703048487646931
-----
Iteration Number: 7767
Loss: 29.83498222146888
12 norm of gradients: 0.19197786725335206
12 norm of weights: 5.702979274781449
-----
Iteration Number: 7768

```


Loss: 29.833134339349357
l2 norm of gradients: 0.19196144144310595
l2 norm of weights: 5.7029100672334705

Iteration Number: 7769
Loss: 29.831286773422324
l2 norm of gradients: 0.19194501900702646
l2 norm of weights: 5.702840865002217

Iteration Number: 7770
Loss: 29.82943952359632
l2 norm of gradients: 0.1919285999437622
l2 norm of weights: 5.702771668086911

Iteration Number: 7771
Loss: 29.827592589777215
l2 norm of gradients: 0.19191218425196277
l2 norm of weights: 5.702702476486778

Iteration Number: 7772
Loss: 29.82574597187786
l2 norm of gradients: 0.19189577193027843
l2 norm of weights: 5.702633290201042

Iteration Number: 7773
Loss: 29.823899669794294
l2 norm of gradients: 0.19187936297736016
l2 norm of weights: 5.702564109228927

Iteration Number: 7774
Loss: 29.82205368344571
l2 norm of gradients: 0.19186295739185982
l2 norm of weights: 5.702494933569657

Iteration Number: 7775
Loss: 29.820208012735446
l2 norm of gradients: 0.1918465551724299
l2 norm of weights: 5.702425763222457

Iteration Number: 7776
Loss: 29.81836265757285
l2 norm of gradients: 0.19183015631772388
l2 norm of weights: 5.70235659818655

Iteration Number: 7777
Loss: 29.816517617861905
l2 norm of gradients: 0.19181376082639573
l2 norm of weights: 5.7022874384611635

```
-----
Iteration Number: 7778
Loss: 29.814672893513553
l2 norm of gradients: 0.19179736869710035
l2 norm of weights: 5.70221828404552
-----
Iteration Number: 7779
Loss: 29.81282848443645
l2 norm of gradients: 0.19178097992849352
l2 norm of weights: 5.702149134938848
-----
Iteration Number: 7780
Loss: 29.810984390536728
l2 norm of gradients: 0.19176459451923153
l2 norm of weights: 5.702079991140371
-----
Iteration Number: 7781
Loss: 29.8091406117226
l2 norm of gradients: 0.19174821246797158
l2 norm of weights: 5.702010852649315
-----
Iteration Number: 7782
Loss: 29.807297147903075
l2 norm of gradients: 0.19173183377337166
l2 norm of weights: 5.701941719464907
-----
Iteration Number: 7783
Loss: 29.80545399898651
l2 norm of gradients: 0.19171545843409046
l2 norm of weights: 5.701872591586373
-----
Iteration Number: 7784
Loss: 29.803611164878443
l2 norm of gradients: 0.19169908644878747
l2 norm of weights: 5.701803469012939
-----
Iteration Number: 7785
Loss: 29.801768645490057
l2 norm of gradients: 0.1916827178161229
l2 norm of weights: 5.701734351743831
-----
Iteration Number: 7786
Loss: 29.79992644073037
l2 norm of gradients: 0.19166635253475778
l2 norm of weights: 5.7016652397782766
-----
Iteration Number: 7787
Loss: 29.798084550502505
```

```

12 norm of gradients: 0.19164999060335386
12 norm of weights: 5.701596133115502
-----
Iteration Number: 7788
Loss: 29.79624297471875
12 norm of gradients: 0.19163363202057376
12 norm of weights: 5.701527031754735
-----
Iteration Number: 7789
Loss: 29.79440171328765
12 norm of gradients: 0.19161727678508067
12 norm of weights: 5.701457935695205
-----
Iteration Number: 7790
Loss: 29.79256076611572
12 norm of gradients: 0.1916009248955387
12 norm of weights: 5.701388844936137
-----
Iteration Number: 7791
Loss: 29.79072013311464
12 norm of gradients: 0.19158457635061268
12 norm of weights: 5.701319759476759
-----
Iteration Number: 7792
Loss: 29.788879814190125
12 norm of gradients: 0.1915682311489682
12 norm of weights: 5.7012506793163
-----
Iteration Number: 7793
Loss: 29.787039809251745
12 norm of gradients: 0.19155188928927158
12 norm of weights: 5.701181604453987
-----
Iteration Number: 7794
Loss: 29.785200118202727
12 norm of gradients: 0.19153555077018994
12 norm of weights: 5.701112534889051
-----
Iteration Number: 7795
Loss: 29.78336074096157
12 norm of gradients: 0.19151921559039115
12 norm of weights: 5.701043470620718
-----
Iteration Number: 7796
Loss: 29.781521677429232
12 norm of gradients: 0.1915028837485438
12 norm of weights: 5.700974411648218
-----

```

Iteration Number: 7797
Loss: 29.779682927519346
l2 norm of gradients: 0.19148655524331734
l2 norm of weights: 5.700905357970781

Iteration Number: 7798
Loss: 29.777844491138826
l2 norm of gradients: 0.19147023007338185
l2 norm of weights: 5.7008363095876335

Iteration Number: 7799
Loss: 29.77600636819342
l2 norm of gradients: 0.19145390823740824
l2 norm of weights: 5.700767266498008

Iteration Number: 7800
Loss: 29.774168558595335
l2 norm of gradients: 0.1914375897340682
l2 norm of weights: 5.700698228701133

Iteration Number: 7801
Loss: 29.77233106225244
l2 norm of gradients: 0.19142127456203414
l2 norm of weights: 5.700629196196237

Iteration Number: 7802
Loss: 29.770493879074987
l2 norm of gradients: 0.19140496271997917
l2 norm of weights: 5.700560168982553

Iteration Number: 7803
Loss: 29.7686570089705
l2 norm of gradients: 0.1913886542065773
l2 norm of weights: 5.700491147059309

Iteration Number: 7804
Loss: 29.766820451846478
l2 norm of gradients: 0.1913723490205032
l2 norm of weights: 5.700422130425736

Iteration Number: 7805
Loss: 29.76498420761642
l2 norm of gradients: 0.1913560471604322
l2 norm of weights: 5.700353119081065

Iteration Number: 7806
Loss: 29.76314827618325
l2 norm of gradients: 0.19133974862504066

```

12 norm of weights: 5.7002841130245265
-----
Iteration Number: 7807
Loss: 29.761312657463804
12 norm of gradients: 0.19132345341300544
12 norm of weights: 5.700215112255352
-----
Iteration Number: 7808
Loss: 29.759477351361106
12 norm of gradients: 0.1913071615230042
12 norm of weights: 5.700146116772772
-----
Iteration Number: 7809
Loss: 29.757642357785517
12 norm of gradients: 0.1912908729537155
12 norm of weights: 5.700077126576019
-----
Iteration Number: 7810
Loss: 29.75580767664836
12 norm of gradients: 0.19127458770381847
12 norm of weights: 5.700008141664325
-----
Iteration Number: 7811
Loss: 29.753973307858313
12 norm of gradients: 0.19125830577199307
12 norm of weights: 5.699939162036921
-----
Iteration Number: 7812
Loss: 29.752139251321772
12 norm of gradients: 0.19124202715692007
12 norm of weights: 5.699870187693039
-----
Iteration Number: 7813
Loss: 29.750305506951058
12 norm of gradients: 0.19122575185728088
12 norm of weights: 5.699801218631912
-----
Iteration Number: 7814
Loss: 29.74847207465434
12 norm of gradients: 0.1912094798717577
12 norm of weights: 5.699732254852772
-----
Iteration Number: 7815
Loss: 29.74663895434351
12 norm of gradients: 0.1911932111990335
12 norm of weights: 5.699663296354851
-----
Iteration Number: 7816

```

Loss: 29.744806145925487
l2 norm of gradients: 0.1911769458377921
l2 norm of weights: 5.699594343137384

Iteration Number: 7817
Loss: 29.742973649309498
l2 norm of gradients: 0.19116068378671786
l2 norm of weights: 5.699525395199602

Iteration Number: 7818
Loss: 29.741141464405615
l2 norm of gradients: 0.19114442504449594
l2 norm of weights: 5.6994564525407405

Iteration Number: 7819
Loss: 29.73930959112573
l2 norm of gradients: 0.19112816960981246
l2 norm of weights: 5.699387515160031

Iteration Number: 7820
Loss: 29.737478029376824
l2 norm of gradients: 0.19111191748135398
l2 norm of weights: 5.699318583056708

Iteration Number: 7821
Loss: 29.735646779067462
l2 norm of gradients: 0.19109566865780805
l2 norm of weights: 5.699249656230005

Iteration Number: 7822
Loss: 29.733815840114744
l2 norm of gradients: 0.1910794231378628
l2 norm of weights: 5.699180734679156

Iteration Number: 7823
Loss: 29.731985212416117
l2 norm of gradients: 0.19106318092020724
l2 norm of weights: 5.699111818403397

Iteration Number: 7824
Loss: 29.73015489589309
l2 norm of gradients: 0.19104694200353112
l2 norm of weights: 5.699042907401961

Iteration Number: 7825
Loss: 29.72832489044824
l2 norm of gradients: 0.19103070638652472
l2 norm of weights: 5.698974001674083

```

-----
Iteration Number: 7826
Loss: 29.726495195997792
l2 norm of gradients: 0.19101447406787941
l2 norm of weights: 5.6989051012189975
-----
Iteration Number: 7827
Loss: 29.724665812443348
l2 norm of gradients: 0.19099824504628699
l2 norm of weights: 5.69883620603594
-----
Iteration Number: 7828
Loss: 29.722836739702075
l2 norm of gradients: 0.19098201932044018
l2 norm of weights: 5.6987673161241466
-----
Iteration Number: 7829
Loss: 29.721007977681566
l2 norm of gradients: 0.19096579688903242
l2 norm of weights: 5.6986984314828515
-----
Iteration Number: 7830
Loss: 29.71917952629098
l2 norm of gradients: 0.19094957775075785
l2 norm of weights: 5.698629552111292
-----
Iteration Number: 7831
Loss: 29.71735138544256
l2 norm of gradients: 0.1909333619043114
l2 norm of weights: 5.698560678008702
-----
Iteration Number: 7832
Loss: 29.71552355504321
l2 norm of gradients: 0.19091714934838874
l2 norm of weights: 5.69849180917432
-----
Iteration Number: 7833
Loss: 29.713696035006826
l2 norm of gradients: 0.19090094008168615
l2 norm of weights: 5.698422945607382
-----
Iteration Number: 7834
Loss: 29.711868825238607
l2 norm of gradients: 0.1908847341029009
l2 norm of weights: 5.698354087307121
-----
Iteration Number: 7835
Loss: 29.710041925657272

```

```

12 norm of gradients: 0.1908685314107309
12 norm of weights: 5.6982852342727774
-----
Iteration Number: 7836
Loss: 29.708215336164887
12 norm of gradients: 0.19085233200387458
12 norm of weights: 5.698216386503587
-----
Iteration Number: 7837
Loss: 29.706389056674855
12 norm of gradients: 0.19083613588103143
12 norm of weights: 5.698147543998787
-----
Iteration Number: 7838
Loss: 29.704563087096872
12 norm of gradients: 0.19081994304090152
12 norm of weights: 5.698078706757614
-----
Iteration Number: 7839
Loss: 29.70273742734403
12 norm of gradients: 0.19080375348218567
12 norm of weights: 5.6980098747793075
-----
Iteration Number: 7840
Loss: 29.70091207732183
12 norm of gradients: 0.1907875672035855
12 norm of weights: 5.697941048063104
-----
Iteration Number: 7841
Loss: 29.699087036949038
12 norm of gradients: 0.1907713842038032
12 norm of weights: 5.697872226608241
-----
Iteration Number: 7842
Loss: 29.697262306126852
12 norm of gradients: 0.19075520448154204
12 norm of weights: 5.697803410413956
-----
Iteration Number: 7843
Loss: 29.69543788477253
12 norm of gradients: 0.1907390280355056
12 norm of weights: 5.69773459947949
-----
Iteration Number: 7844
Loss: 29.69361377279107
12 norm of gradients: 0.19072285486439855
12 norm of weights: 5.697665793804078
-----

```


Iteration Number: 7845
Loss: 29.691789970099396
l2 norm of gradients: 0.19070668496692605
l2 norm of weights: 5.697596993386962

Iteration Number: 7846
Loss: 29.68996647660359
l2 norm of gradients: 0.19069051834179418
l2 norm of weights: 5.697528198227379

Iteration Number: 7847
Loss: 29.688143292216072
l2 norm of gradients: 0.19067435498770963
l2 norm of weights: 5.697459408324568

Iteration Number: 7848
Loss: 29.68632041684782
l2 norm of gradients: 0.1906581949033799
l2 norm of weights: 5.697390623677769

Iteration Number: 7849
Loss: 29.684497850410697
l2 norm of gradients: 0.19064203808751318
l2 norm of weights: 5.697321844286222

Iteration Number: 7850
Loss: 29.682675592813727
l2 norm of gradients: 0.19062588453881837
l2 norm of weights: 5.697253070149166

Iteration Number: 7851
Loss: 29.680853643970217
l2 norm of gradients: 0.19060973425600516
l2 norm of weights: 5.697184301265841

Iteration Number: 7852
Loss: 29.67903200378611
l2 norm of gradients: 0.19059358723778405
l2 norm of weights: 5.697115537635487

Iteration Number: 7853
Loss: 29.677210672175338
l2 norm of gradients: 0.19057744348286607
l2 norm of weights: 5.697046779257344

Iteration Number: 7854
Loss: 29.675389649050107
l2 norm of gradients: 0.1905613029899632

```

12 norm of weights: 5.696978026130654
-----
Iteration Number: 7855
Loss: 29.673568934325033
12 norm of gradients: 0.19054516575778796
12 norm of weights: 5.696909278254656
-----
Iteration Number: 7856
Loss: 29.6717485279039
12 norm of gradients: 0.19052903178505365
12 norm of weights: 5.696840535628592
-----
Iteration Number: 7857
Loss: 29.669928429699166
12 norm of gradients: 0.19051290107047447
12 norm of weights: 5.696771798251703
-----
Iteration Number: 7858
Loss: 29.668108639624705
12 norm of gradients: 0.19049677361276512
12 norm of weights: 5.696703066123228
-----
Iteration Number: 7859
Loss: 29.66628915759188
12 norm of gradients: 0.19048064941064122
12 norm of weights: 5.696634339242412
-----
Iteration Number: 7860
Loss: 29.664469983512802
12 norm of gradients: 0.1904645284628189
12 norm of weights: 5.696565617608496
-----
Iteration Number: 7861
Loss: 29.66265111729536
12 norm of gradients: 0.19044841076801527
12 norm of weights: 5.69649690122072
-----
Iteration Number: 7862
Loss: 29.66083255885178
12 norm of gradients: 0.19043229632494804
12 norm of weights: 5.6964281900783265
-----
Iteration Number: 7863
Loss: 29.659014308093997
12 norm of gradients: 0.19041618513233557
12 norm of weights: 5.696359484180559
-----
Iteration Number: 7864

```

Loss: 29.65719636493589
l2 norm of gradients: 0.19040007718889712
l2 norm of weights: 5.696290783526659

Iteration Number: 7865
Loss: 29.655378729285317
l2 norm of gradients: 0.19038397249335257
l2 norm of weights: 5.696222088115871

Iteration Number: 7866
Loss: 29.65356140105614
l2 norm of gradients: 0.19036787104442257
l2 norm of weights: 5.696153397947436

Iteration Number: 7867
Loss: 29.65174438015713
l2 norm of gradients: 0.19035177284082844
l2 norm of weights: 5.696084713020597

Iteration Number: 7868
Loss: 29.649927666505377
l2 norm of gradients: 0.19033567788129227
l2 norm of weights: 5.6960160333345975

Iteration Number: 7869
Loss: 29.648111260004992
l2 norm of gradients: 0.1903195861645369
l2 norm of weights: 5.695947358888682

Iteration Number: 7870
Loss: 29.646295160574844
l2 norm of gradients: 0.19030349768928584
l2 norm of weights: 5.695878689682092

Iteration Number: 7871
Loss: 29.644479368121697
l2 norm of gradients: 0.19028741245426342
l2 norm of weights: 5.695810025714074

Iteration Number: 7872
Loss: 29.642663882560527
l2 norm of gradients: 0.19027133045819455
l2 norm of weights: 5.695741366983872

Iteration Number: 7873
Loss: 29.64084870380295
l2 norm of gradients: 0.19025525169980498
l2 norm of weights: 5.695672713490727

```

-----
Iteration Number: 7874
Loss: 29.63903383175736
l2 norm of gradients: 0.1902391761778211
l2 norm of weights: 5.695604065233886
-----
Iteration Number: 7875
Loss: 29.63721926633878
l2 norm of gradients: 0.19022310389097016
l2 norm of weights: 5.695535422212593
-----
Iteration Number: 7876
Loss: 29.6354050074618
l2 norm of gradients: 0.19020703483798
l2 norm of weights: 5.695466784426094
-----
Iteration Number: 7877
Loss: 29.63359105502942
l2 norm of gradients: 0.19019096901757918
l2 norm of weights: 5.695398151873633
-----
Iteration Number: 7878
Loss: 29.631777408962876
l2 norm of gradients: 0.19017490642849708
l2 norm of weights: 5.695329524554455
-----
Iteration Number: 7879
Loss: 29.62996406917021
l2 norm of gradients: 0.19015884706946376
l2 norm of weights: 5.695260902467806
-----
Iteration Number: 7880
Loss: 29.62815103556335
l2 norm of gradients: 0.19014279093920994
l2 norm of weights: 5.695192285612931
-----
Iteration Number: 7881
Loss: 29.626338308051945
l2 norm of gradients: 0.19012673803646715
l2 norm of weights: 5.695123673989076
-----
Iteration Number: 7882
Loss: 29.62452588655644
l2 norm of gradients: 0.19011068835996758
l2 norm of weights: 5.6950550675954865
-----
Iteration Number: 7883
Loss: 29.622713770981402

```

```

12 norm of gradients: 0.19009464190844422
12 norm of weights: 5.694986466431411
-----
Iteration Number: 7884
Loss: 29.62090196124201
12 norm of gradients: 0.19007859868063062
12 norm of weights: 5.694917870496094
-----
Iteration Number: 7885
Loss: 29.6190904572502
12 norm of gradients: 0.19006255867526126
12 norm of weights: 5.694849279788781
-----
Iteration Number: 7886
Loss: 29.617279258919
12 norm of gradients: 0.19004652189107116
12 norm of weights: 5.694780694308721
-----
Iteration Number: 7887
Loss: 29.61546836615802
12 norm of gradients: 0.19003048832679617
12 norm of weights: 5.694712114055161
-----
Iteration Number: 7888
Loss: 29.61365777888348
12 norm of gradients: 0.19001445798117278
12 norm of weights: 5.694643539027347
-----
Iteration Number: 7889
Loss: 29.611847497005446
12 norm of gradients: 0.1899984308529383
12 norm of weights: 5.694574969224526
-----
Iteration Number: 7890
Loss: 29.610037520437864
12 norm of gradients: 0.18998240694083063
12 norm of weights: 5.6945064046459475
-----
Iteration Number: 7891
Loss: 29.60822784909177
12 norm of gradients: 0.18996638624358847
12 norm of weights: 5.6944378452908575
-----
Iteration Number: 7892
Loss: 29.606418482880844
12 norm of gradients: 0.1899503687599512
12 norm of weights: 5.694369291158504
-----

```

Iteration Number: 7893
Loss: 29.60460942171548
l2 norm of gradients: 0.18993435448865892
l2 norm of weights: 5.694300742248137

Iteration Number: 7894
Loss: 29.602800665509687
l2 norm of gradients: 0.18991834342845254
l2 norm of weights: 5.6942321985590025

Iteration Number: 7895
Loss: 29.600992214179605
l2 norm of gradients: 0.18990233557807348
l2 norm of weights: 5.69416366009035

Iteration Number: 7896
Loss: 29.59918406763331
l2 norm of gradients: 0.18988633093626414
l2 norm of weights: 5.69409512684143

Iteration Number: 7897
Loss: 29.597376225787055
l2 norm of gradients: 0.1898703295017674
l2 norm of weights: 5.694026598811488

Iteration Number: 7898
Loss: 29.595568688548827
l2 norm of gradients: 0.1898543312733269
l2 norm of weights: 5.6939580759997765

Iteration Number: 7899
Loss: 29.593761455835693
l2 norm of gradients: 0.1898383362496871
l2 norm of weights: 5.693889558405543

Iteration Number: 7900
Loss: 29.591954527560073
l2 norm of gradients: 0.18982234442959312
l2 norm of weights: 5.693821046028036

Iteration Number: 7901
Loss: 29.590147903633053
l2 norm of gradients: 0.18980635581179078
l2 norm of weights: 5.693752538866508

Iteration Number: 7902
Loss: 29.588341583969758
l2 norm of gradients: 0.18979037039502664

```

12 norm of weights: 5.693684036920207
-----
Iteration Number: 7903
Loss: 29.58653556848056
12 norm of gradients: 0.18977438817804787
12 norm of weights: 5.693615540188383
-----
Iteration Number: 7904
Loss: 29.584729857084888
12 norm of gradients: 0.18975840915960243
12 norm of weights: 5.693547048670288
-----
Iteration Number: 7905
Loss: 29.582924449684654
12 norm of gradients: 0.18974243333843907
12 norm of weights: 5.693478562365171
-----
Iteration Number: 7906
Loss: 29.581119346204623
12 norm of gradients: 0.18972646071330718
12 norm of weights: 5.693410081272283
-----
Iteration Number: 7907
Loss: 29.57931454654897
12 norm of gradients: 0.18971049128295675
12 norm of weights: 5.693341605390874
-----
Iteration Number: 7908
Loss: 29.577510050636377
12 norm of gradients: 0.18969452504613862
12 norm of weights: 5.693273134720197
-----
Iteration Number: 7909
Loss: 29.57570585837994
12 norm of gradients: 0.1896785620016043
12 norm of weights: 5.693204669259501
-----
Iteration Number: 7910
Loss: 29.573901969686553
12 norm of gradients: 0.189662602148106
12 norm of weights: 5.69313620900804
-----
Iteration Number: 7911
Loss: 29.572098384477485
12 norm of gradients: 0.18964664548439672
12 norm of weights: 5.693067753965065
-----
Iteration Number: 7912

```

Loss: 29.57029510266516
l2 norm of gradients: 0.18963069200923002
l2 norm of weights: 5.692999304129825

Iteration Number: 7913
Loss: 29.568492124157274
l2 norm of gradients: 0.18961474172136028
l2 norm of weights: 5.692930859501576

Iteration Number: 7914
Loss: 29.566689448875167
l2 norm of gradients: 0.18959879461954252
l2 norm of weights: 5.692862420079567

Iteration Number: 7915
Loss: 29.56488707672333
l2 norm of gradients: 0.18958285070253247
l2 norm of weights: 5.692793985863053

Iteration Number: 7916
Loss: 29.563085007623066
l2 norm of gradients: 0.18956690996908668
l2 norm of weights: 5.692725556851285

Iteration Number: 7917
Loss: 29.56128324148362
l2 norm of gradients: 0.1895509724179622
l2 norm of weights: 5.692657133043516

Iteration Number: 7918
Loss: 29.559481778219237
l2 norm of gradients: 0.189535038047917
l2 norm of weights: 5.692588714438999

Iteration Number: 7919
Loss: 29.55768061774686
l2 norm of gradients: 0.18951910685770962
l2 norm of weights: 5.692520301036989

Iteration Number: 7920
Loss: 29.55587975997415
l2 norm of gradients: 0.18950317884609935
l2 norm of weights: 5.692451892836737

Iteration Number: 7921
Loss: 29.554079204819406
l2 norm of gradients: 0.18948725401184618
l2 norm of weights: 5.692383489837497


```

-----
Iteration Number: 7922
Loss: 29.552278952192957
l2 norm of gradients: 0.18947133235371086
l2 norm of weights: 5.692315092038524
-----
Iteration Number: 7923
Loss: 29.550479002013617
l2 norm of gradients: 0.18945541387045464
l2 norm of weights: 5.692246699439071
-----
Iteration Number: 7924
Loss: 29.548679354194736
l2 norm of gradients: 0.1894394985608398
l2 norm of weights: 5.692178312038392
-----
Iteration Number: 7925
Loss: 29.54688000864264
l2 norm of gradients: 0.18942358642362894
l2 norm of weights: 5.692109929835741
-----
Iteration Number: 7926
Loss: 29.545080965279098
l2 norm of gradients: 0.18940767745758574
l2 norm of weights: 5.692041552830374
-----
Iteration Number: 7927
Loss: 29.543282224011666
l2 norm of gradients: 0.18939177166147436
l2 norm of weights: 5.691973181021544
-----
Iteration Number: 7928
Loss: 29.54148378476295
l2 norm of gradients: 0.18937586903405956
l2 norm of weights: 5.691904814408509
-----
Iteration Number: 7929
Loss: 29.53968564744148
l2 norm of gradients: 0.1893599695741072
l2 norm of weights: 5.69183645299052
-----
Iteration Number: 7930
Loss: 29.537887811960623
l2 norm of gradients: 0.1893440732803834
l2 norm of weights: 5.691768096766835
-----
Iteration Number: 7931
Loss: 29.53609027823575

```

```

12 norm of gradients: 0.18932818015165526
12 norm of weights: 5.691699745736709
-----
Iteration Number: 7932
Loss: 29.534293046181112
12 norm of gradients: 0.18931229018669044
12 norm of weights: 5.691631399899396
-----
Iteration Number: 7933
Loss: 29.532496115711147
12 norm of gradients: 0.18929640338425735
12 norm of weights: 5.691563059254155
-----
Iteration Number: 7934
Loss: 29.53069948673642
12 norm of gradients: 0.1892805197431251
12 norm of weights: 5.691494723800239
-----
Iteration Number: 7935
Loss: 29.528903159179485
12 norm of gradients: 0.18926463926206355
12 norm of weights: 5.691426393536906
-----
Iteration Number: 7936
Loss: 29.527107132947386
12 norm of gradients: 0.18924876193984316
12 norm of weights: 5.691358068463413
-----
Iteration Number: 7937
Loss: 29.525311407957382
12 norm of gradients: 0.18923288777523511
12 norm of weights: 5.691289748579015
-----
Iteration Number: 7938
Loss: 29.523515984121744
12 norm of gradients: 0.1892170167670114
12 norm of weights: 5.691221433882969
-----
Iteration Number: 7939
Loss: 29.521720861357302
12 norm of gradients: 0.1892011489139445
12 norm of weights: 5.691153124374534
-----
Iteration Number: 7940
Loss: 29.51992603957842
12 norm of gradients: 0.18918528421480776
12 norm of weights: 5.691084820052965
-----

```

Iteration Number: 7941
Loss: 29.518131518696297
l2 norm of gradients: 0.1891694226683752
l2 norm of weights: 5.691016520917519

Iteration Number: 7942
Loss: 29.516337298630987
l2 norm of gradients: 0.18915356427342145
l2 norm of weights: 5.690948226967456

Iteration Number: 7943
Loss: 29.51454337929216
l2 norm of gradients: 0.1891377090287219
l2 norm of weights: 5.690879938202032

Iteration Number: 7944
Loss: 29.512749760598172
l2 norm of gradients: 0.18912185693305267
l2 norm of weights: 5.690811654620506

Iteration Number: 7945
Loss: 29.510956442457843
l2 norm of gradients: 0.18910600798519048
l2 norm of weights: 5.690743376222135

Iteration Number: 7946
Loss: 29.50916342479389
l2 norm of gradients: 0.18909016218391284
l2 norm of weights: 5.690675103006178

Iteration Number: 7947
Loss: 29.507370707516355
l2 norm of gradients: 0.18907431952799794
l2 norm of weights: 5.690606834971895

Iteration Number: 7948
Loss: 29.50557829053712
l2 norm of gradients: 0.18905848001622447
l2 norm of weights: 5.690538572118542

Iteration Number: 7949
Loss: 29.50378617377651
l2 norm of gradients: 0.1890426436473722
l2 norm of weights: 5.690470314445379

Iteration Number: 7950
Loss: 29.501994357149194
l2 norm of gradients: 0.18902681042022118

```

12 norm of weights: 5.690402061951666
-----
Iteration Number: 7951
Loss: 29.500202840564423
12 norm of gradients: 0.18901098033355246
12 norm of weights: 5.690333814636662
-----
Iteration Number: 7952
Loss: 29.49841162394337
12 norm of gradients: 0.18899515338614759
12 norm of weights: 5.690265572499625
-----
Iteration Number: 7953
Loss: 29.496620707198332
12 norm of gradients: 0.18897932957678895
12 norm of weights: 5.690197335539817
-----
Iteration Number: 7954
Loss: 29.49483009024372
12 norm of gradients: 0.18896350890425953
12 norm of weights: 5.690129103756495
-----
Iteration Number: 7955
Loss: 29.493039772995598
12 norm of gradients: 0.18894769136734296
12 norm of weights: 5.690060877148921
-----
Iteration Number: 7956
Loss: 29.491249755366674
12 norm of gradients: 0.18893187696482375
12 norm of weights: 5.689992655716356
-----
Iteration Number: 7957
Loss: 29.489460037278626
12 norm of gradients: 0.18891606569548686
12 norm of weights: 5.689924439458058
-----
Iteration Number: 7958
Loss: 29.487670618637758
12 norm of gradients: 0.18890025755811807
12 norm of weights: 5.68985622837329
-----
Iteration Number: 7959
Loss: 29.485881499363977
12 norm of gradients: 0.188884452551504
12 norm of weights: 5.689788022461312
-----
Iteration Number: 7960

```

```

Loss: 29.484092679374694
l2 norm of gradients: 0.18886865067443154
l2 norm of weights: 5.689719821721385
-----
Iteration Number: 7961
Loss: 29.48230415858036
l2 norm of gradients: 0.18885285192568865
l2 norm of weights: 5.689651626152769
-----
Iteration Number: 7962
Loss: 29.480515936898307
l2 norm of gradients: 0.18883705630406392
l2 norm of weights: 5.689583435754726
-----
Iteration Number: 7963
Loss: 29.478728014244716
l2 norm of gradients: 0.18882126380834646
l2 norm of weights: 5.68951525052652
-----
Iteration Number: 7964
Loss: 29.476940390532935
l2 norm of gradients: 0.18880547443732618
l2 norm of weights: 5.689447070467408
-----
Iteration Number: 7965
Loss: 29.475153065682587
l2 norm of gradients: 0.18878968818979375
l2 norm of weights: 5.689378895576657
-----
Iteration Number: 7966
Loss: 29.47336603960284
l2 norm of gradients: 0.18877390506454025
l2 norm of weights: 5.689310725853526
-----
Iteration Number: 7967
Loss: 29.47157931221574
l2 norm of gradients: 0.18875812506035783
l2 norm of weights: 5.689242561297277
-----
Iteration Number: 7968
Loss: 29.46979288343019
l2 norm of gradients: 0.18874234817603908
l2 norm of weights: 5.689174401907175
-----
Iteration Number: 7969
Loss: 29.46800675316862
l2 norm of gradients: 0.18872657441037724
l2 norm of weights: 5.689106247682481

```

```

-----
Iteration Number: 7970
Loss: 29.466220921339954
l2 norm of gradients: 0.18871080376216642
l2 norm of weights: 5.689038098622458
-----
Iteration Number: 7971
Loss: 29.464435387863844
l2 norm of gradients: 0.18869503623020129
l2 norm of weights: 5.688969954726369
-----
Iteration Number: 7972
Loss: 29.462650152656465
l2 norm of gradients: 0.18867927181327723
l2 norm of weights: 5.6889018159934786
-----
Iteration Number: 7973
Loss: 29.46086521562981
l2 norm of gradients: 0.1886635105101903
l2 norm of weights: 5.6888336824230485
-----
Iteration Number: 7974
Loss: 29.459080576705393
l2 norm of gradients: 0.18864775231973727
l2 norm of weights: 5.688765554014343
-----
Iteration Number: 7975
Loss: 29.457296235793542
l2 norm of gradients: 0.1886319972407155
l2 norm of weights: 5.688697430766625
-----
Iteration Number: 7976
Loss: 29.455512192812716
l2 norm of gradients: 0.18861624527192317
l2 norm of weights: 5.6886293126791605
-----
Iteration Number: 7977
Loss: 29.453728447678913
l2 norm of gradients: 0.18860049641215906
l2 norm of weights: 5.688561199751213
-----
Iteration Number: 7978
Loss: 29.451945000305624
l2 norm of gradients: 0.1885847506602227
l2 norm of weights: 5.6884930919820444
-----
Iteration Number: 7979
Loss: 29.45016185061237

```

```

12 norm of gradients: 0.1885690080149142
12 norm of weights: 5.688424989370923
-----
Iteration Number: 7980
Loss: 29.448378998511643
12 norm of gradients: 0.1885532684750343
12 norm of weights: 5.688356891917112
-----
Iteration Number: 7981
Loss: 29.446596443922402
12 norm of gradients: 0.1885375320393847
12 norm of weights: 5.688288799619876
-----
Iteration Number: 7982
Loss: 29.444814186758155
12 norm of gradients: 0.18852179870676752
12 norm of weights: 5.68822071247848
-----
Iteration Number: 7983
Loss: 29.443032226938254
12 norm of gradients: 0.18850606847598567
12 norm of weights: 5.688152630492191
-----
Iteration Number: 7984
Loss: 29.441250564374702
12 norm of gradients: 0.18849034134584267
12 norm of weights: 5.688084553660273
-----
Iteration Number: 7985
Loss: 29.439469198984465
12 norm of gradients: 0.18847461731514276
12 norm of weights: 5.688016481981991
-----
Iteration Number: 7986
Loss: 29.437688130688784
12 norm of gradients: 0.1884588963826909
12 norm of weights: 5.687948415456613
-----
Iteration Number: 7987
Loss: 29.435907359398936
12 norm of gradients: 0.1884431785472927
12 norm of weights: 5.687880354083402
-----
Iteration Number: 7988
Loss: 29.434126885029716
12 norm of gradients: 0.18842746380775435
12 norm of weights: 5.687812297861628
-----

```

Iteration Number: 7989
Loss: 29.432346707505083
l2 norm of gradients: 0.1884117521628829
l2 norm of weights: 5.687744246790555

Iteration Number: 7990
Loss: 29.430566826730757
l2 norm of gradients: 0.18839604361148593
l2 norm of weights: 5.68767620086945

Iteration Number: 7991
Loss: 29.428787242634183
l2 norm of gradients: 0.1883803381523718
l2 norm of weights: 5.68760816009758

Iteration Number: 7992
Loss: 29.42700795512236
l2 norm of gradients: 0.18836463578434942
l2 norm of weights: 5.687540124474212

Iteration Number: 7993
Loss: 29.42522896411799
l2 norm of gradients: 0.18834893650622853
l2 norm of weights: 5.687472093998613

Iteration Number: 7994
Loss: 29.423450269531635
l2 norm of gradients: 0.1883332403168194
l2 norm of weights: 5.687404068670052

Iteration Number: 7995
Loss: 29.421671871287145
l2 norm of gradients: 0.18831754721493307
l2 norm of weights: 5.6873360484877935

Iteration Number: 7996
Loss: 29.41989376929842
l2 norm of gradients: 0.18830185719938122
l2 norm of weights: 5.687268033451106

Iteration Number: 7997
Loss: 29.418115963477
l2 norm of gradients: 0.1882861702689762
l2 norm of weights: 5.687200023559261

Iteration Number: 7998
Loss: 29.41633845374704
l2 norm of gradients: 0.18827048642253114


```

12 norm of weights: 5.687132018811522
-----
Iteration Number: 7999
Loss: 29.414561240021158
12 norm of gradients: 0.18825480565885966
12 norm of weights: 5.68706401920716
-----
Iteration Number: 8000
Loss: 29.412784322216574
12 norm of gradients: 0.18823912797677614
12 norm of weights: 5.686996024745442
-----
Iteration Number: 8001
Loss: 29.411007700251155
12 norm of gradients: 0.1882234533750957
12 norm of weights: 5.6869280354256375
-----
Iteration Number: 8002
Loss: 29.409231374039603
12 norm of gradients: 0.188207781852634
12 norm of weights: 5.686860051247016
-----
Iteration Number: 8003
Loss: 29.407455343499258
12 norm of gradients: 0.18819211340820752
12 norm of weights: 5.686792072208845
-----
Iteration Number: 8004
Loss: 29.40567960854938
12 norm of gradients: 0.1881764480406333
12 norm of weights: 5.686724098310394
-----
Iteration Number: 8005
Loss: 29.403904169103257
12 norm of gradients: 0.18816078574872905
12 norm of weights: 5.686656129550934
-----
Iteration Number: 8006
Loss: 29.40212902508094
12 norm of gradients: 0.18814512653131324
12 norm of weights: 5.6865881659297335
-----
Iteration Number: 8007
Loss: 29.40035417639715
12 norm of gradients: 0.18812947038720496
12 norm of weights: 5.686520207446061
-----
Iteration Number: 8008

```

```

Loss: 29.39857962297081
12 norm of gradients: 0.18811381731522395
12 norm of weights: 5.686452254099189
-----
Iteration Number: 8009
Loss: 29.396805364717842
12 norm of gradients: 0.18809816731419063
12 norm of weights: 5.686384305888386
-----
Iteration Number: 8010
Loss: 29.395031401557077
12 norm of gradients: 0.18808252038292614
12 norm of weights: 5.6863163628129225
-----
Iteration Number: 8011
Loss: 29.393257733403576
12 norm of gradients: 0.18806687652025225
12 norm of weights: 5.686248424872069
-----
Iteration Number: 8012
Loss: 29.391484360173408
12 norm of gradients: 0.18805123572499133
12 norm of weights: 5.6861804920650965
-----
Iteration Number: 8013
Loss: 29.389711281787296
12 norm of gradients: 0.18803559799596659
12 norm of weights: 5.6861125643912755
-----
Iteration Number: 8014
Loss: 29.387938498157467
12 norm of gradients: 0.1880199633320017
12 norm of weights: 5.686044641849877
-----
Iteration Number: 8015
Loss: 29.3861660092072
12 norm of gradients: 0.18800433173192121
12 norm of weights: 5.685976724440175
-----
Iteration Number: 8016
Loss: 29.384393814851027
12 norm of gradients: 0.18798870319455016
12 norm of weights: 5.685908812161436
-----
Iteration Number: 8017
Loss: 29.38262191500649
12 norm of gradients: 0.18797307771871435
12 norm of weights: 5.685840905012935

```

```
-----  
Iteration Number: 8018  
Loss: 29.380850309588865  
l2 norm of gradients: 0.18795745530324026  
l2 norm of weights: 5.685773002993944  
-----  
Iteration Number: 8019  
Loss: 29.379078998519  
l2 norm of gradients: 0.18794183594695496  
l2 norm of weights: 5.6857051061037325  
-----  
Iteration Number: 8020  
Loss: 29.377307981711354  
l2 norm of gradients: 0.18792621964868622  
l2 norm of weights: 5.6856372143415745  
-----  
Iteration Number: 8021  
Loss: 29.37553725908513  
l2 norm of gradients: 0.18791060640726254  
l2 norm of weights: 5.685569327706742  
-----  
Iteration Number: 8022  
Loss: 29.373766830560406  
l2 norm of gradients: 0.187894996221513  
l2 norm of weights: 5.685501446198508  
-----  
Iteration Number: 8023  
Loss: 29.37199669604926  
l2 norm of gradients: 0.1878793890902674  
l2 norm of weights: 5.685433569816144  
-----  
Iteration Number: 8024  
Loss: 29.370226855472215  
l2 norm of gradients: 0.18786378501235618  
l2 norm of weights: 5.685365698558925  
-----  
Iteration Number: 8025  
Loss: 29.368457308745285  
l2 norm of gradients: 0.18784818398661038  
l2 norm of weights: 5.685297832426121  
-----  
Iteration Number: 8026  
Loss: 29.366688055789115  
l2 norm of gradients: 0.18783258601186187  
l2 norm of weights: 5.685229971417009  
-----  
Iteration Number: 8027  
Loss: 29.36491909651844
```

```

12 norm of gradients: 0.187816991086943
12 norm of weights: 5.68516211553086
-----
Iteration Number: 8028
Loss: 29.363150430853953
12 norm of gradients: 0.18780139921068692
12 norm of weights: 5.685094264766947
-----
Iteration Number: 8029
Loss: 29.361382058712238
12 norm of gradients: 0.18778581038192732
12 norm of weights: 5.685026419124545
-----
Iteration Number: 8030
Loss: 29.35961398000832
12 norm of gradients: 0.18777022459949874
12 norm of weights: 5.684958578602929
-----
Iteration Number: 8031
Loss: 29.35784619466512
12 norm of gradients: 0.18775464186223617
12 norm of weights: 5.684890743201371
-----
Iteration Number: 8032
Loss: 29.356078702597394
12 norm of gradients: 0.18773906216897537
12 norm of weights: 5.684822912919148
-----
Iteration Number: 8033
Loss: 29.354311503723125
12 norm of gradients: 0.1877234855185528
12 norm of weights: 5.684755087755532
-----
Iteration Number: 8034
Loss: 29.352544597961007
12 norm of gradients: 0.18770791190980549
12 norm of weights: 5.684687267709799
-----
Iteration Number: 8035
Loss: 29.35077798522645
12 norm of gradients: 0.1876923413415712
12 norm of weights: 5.684619452781224
-----
Iteration Number: 8036
Loss: 29.349011665442127
12 norm of gradients: 0.18767677381268827
12 norm of weights: 5.684551642969081
-----

```

Iteration Number: 8037
Loss: 29.34724563852276
l2 norm of gradients: 0.18766120932199581
l2 norm of weights: 5.684483838272647

Iteration Number: 8038
Loss: 29.345479904388483
l2 norm of gradients: 0.18764564786833346
l2 norm of weights: 5.684416038691197

Iteration Number: 8039
Loss: 29.343714462957077
l2 norm of gradients: 0.18763008945054163
l2 norm of weights: 5.684348244224005

Iteration Number: 8040
Loss: 29.34194931414579
l2 norm of gradients: 0.1876145340674614
l2 norm of weights: 5.684280454870349

Iteration Number: 8041
Loss: 29.34018445787038
l2 norm of gradients: 0.18759898171793438
l2 norm of weights: 5.6842126706295035

Iteration Number: 8042
Loss: 29.338419894054798
l2 norm of gradients: 0.18758343240080294
l2 norm of weights: 5.684144891500746

Iteration Number: 8043
Loss: 29.33665562261256
l2 norm of gradients: 0.1875678861149101
l2 norm of weights: 5.684077117483351

Iteration Number: 8044
Loss: 29.3348916434651
l2 norm of gradients: 0.18755234285909947
l2 norm of weights: 5.6840093485765975

Iteration Number: 8045
Loss: 29.3331279565311
l2 norm of gradients: 0.18753680263221545
l2 norm of weights: 5.6839415847797605

Iteration Number: 8046
Loss: 29.33136456172398
l2 norm of gradients: 0.18752126543310296

```

12 norm of weights: 5.683873826092117
-----
Iteration Number: 8047
Loss: 29.32960145896792
12 norm of gradients: 0.18750573126060766
12 norm of weights: 5.683806072512945
-----
Iteration Number: 8048
Loss: 29.327838648177273
12 norm of gradients: 0.18749020011357587
12 norm of weights: 5.683738324041521
-----
Iteration Number: 8049
Loss: 29.32607612927367
12 norm of gradients: 0.1874746719908544
12 norm of weights: 5.683670580677123
-----
Iteration Number: 8050
Loss: 29.3243139021751
12 norm of gradients: 0.18745914689129098
12 norm of weights: 5.683602842419028
-----
Iteration Number: 8051
Loss: 29.322551966798134
12 norm of gradients: 0.18744362481373386
12 norm of weights: 5.683535109266515
-----
Iteration Number: 8052
Loss: 29.320790323060002
12 norm of gradients: 0.18742810575703192
12 norm of weights: 5.68346738121886
-----
Iteration Number: 8053
Loss: 29.31902897088631
12 norm of gradients: 0.1874125897200347
12 norm of weights: 5.683399658275343
-----
Iteration Number: 8054
Loss: 29.317267910189216
12 norm of gradients: 0.18739707670159245
12 norm of weights: 5.683331940435242
-----
Iteration Number: 8055
Loss: 29.31550714089
12 norm of gradients: 0.187381566700556
12 norm of weights: 5.683264227697835
-----
Iteration Number: 8056

```

Loss: 29.313746662908322
l2 norm of gradients: 0.18736605971577694
l2 norm of weights: 5.683196520062401

Iteration Number: 8057
Loss: 29.311986476160552
l2 norm of gradients: 0.18735055574610743
l2 norm of weights: 5.683128817528219

Iteration Number: 8058
Loss: 29.31022658056672
l2 norm of gradients: 0.1873350547904003
l2 norm of weights: 5.683061120094568

Iteration Number: 8059
Loss: 29.30846697604551
l2 norm of gradients: 0.18731955684750898
l2 norm of weights: 5.682993427760728

Iteration Number: 8060
Loss: 29.30670766251658
l2 norm of gradients: 0.18730406191628762
l2 norm of weights: 5.6829257405259765

Iteration Number: 8061
Loss: 29.304948639897834
l2 norm of gradients: 0.18728856999559107
l2 norm of weights: 5.682858058389596

Iteration Number: 8062
Loss: 29.303189908108113
l2 norm of gradients: 0.1872730810842747
l2 norm of weights: 5.682790381350864

Iteration Number: 8063
Loss: 29.301431467068628
l2 norm of gradients: 0.1872575951811947
l2 norm of weights: 5.682722709409061

Iteration Number: 8064
Loss: 29.299673316694886
l2 norm of gradients: 0.18724211228520768
l2 norm of weights: 5.682655042563469

Iteration Number: 8065
Loss: 29.297915456909337
l2 norm of gradients: 0.18722663239517107
l2 norm of weights: 5.682587380813366

```
-----  
Iteration Number: 8066  
Loss: 29.29615788762876  
l2 norm of gradients: 0.18721115550994297  
l2 norm of weights: 5.682519724158032  
-----  
Iteration Number: 8067  
Loss: 29.29440060877273  
l2 norm of gradients: 0.18719568162838202  
l2 norm of weights: 5.6824520725967504  
-----  
Iteration Number: 8068  
Loss: 29.292643620260204  
l2 norm of gradients: 0.18718021074934754  
l2 norm of weights: 5.682384426128801  
-----  
Iteration Number: 8069  
Loss: 29.29088692201417  
l2 norm of gradients: 0.1871647428716995  
l2 norm of weights: 5.682316784753463  
-----  
Iteration Number: 8070  
Loss: 29.2891305139482  
l2 norm of gradients: 0.18714927799429867  
l2 norm of weights: 5.682249148470022  
-----  
Iteration Number: 8071  
Loss: 29.287374395985225  
l2 norm of gradients: 0.18713381611600619  
l2 norm of weights: 5.682181517277754  
-----  
Iteration Number: 8072  
Loss: 29.28561856804435  
l2 norm of gradients: 0.187118357235684  
l2 norm of weights: 5.682113891175945  
-----  
Iteration Number: 8073  
Loss: 29.28386303004302  
l2 norm of gradients: 0.18710290135219484  
l2 norm of weights: 5.682046270163873  
-----  
Iteration Number: 8074  
Loss: 29.282107781900034  
l2 norm of gradients: 0.18708744846440167  
l2 norm of weights: 5.681978654240824  
-----  
Iteration Number: 8075  
Loss: 29.280352823540547
```



```

12 norm of gradients: 0.18707199857116857
12 norm of weights: 5.681911043406078
-----
Iteration Number: 8076
Loss: 29.278598154876775
12 norm of gradients: 0.18705655167135998
12 norm of weights: 5.681843437658918
-----
Iteration Number: 8077
Loss: 29.276843775834095
12 norm of gradients: 0.18704110776384109
12 norm of weights: 5.681775836998625
-----
Iteration Number: 8078
Loss: 29.27508968632799
12 norm of gradients: 0.18702566684747765
12 norm of weights: 5.6817082414244835
-----
Iteration Number: 8079
Loss: 29.273335886278744
12 norm of gradients: 0.1870102289211362
12 norm of weights: 5.681640650935775
-----
Iteration Number: 8080
Loss: 29.271582375607455
12 norm of gradients: 0.18699479398368377
12 norm of weights: 5.681573065531785
-----
Iteration Number: 8081
Loss: 29.269829154231594
12 norm of gradients: 0.18697936203398816
12 norm of weights: 5.681505485211794
-----
Iteration Number: 8082
Loss: 29.268076222073965
12 norm of gradients: 0.18696393307091771
12 norm of weights: 5.681437909975085
-----
Iteration Number: 8083
Loss: 29.266323579052752
12 norm of gradients: 0.18694850709334146
12 norm of weights: 5.681370339820945
-----
Iteration Number: 8084
Loss: 29.264571225086453
12 norm of gradients: 0.18693308410012915
12 norm of weights: 5.681302774748654
-----

```

Iteration Number: 8085
Loss: 29.26281916009655
l2 norm of gradients: 0.186917664090151
l2 norm of weights: 5.681235214757499

Iteration Number: 8086
Loss: 29.2610673840015
l2 norm of gradients: 0.18690224706227807
l2 norm of weights: 5.6811676598467615

Iteration Number: 8087
Loss: 29.259315896722054
l2 norm of gradients: 0.1868868330153818
l2 norm of weights: 5.681100110015728

Iteration Number: 8088
Loss: 29.25756469818043
l2 norm of gradients: 0.18687142194833464
l2 norm of weights: 5.681032565263681

Iteration Number: 8089
Loss: 29.25581378829151
l2 norm of gradients: 0.18685601386000944
l2 norm of weights: 5.680965025589908

Iteration Number: 8090
Loss: 29.25406316697821
l2 norm of gradients: 0.18684060874927963
l2 norm of weights: 5.680897490993692

Iteration Number: 8091
Loss: 29.25231283416143
l2 norm of gradients: 0.18682520661501936
l2 norm of weights: 5.680829961474316

Iteration Number: 8092
Loss: 29.25056278975947
l2 norm of gradients: 0.1868098074561036
l2 norm of weights: 5.680762437031069

Iteration Number: 8093
Loss: 29.24881303369334
l2 norm of gradients: 0.1867944112714077
l2 norm of weights: 5.680694917663233

Iteration Number: 8094
Loss: 29.247063565880378
l2 norm of gradients: 0.18677901805980773

```

12 norm of weights: 5.680627403370096
-----
Iteration Number: 8095
Loss: 29.245314386244956
12 norm of gradients: 0.18676362782018047
12 norm of weights: 5.680559894150943
-----
Iteration Number: 8096
Loss: 29.24356549470246
12 norm of gradients: 0.1867482405514033
12 norm of weights: 5.68049239000506
-----
Iteration Number: 8097
Loss: 29.24181689117875
12 norm of gradients: 0.18673285625235422
12 norm of weights: 5.680424890931733
-----
Iteration Number: 8098
Loss: 29.240068575591305
12 norm of gradients: 0.1867174749219118
12 norm of weights: 5.680357396930248
-----
Iteration Number: 8099
Loss: 29.238320547861147
12 norm of gradients: 0.18670209655895542
12 norm of weights: 5.680289907999891
-----
Iteration Number: 8100
Loss: 29.23657280790634
12 norm of gradients: 0.18668672116236504
12 norm of weights: 5.6802224241399495
-----
Iteration Number: 8101
Loss: 29.23482535564832
12 norm of gradients: 0.1866713487310211
12 norm of weights: 5.6801549453497095
-----
Iteration Number: 8102
Loss: 29.233078191009316
12 norm of gradients: 0.18665597926380484
12 norm of weights: 5.680087471628459
-----
Iteration Number: 8103
Loss: 29.231331313907543
12 norm of gradients: 0.18664061275959823
12 norm of weights: 5.680020002975485
-----
Iteration Number: 8104

```

```

Loss: 29.22958472426338
l2 norm of gradients: 0.1866252492172836
l2 norm of weights: 5.679952539390074
-----
Iteration Number: 8105
Loss: 29.22783842199854
l2 norm of gradients: 0.18660988863574407
l2 norm of weights: 5.679885080871513
-----
Iteration Number: 8106
Loss: 29.22609240703443
l2 norm of gradients: 0.18659453101386342
l2 norm of weights: 5.679817627419092
-----
Iteration Number: 8107
Loss: 29.224346679289166
l2 norm of gradients: 0.18657917635052598
l2 norm of weights: 5.679750179032095
-----
Iteration Number: 8108
Loss: 29.22260123868476
l2 norm of gradients: 0.18656382464461693
l2 norm of weights: 5.6796827357098145
-----
Iteration Number: 8109
Loss: 29.220856085142998
l2 norm of gradients: 0.18654847589502171
l2 norm of weights: 5.6796152974515355
-----
Iteration Number: 8110
Loss: 29.219111218581943
l2 norm of gradients: 0.18653313010062675
l2 norm of weights: 5.6795478642565485
-----
Iteration Number: 8111
Loss: 29.217366638920474
l2 norm of gradients: 0.1865177872603189
l2 norm of weights: 5.679480436124141
-----
Iteration Number: 8112
Loss: 29.215622346086285
l2 norm of gradients: 0.18650244737298574
l2 norm of weights: 5.679413013053601
-----
Iteration Number: 8113
Loss: 29.21387833999609
l2 norm of gradients: 0.18648711043751548
l2 norm of weights: 5.679345595044219

```

```

-----
Iteration Number: 8114
Loss: 29.2121346205697
l2 norm of gradients: 0.1864717764527969
l2 norm of weights: 5.679278182095283
-----
Iteration Number: 8115
Loss: 29.2103911877302
l2 norm of gradients: 0.18645644541771947
l2 norm of weights: 5.679210774206083
-----
Iteration Number: 8116
Loss: 29.20864804139366
l2 norm of gradients: 0.18644111733117333
l2 norm of weights: 5.679143371375907
-----
Iteration Number: 8117
Loss: 29.206905181488658
l2 norm of gradients: 0.18642579219204905
l2 norm of weights: 5.679075973604046
-----
Iteration Number: 8118
Loss: 29.205162607928433
l2 norm of gradients: 0.1864104699992382
l2 norm of weights: 5.67900858088979
-----
Iteration Number: 8119
Loss: 29.203420320642444
l2 norm of gradients: 0.1863951507516326
l2 norm of weights: 5.6789411932324265
-----
Iteration Number: 8120
Loss: 29.201678319544303
l2 norm of gradients: 0.18637983444812492
l2 norm of weights: 5.6788738106312495
-----
Iteration Number: 8121
Loss: 29.199936604557028
l2 norm of gradients: 0.18636452108760837
l2 norm of weights: 5.678806433085547
-----
Iteration Number: 8122
Loss: 29.19819517560125
l2 norm of gradients: 0.18634921066897686
l2 norm of weights: 5.678739060594609
-----
Iteration Number: 8123
Loss: 29.196454032599533

```

```

12 norm of gradients: 0.18633390319112492
12 norm of weights: 5.678671693157728
-----
Iteration Number: 8124
Loss: 29.194713175476156
12 norm of gradients: 0.18631859865294761
12 norm of weights: 5.678604330774193
-----
Iteration Number: 8125
Loss: 29.192972604144266
12 norm of gradients: 0.18630329705334076
12 norm of weights: 5.678536973443296
-----
Iteration Number: 8126
Loss: 29.19123231853165
12 norm of gradients: 0.18628799839120075
12 norm of weights: 5.678469621164328
-----
Iteration Number: 8127
Loss: 29.18949231855877
12 norm of gradients: 0.18627270266542464
12 norm of weights: 5.67840227393658
-----
Iteration Number: 8128
Loss: 29.187752604143736
12 norm of gradients: 0.18625740987491
12 norm of weights: 5.678334931759345
-----
Iteration Number: 8129
Loss: 29.186013175209904
12 norm of gradients: 0.18624212001855517
12 norm of weights: 5.6782675946319126
-----
Iteration Number: 8130
Loss: 29.184274031677997
12 norm of gradients: 0.18622683309525898
12 norm of weights: 5.678200262553576
-----
Iteration Number: 8131
Loss: 29.182535173469812
12 norm of gradients: 0.18621154910392113
12 norm of weights: 5.678132935523627
-----
Iteration Number: 8132
Loss: 29.180796600509446
12 norm of gradients: 0.1861962680434416
12 norm of weights: 5.678065613541358
-----

```

Iteration Number: 8133
Loss: 29.179058312711764
l2 norm of gradients: 0.18618098991272128
l2 norm of weights: 5.67799829660606

Iteration Number: 8134
Loss: 29.177320310004752
l2 norm of gradients: 0.18616571471066157
l2 norm of weights: 5.67793098471703

Iteration Number: 8135
Loss: 29.17558259230725
l2 norm of gradients: 0.1861504424361645
l2 norm of weights: 5.677863677873556

Iteration Number: 8136
Loss: 29.17384515953968
l2 norm of gradients: 0.18613517308813277
l2 norm of weights: 5.677796376074932

Iteration Number: 8137
Loss: 29.17210801162683
l2 norm of gradients: 0.18611990666546965
l2 norm of weights: 5.677729079320452

Iteration Number: 8138
Loss: 29.17037114848691
l2 norm of gradients: 0.18610464316707903
l2 norm of weights: 5.677661787609409

Iteration Number: 8139
Loss: 29.168634570042794
l2 norm of gradients: 0.18608938259186553
l2 norm of weights: 5.677594500941096

Iteration Number: 8140
Loss: 29.166898276217335
l2 norm of gradients: 0.18607412493873424
l2 norm of weights: 5.6775272193148085

Iteration Number: 8141
Loss: 29.165162266930672
l2 norm of gradients: 0.18605887020659098
l2 norm of weights: 5.677459942729837

Iteration Number: 8142
Loss: 29.163426542105665
l2 norm of gradients: 0.1860436183943422

```

12 norm of weights: 5.677392671185479
-----
Iteration Number: 8143
Loss: 29.161691101663386
12 norm of gradients: 0.1860283695008949
12 norm of weights: 5.677325404681027
-----
Iteration Number: 8144
Loss: 29.159955945525713
12 norm of gradients: 0.18601312352515675
12 norm of weights: 5.6772581432157745
-----
Iteration Number: 8145
Loss: 29.15822107361459
12 norm of gradients: 0.18599788046603608
12 norm of weights: 5.677190886789017
-----
Iteration Number: 8146
Loss: 29.156486485853193
12 norm of gradients: 0.18598264032244172
12 norm of weights: 5.67712363540005
-----
Iteration Number: 8147
Loss: 29.15475218216427
12 norm of gradients: 0.18596740309328322
12 norm of weights: 5.677056389048167
-----
Iteration Number: 8148
Loss: 29.1530181624621
12 norm of gradients: 0.1859521687774708
12 norm of weights: 5.676989147732663
-----
Iteration Number: 8149
Loss: 29.15128442667899
12 norm of gradients: 0.18593693737391517
12 norm of weights: 5.676921911452833
-----
Iteration Number: 8150
Loss: 29.149550974730186
12 norm of gradients: 0.1859217088815278
12 norm of weights: 5.676854680207975
-----
Iteration Number: 8151
Loss: 29.147817806540928
12 norm of gradients: 0.1859064832992206
12 norm of weights: 5.676787453997381
-----
Iteration Number: 8152

```


Loss: 29.146084922032546
l2 norm of gradients: 0.18589126062590627
l2 norm of weights: 5.67672023282035

Iteration Number: 8153
Loss: 29.144352321127904
l2 norm of gradients: 0.18587604086049808
l2 norm of weights: 5.676653016676175

Iteration Number: 8154
Loss: 29.142620003746334
l2 norm of gradients: 0.1858608240019098
l2 norm of weights: 5.676585805564153

Iteration Number: 8155
Loss: 29.14088796981195
l2 norm of gradients: 0.18584561004905606
l2 norm of weights: 5.676518599483582

Iteration Number: 8156
Loss: 29.13915621924863
l2 norm of gradients: 0.18583039900085194
l2 norm of weights: 5.676451398433756

Iteration Number: 8157
Loss: 29.137424751974642
l2 norm of gradients: 0.18581519085621317
l2 norm of weights: 5.676384202413972

Iteration Number: 8158
Loss: 29.135693567916228
l2 norm of gradients: 0.18579998561405606
l2 norm of weights: 5.676317011423528

Iteration Number: 8159
Loss: 29.13396266699514
l2 norm of gradients: 0.18578478327329762
l2 norm of weights: 5.676249825461721

Iteration Number: 8160
Loss: 29.13223204913071
l2 norm of gradients: 0.18576958383285544
l2 norm of weights: 5.676182644527846

Iteration Number: 8161
Loss: 29.13050171424819
l2 norm of gradients: 0.18575438729164773
l2 norm of weights: 5.6761154686212025

```

-----
Iteration Number: 8162
Loss: 29.12877166227031
l2 norm of gradients: 0.18573919364859334
l2 norm of weights: 5.676048297741087
-----
Iteration Number: 8163
Loss: 29.12704189311532
l2 norm of gradients: 0.18572400290261157
l2 norm of weights: 5.675981131886798
-----
Iteration Number: 8164
Loss: 29.125312406710737
l2 norm of gradients: 0.18570881505262274
l2 norm of weights: 5.67591397105763
-----
Iteration Number: 8165
Loss: 29.123583202977294
l2 norm of gradients: 0.18569363009754725
l2 norm of weights: 5.675846815252886
-----
Iteration Number: 8166
Loss: 29.121854281837177
l2 norm of gradients: 0.18567844803630654
l2 norm of weights: 5.67577966447186
-----
Iteration Number: 8167
Loss: 29.120125643212457
l2 norm of gradients: 0.1856632688678225
l2 norm of weights: 5.675712518713853
-----
Iteration Number: 8168
Loss: 29.118397287028863
l2 norm of gradients: 0.18564809259101758
l2 norm of weights: 5.675645377978162
-----
Iteration Number: 8169
Loss: 29.116669213204116
l2 norm of gradients: 0.185632919204815
l2 norm of weights: 5.675578242264086
-----
Iteration Number: 8170
Loss: 29.11494142166603
l2 norm of gradients: 0.18561774870813855
l2 norm of weights: 5.675511111570924
-----
Iteration Number: 8171
Loss: 29.11321391233205

```

```

12 norm of gradients: 0.18560258109991243
12 norm of weights: 5.675443985897975
-----
Iteration Number: 8172
Loss: 29.11148668513106
12 norm of gradients: 0.18558741637906176
12 norm of weights: 5.675376865244539
-----
Iteration Number: 8173
Loss: 29.109759739979964
12 norm of gradients: 0.1855722545445121
12 norm of weights: 5.675309749609913
-----
Iteration Number: 8174
Loss: 29.10803307680469
12 norm of gradients: 0.18555709559518965
12 norm of weights: 5.6752426389933985
-----
Iteration Number: 8175
Loss: 29.10630669552797
12 norm of gradients: 0.18554193953002118
12 norm of weights: 5.675175533394295
-----
Iteration Number: 8176
Loss: 29.104580596072463
12 norm of gradients: 0.18552678634793415
12 norm of weights: 5.675108432811902
-----
Iteration Number: 8177
Loss: 29.102854778361557
12 norm of gradients: 0.18551163604785667
12 norm of weights: 5.675041337245519
-----
Iteration Number: 8178
Loss: 29.101129242316354
12 norm of gradients: 0.1854964886287173
12 norm of weights: 5.674974246694447
-----
Iteration Number: 8179
Loss: 29.09940398786256
12 norm of gradients: 0.1854813440894454
12 norm of weights: 5.674907161157986
-----
Iteration Number: 8180
Loss: 29.09767901491975
12 norm of gradients: 0.18546620242897074
12 norm of weights: 5.674840080635437
-----

```

Iteration Number: 8181
Loss: 29.09595432341443
l2 norm of gradients: 0.18545106364622388
l2 norm of weights: 5.6747730051261005

Iteration Number: 8182
Loss: 29.09422991326697
l2 norm of gradients: 0.18543592774013595
l2 norm of weights: 5.6747059346292765

Iteration Number: 8183
Loss: 29.09250578440245
l2 norm of gradients: 0.18542079470963854
l2 norm of weights: 5.674638869144268

Iteration Number: 8184
Loss: 29.09078193674636
l2 norm of gradients: 0.18540566455366414
l2 norm of weights: 5.6745718086703745

Iteration Number: 8185
Loss: 29.089058370215263
l2 norm of gradients: 0.18539053727114554
l2 norm of weights: 5.6745047532068975

Iteration Number: 8186
Loss: 29.087335084737123
l2 norm of gradients: 0.18537541286101633
l2 norm of weights: 5.67443770275314

Iteration Number: 8187
Loss: 29.085612080234128
l2 norm of gradients: 0.18536029132221063
l2 norm of weights: 5.674370657308402

Iteration Number: 8188
Loss: 29.083889356631317
l2 norm of gradients: 0.18534517265366326
l2 norm of weights: 5.6743036168719865

Iteration Number: 8189
Loss: 29.082166913850298
l2 norm of gradients: 0.18533005685430956
l2 norm of weights: 5.674236581443195

Iteration Number: 8190
Loss: 29.080444751810337
l2 norm of gradients: 0.18531494392308548

```
l2 norm of weights: 5.674169551021332
-----
Iteration Number: 8191
Loss: 29.078722870443674
l2 norm of gradients: 0.18529983385892762
l2 norm of weights: 5.674102525605695
-----
Iteration Number: 8192
Loss: 29.077001269667853
l2 norm of gradients: 0.18528472666077314
l2 norm of weights: 5.674035505195591
-----
Iteration Number: 8193
Loss: 29.07527994940581
l2 norm of gradients: 0.1852696223275599
l2 norm of weights: 5.673968489790322
-----
Iteration Number: 8194
Loss: 29.073558909584968
l2 norm of gradients: 0.1852545208582262
l2 norm of weights: 5.6739014793891895
-----
Iteration Number: 8195
Loss: 29.071838150124236
l2 norm of gradients: 0.18523942225171122
l2 norm of weights: 5.673834473991497
-----
Iteration Number: 8196
Loss: 29.07011767094958
l2 norm of gradients: 0.18522432650695442
l2 norm of weights: 5.673767473596548
-----
Iteration Number: 8197
Loss: 29.06839747198888
l2 norm of gradients: 0.185209233622896
l2 norm of weights: 5.673700478203647
-----
Iteration Number: 8198
Loss: 29.066677553157344
l2 norm of gradients: 0.185194143598477
l2 norm of weights: 5.673633487812096
-----
Iteration Number: 8199
Loss: 29.064957914384166
l2 norm of gradients: 0.18517905643263863
l2 norm of weights: 5.6735665024212
-----
Iteration Number: 8200
```

```

Loss: 29.063238555591482
12 norm of gradients: 0.185163972124323
12 norm of weights: 5.673499522030263
-----
Iteration Number: 8201
Loss: 29.06151947670265
12 norm of gradients: 0.18514889067247278
12 norm of weights: 5.673432546638587
-----
Iteration Number: 8202
Loss: 29.059800677642336
12 norm of gradients: 0.1851338120760312
12 norm of weights: 5.673365576245479
-----
Iteration Number: 8203
Loss: 29.058082158334827
12 norm of gradients: 0.1851187363339421
12 norm of weights: 5.673298610850243
-----
Iteration Number: 8204
Loss: 29.056363918702584
12 norm of gradients: 0.18510366344514995
12 norm of weights: 5.673231650452182
-----
Iteration Number: 8205
Loss: 29.054645958669777
12 norm of gradients: 0.1850885934085998
12 norm of weights: 5.673164695050602
-----
Iteration Number: 8206
Loss: 29.05292827815897
12 norm of gradients: 0.1850735262232373
12 norm of weights: 5.673097744644807
-----
Iteration Number: 8207
Loss: 29.0512108770985
12 norm of gradients: 0.1850584618880087
12 norm of weights: 5.6730307992341045
-----
Iteration Number: 8208
Loss: 29.049493755408044
12 norm of gradients: 0.18504340040186093
12 norm of weights: 5.672963858817797
-----
Iteration Number: 8209
Loss: 29.047776913012
12 norm of gradients: 0.1850283417637414
12 norm of weights: 5.672896923395191

```

```

-----
Iteration Number: 8210
Loss: 29.046060349838044
l2 norm of gradients: 0.18501328597259817
l2 norm of weights: 5.6728299929655925
-----
Iteration Number: 8211
Loss: 29.044344065805912
l2 norm of gradients: 0.18499823302737992
l2 norm of weights: 5.672763067528307
-----
Iteration Number: 8212
Loss: 29.042628060840702
l2 norm of gradients: 0.18498318292703594
l2 norm of weights: 5.672696147082642
-----
Iteration Number: 8213
Loss: 29.040912334867034
l2 norm of gradients: 0.1849681356705161
l2 norm of weights: 5.6726292316279014
-----
Iteration Number: 8214
Loss: 29.03919688781054
l2 norm of gradients: 0.1849530912567708
l2 norm of weights: 5.672562321163392
-----
Iteration Number: 8215
Loss: 29.037481719593217
l2 norm of gradients: 0.18493804968475125
l2 norm of weights: 5.672495415688421
-----
Iteration Number: 8216
Loss: 29.03576683013995
l2 norm of gradients: 0.18492301095340902
l2 norm of weights: 5.672428515202294
-----
Iteration Number: 8217
Loss: 29.03405221937345
l2 norm of gradients: 0.18490797506169634
l2 norm of weights: 5.672361619704319
-----
Iteration Number: 8218
Loss: 29.032337887224017
l2 norm of gradients: 0.18489294200856624
l2 norm of weights: 5.672294729193804
-----
Iteration Number: 8219
Loss: 29.03062383360865

```

```

12 norm of gradients: 0.184877911792972
12 norm of weights: 5.672227843670053
-----
Iteration Number: 8220
Loss: 29.02891005845523
12 norm of gradients: 0.18486288441386783
12 norm of weights: 5.672160963132376
-----
Iteration Number: 8221
Loss: 29.027196561687255
12 norm of gradients: 0.18484785987020833
12 norm of weights: 5.67209408758008
-----
Iteration Number: 8222
Loss: 29.02548334323024
12 norm of gradients: 0.18483283816094878
12 norm of weights: 5.672027217012471
-----
Iteration Number: 8223
Loss: 29.02377040300886
12 norm of gradients: 0.18481781928504504
12 norm of weights: 5.67196035142886
-----
Iteration Number: 8224
Loss: 29.02205774094244
12 norm of gradients: 0.18480280324145354
12 norm of weights: 5.671893490828551
-----
Iteration Number: 8225
Loss: 29.020345356963993
12 norm of gradients: 0.1847877900291314
12 norm of weights: 5.671826635210855
-----
Iteration Number: 8226
Loss: 29.018633250991055
12 norm of gradients: 0.18477277964703617
12 norm of weights: 5.67175978457508
-----
Iteration Number: 8227
Loss: 29.016921422953192
12 norm of gradients: 0.1847577720941262
12 norm of weights: 5.6716929389205335
-----
Iteration Number: 8228
Loss: 29.015209872770182
12 norm of gradients: 0.18474276736936024
12 norm of weights: 5.671626098246525
-----

```


Iteration Number: 8229
Loss: 29.013498600370784
l2 norm of gradients: 0.18472776547169778
l2 norm of weights: 5.671559262552364

Iteration Number: 8230
Loss: 29.011787605678037
l2 norm of gradients: 0.18471276640009887
l2 norm of weights: 5.671492431837358

Iteration Number: 8231
Loss: 29.01007688861639
l2 norm of gradients: 0.1846977701535241
l2 norm of weights: 5.671425606100817

Iteration Number: 8232
Loss: 29.008366449110138
l2 norm of gradients: 0.1846827767309347
l2 norm of weights: 5.67135878534205

Iteration Number: 8233
Loss: 29.006656287085498
l2 norm of gradients: 0.1846677861312925
l2 norm of weights: 5.671291969560366

Iteration Number: 8234
Loss: 29.004946402464828
l2 norm of gradients: 0.18465279835355994
l2 norm of weights: 5.671225158755076

Iteration Number: 8235
Loss: 29.003236795176246
l2 norm of gradients: 0.18463781339669996
l2 norm of weights: 5.671158352925491

Iteration Number: 8236
Loss: 29.00152746514338
l2 norm of gradients: 0.18462283125967618
l2 norm of weights: 5.6710915520709175

Iteration Number: 8237
Loss: 28.999818412287354
l2 norm of gradients: 0.1846078519414528
l2 norm of weights: 5.6710247561906675

Iteration Number: 8238
Loss: 28.998109636541187
l2 norm of gradients: 0.18459287544099462

```

12 norm of weights: 5.670957965284051
-----
Iteration Number: 8239
Loss: 28.996401137822353
12 norm of gradients: 0.18457790175726696
12 norm of weights: 5.67089117935038
-----
Iteration Number: 8240
Loss: 28.994692916058565
12 norm of gradients: 0.18456293088923587
12 norm of weights: 5.670824398388963
-----
Iteration Number: 8241
Loss: 28.992984971175737
12 norm of gradients: 0.18454796283586786
12 norm of weights: 5.670757622399112
-----
Iteration Number: 8242
Loss: 28.991277303097924
12 norm of gradients: 0.18453299759613007
12 norm of weights: 5.670690851380138
-----
Iteration Number: 8243
Loss: 28.98956991175136
12 norm of gradients: 0.18451803516899024
12 norm of weights: 5.670624085331352
-----
Iteration Number: 8244
Loss: 28.987862797058426
12 norm of gradients: 0.18450307555341675
12 norm of weights: 5.670557324252066
-----
Iteration Number: 8245
Loss: 28.986155958945442
12 norm of gradients: 0.1844881187483785
12 norm of weights: 5.670490568141591
-----
Iteration Number: 8246
Loss: 28.98444939734131
12 norm of gradients: 0.18447316475284498
12 norm of weights: 5.670423816999238
-----
Iteration Number: 8247
Loss: 28.982743112166105
12 norm of gradients: 0.18445821356578632
12 norm of weights: 5.67035707082432
-----
Iteration Number: 8248

```

Loss: 28.981037103345663
l2 norm of gradients: 0.18444326518617318
l2 norm of weights: 5.670290329616147

Iteration Number: 8249
Loss: 28.979331370808207
l2 norm of gradients: 0.1844283196129769
l2 norm of weights: 5.670223593374034

Iteration Number: 8250
Loss: 28.97762591447812
l2 norm of gradients: 0.18441337684516934
l2 norm of weights: 5.670156862097292

Iteration Number: 8251
Loss: 28.975920734278905
l2 norm of gradients: 0.18439843688172283
l2 norm of weights: 5.670090135785233

Iteration Number: 8252
Loss: 28.974215830137304
l2 norm of gradients: 0.1843834997216106
l2 norm of weights: 5.670023414437171

Iteration Number: 8253
Loss: 28.972511201979234
l2 norm of gradients: 0.18436856536380625
l2 norm of weights: 5.669956698052418

Iteration Number: 8254
Loss: 28.97080684972974
l2 norm of gradients: 0.18435363380728387
l2 norm of weights: 5.669889986630286

Iteration Number: 8255
Loss: 28.96910277331287
l2 norm of gradients: 0.18433870505101846
l2 norm of weights: 5.66982328017009

Iteration Number: 8256
Loss: 28.96739897265307
l2 norm of gradients: 0.18432377909398523
l2 norm of weights: 5.669756578671143

Iteration Number: 8257
Loss: 28.965695447682997
l2 norm of gradients: 0.1843088559351603
l2 norm of weights: 5.669689882132757

```

-----
Iteration Number: 8258
Loss: 28.96399219831864
l2 norm of gradients: 0.1842939355735202
l2 norm of weights: 5.6696231905542485
-----
Iteration Number: 8259
Loss: 28.962289224493645
l2 norm of gradients: 0.18427901800804206
l2 norm of weights: 5.6695565039349285
-----
Iteration Number: 8260
Loss: 28.96058652612913
l2 norm of gradients: 0.18426410323770367
l2 norm of weights: 5.6694898222741115
-----
Iteration Number: 8261
Loss: 28.95888410315176
l2 norm of gradients: 0.1842491912614833
l2 norm of weights: 5.669423145571113
-----
Iteration Number: 8262
Loss: 28.957181955487673
l2 norm of gradients: 0.18423428207835993
l2 norm of weights: 5.669356473825246
-----
Iteration Number: 8263
Loss: 28.955480083061556
l2 norm of gradients: 0.18421937568731295
l2 norm of weights: 5.669289807035827
-----
Iteration Number: 8264
Loss: 28.953778485800342
l2 norm of gradients: 0.18420447208732257
l2 norm of weights: 5.669223145202168
-----
Iteration Number: 8265
Loss: 28.952077163626903
l2 norm of gradients: 0.18418957127736937
l2 norm of weights: 5.669156488323586
-----
Iteration Number: 8266
Loss: 28.95037611647137
l2 norm of gradients: 0.18417467325643466
l2 norm of weights: 5.669089836399395
-----
Iteration Number: 8267
Loss: 28.94867534425756

```

```

12 norm of gradients: 0.18415977802350017
12 norm of weights: 5.66902318942891
-----
Iteration Number: 8268
Loss: 28.9469748469126
12 norm of gradients: 0.1841448855775485
12 norm of weights: 5.668956547411446
-----
Iteration Number: 8269
Loss: 28.94527462435965
12 norm of gradients: 0.1841299959175624
12 norm of weights: 5.66888991034632
-----
Iteration Number: 8270
Loss: 28.94357467652712
12 norm of gradients: 0.18411510904252565
12 norm of weights: 5.668823278232847
-----
Iteration Number: 8271
Loss: 28.941875003340808
12 norm of gradients: 0.18410022495142236
12 norm of weights: 5.668756651070342
-----
Iteration Number: 8272
Loss: 28.94017560472411
12 norm of gradients: 0.18408534364323725
12 norm of weights: 5.668690028858122
-----
Iteration Number: 8273
Loss: 28.93847648060531
12 norm of gradients: 0.18407046511695563
12 norm of weights: 5.668623411595503
-----
Iteration Number: 8274
Loss: 28.93677763090981
12 norm of gradients: 0.1840555893715635
12 norm of weights: 5.6685567992818
-----
Iteration Number: 8275
Loss: 28.935079055566554
12 norm of gradients: 0.18404071640604727
12 norm of weights: 5.668490191916331
-----
Iteration Number: 8276
Loss: 28.933380754497453
12 norm of gradients: 0.18402584621939402
12 norm of weights: 5.6684235894984125
-----

```

Iteration Number: 8277
Loss: 28.93168272763043
l2 norm of gradients: 0.1840109788105914
l2 norm of weights: 5.668356992027361

Iteration Number: 8278
Loss: 28.92998497489039
l2 norm of gradients: 0.18399611417862766
l2 norm of weights: 5.668290399502493

Iteration Number: 8279
Loss: 28.928287496205357
l2 norm of gradients: 0.18398125232249163
l2 norm of weights: 5.668223811923126

Iteration Number: 8280
Loss: 28.926590291502876
l2 norm of gradients: 0.18396639324117264
l2 norm of weights: 5.668157229288577

Iteration Number: 8281
Loss: 28.92489336070503
l2 norm of gradients: 0.1839515369336607
l2 norm of weights: 5.668090651598166

Iteration Number: 8282
Loss: 28.92319670374203
l2 norm of gradients: 0.18393668339894634
l2 norm of weights: 5.668024078851207

Iteration Number: 8283
Loss: 28.921500320538108
l2 norm of gradients: 0.18392183263602072
l2 norm of weights: 5.667957511047019

Iteration Number: 8284
Loss: 28.91980421101968
l2 norm of gradients: 0.18390698464387548
l2 norm of weights: 5.667890948184921

Iteration Number: 8285
Loss: 28.91810837511374
l2 norm of gradients: 0.18389213942150295
l2 norm of weights: 5.66782439026423

Iteration Number: 8286
Loss: 28.91641281274593
l2 norm of gradients: 0.18387729696789598

```

12 norm of weights: 5.667757837284265
-----
Iteration Number: 8287
Loss: 28.914717523843148
12 norm of gradients: 0.18386245728204798
12 norm of weights: 5.667691289244344
-----
Iteration Number: 8288
Loss: 28.913022508333764
12 norm of gradients: 0.18384762036295302
12 norm of weights: 5.667624746143784
-----
Iteration Number: 8289
Loss: 28.91132776614272
12 norm of gradients: 0.18383278620960564
12 norm of weights: 5.667558207981908
-----
Iteration Number: 8290
Loss: 28.90963329719651
12 norm of gradients: 0.18381795482100105
12 norm of weights: 5.667491674758031
-----
Iteration Number: 8291
Loss: 28.907939101420645
12 norm of gradients: 0.1838031261961349
12 norm of weights: 5.667425146471473
-----
Iteration Number: 8292
Loss: 28.9062451787431
12 norm of gradients: 0.18378830033400367
12 norm of weights: 5.667358623121555
-----
Iteration Number: 8293
Loss: 28.904551529090437
12 norm of gradients: 0.1837734772336041
12 norm of weights: 5.667292104707594
-----
Iteration Number: 8294
Loss: 28.90285815239345
12 norm of gradients: 0.1837586568939337
12 norm of weights: 5.667225591228912
-----
Iteration Number: 8295
Loss: 28.901165048568753
12 norm of gradients: 0.18374383931399052
12 norm of weights: 5.667159082684827
-----
Iteration Number: 8296

```

Loss: 28.899472217552294
l2 norm of gradients: 0.18372902449277323
l2 norm of weights: 5.66709257907466

Iteration Number: 8297
Loss: 28.89777965926541
l2 norm of gradients: 0.18371421242928093
l2 norm of weights: 5.667026080397729

Iteration Number: 8298
Loss: 28.89608737363859
l2 norm of gradients: 0.18369940312251343
l2 norm of weights: 5.666959586653357

Iteration Number: 8299
Loss: 28.894395360597088
l2 norm of gradients: 0.18368459657147107
l2 norm of weights: 5.666893097840863

Iteration Number: 8300
Loss: 28.89270362006541
l2 norm of gradients: 0.18366979277515477
l2 norm of weights: 5.666826613959568

Iteration Number: 8301
Loss: 28.891012151974554
l2 norm of gradients: 0.183654991732566
l2 norm of weights: 5.666760135008792

Iteration Number: 8302
Loss: 28.889320956250334
l2 norm of gradients: 0.18364019344270685
l2 norm of weights: 5.666693660987856

Iteration Number: 8303
Loss: 28.887630032819
l2 norm of gradients: 0.18362539790457985
l2 norm of weights: 5.666627191896081

Iteration Number: 8304
Loss: 28.885939381607756
l2 norm of gradients: 0.1836106051171883
l2 norm of weights: 5.666560727732791

Iteration Number: 8305
Loss: 28.884249002542767
l2 norm of gradients: 0.18359581507953596
l2 norm of weights: 5.666494268497303


```

-----
Iteration Number: 8306
Loss: 28.8825588955534
l2 norm of gradients: 0.1835810277906271
l2 norm of weights: 5.666427814188941
-----
Iteration Number: 8307
Loss: 28.880869060566
l2 norm of gradients: 0.18356624324946674
l2 norm of weights: 5.666361364807026
-----
Iteration Number: 8308
Loss: 28.87917949750446
l2 norm of gradients: 0.1835514614550603
l2 norm of weights: 5.666294920350881
-----
Iteration Number: 8309
Loss: 28.877490206299264
l2 norm of gradients: 0.18353668240641383
l2 norm of weights: 5.666228480819827
-----
Iteration Number: 8310
Loss: 28.875801186877464
l2 norm of gradients: 0.183521906102534
l2 norm of weights: 5.666162046213185
-----
Iteration Number: 8311
Loss: 28.874112439164033
l2 norm of gradients: 0.18350713254242795
l2 norm of weights: 5.66609561653028
-----
Iteration Number: 8312
Loss: 28.872423963087588
l2 norm of gradients: 0.18349236172510347
l2 norm of weights: 5.666029191770433
-----
Iteration Number: 8313
Loss: 28.870735758577283
l2 norm of gradients: 0.18347759364956892
l2 norm of weights: 5.665962771932965
-----
Iteration Number: 8314
Loss: 28.869047825557764
l2 norm of gradients: 0.18346282831483315
l2 norm of weights: 5.6658963570172025
-----
Iteration Number: 8315
Loss: 28.86736016395737

```

```

12 norm of gradients: 0.18344806571990566
12 norm of weights: 5.665829947022466
-----
Iteration Number: 8316
Loss: 28.865672773703572
12 norm of gradients: 0.18343330586379653
12 norm of weights: 5.665763541948081
-----
Iteration Number: 8317
Loss: 28.863985654722992
12 norm of gradients: 0.18341854874551627
12 norm of weights: 5.665697141793368
-----
Iteration Number: 8318
Loss: 28.86229880694373
12 norm of gradients: 0.18340379436407608
12 norm of weights: 5.665630746557652
-----
Iteration Number: 8319
Loss: 28.860612230294638
12 norm of gradients: 0.18338904271848783
12 norm of weights: 5.665564356240255
-----
Iteration Number: 8320
Loss: 28.858925924699292
12 norm of gradients: 0.18337429380776366
12 norm of weights: 5.665497970840503
-----
Iteration Number: 8321
Loss: 28.857239890089737
12 norm of gradients: 0.1833595476309165
12 norm of weights: 5.66543159035772
-----
Iteration Number: 8322
Loss: 28.855554126390963
12 norm of gradients: 0.18334480418695978
12 norm of weights: 5.665365214791228
-----
Iteration Number: 8323
Loss: 28.853868633531455
12 norm of gradients: 0.1833300634749076
12 norm of weights: 5.665298844140353
-----
Iteration Number: 8324
Loss: 28.8521834114375
12 norm of gradients: 0.18331532549377438
12 norm of weights: 5.665232478404419
-----

```

Iteration Number: 8325
Loss: 28.850498460036277
l2 norm of gradients: 0.1833005902425754
l2 norm of weights: 5.66516611758275

Iteration Number: 8326
Loss: 28.84881377925865
l2 norm of gradients: 0.1832858577203262
l2 norm of weights: 5.665099761674671

Iteration Number: 8327
Loss: 28.847129369032203
l2 norm of gradients: 0.18327112792604322
l2 norm of weights: 5.665033410679508

Iteration Number: 8328
Loss: 28.845445229278447
l2 norm of gradients: 0.18325640085874323
l2 norm of weights: 5.664967064596585

Iteration Number: 8329
Loss: 28.84376135993147
l2 norm of gradients: 0.18324167651744358
l2 norm of weights: 5.664900723425228

Iteration Number: 8330
Loss: 28.842077760918524
l2 norm of gradients: 0.18322695490116228
l2 norm of weights: 5.664834387164762

Iteration Number: 8331
Loss: 28.84039443216444
l2 norm of gradients: 0.18321223600891778
l2 norm of weights: 5.664768055814513

Iteration Number: 8332
Loss: 28.838711373599015
l2 norm of gradients: 0.18319751983972932
l2 norm of weights: 5.664701729373804

Iteration Number: 8333
Loss: 28.837028585150957
l2 norm of gradients: 0.18318280639261636
l2 norm of weights: 5.664635407841966

Iteration Number: 8334
Loss: 28.83534606674588
l2 norm of gradients: 0.18316809566659922

```

12 norm of weights: 5.66456909121832
-----
Iteration Number: 8335
Loss: 28.833663818314125
12 norm of gradients: 0.18315338766069866
12 norm of weights: 5.664502779502196
-----
Iteration Number: 8336
Loss: 28.83198183978043
12 norm of gradients: 0.183138682373936
12 norm of weights: 5.664436472692918
-----
Iteration Number: 8337
Loss: 28.830300131078133
12 norm of gradients: 0.18312397980533315
12 norm of weights: 5.664370170789813
-----
Iteration Number: 8338
Loss: 28.828618692128018
12 norm of gradients: 0.18310927995391252
12 norm of weights: 5.66430387379221
-----
Iteration Number: 8339
Loss: 28.82693752286347
12 norm of gradients: 0.1830945828186972
12 norm of weights: 5.664237581699433
-----
Iteration Number: 8340
Loss: 28.82525662321314
12 norm of gradients: 0.18307988839871076
12 norm of weights: 5.66417129451081
-----
Iteration Number: 8341
Loss: 28.82357599310108
12 norm of gradients: 0.18306519669297733
12 norm of weights: 5.664105012225667
-----
Iteration Number: 8342
Loss: 28.82189563245718
12 norm of gradients: 0.18305050770052153
12 norm of weights: 5.664038734843333
-----
Iteration Number: 8343
Loss: 28.82021554121324
12 norm of gradients: 0.18303582142036876
12 norm of weights: 5.663972462363136
-----
Iteration Number: 8344

```

Loss: 28.81853571928964
l2 norm of gradients: 0.18302113785154472
l2 norm of weights: 5.663906194784402

Iteration Number: 8345
Loss: 28.816856166622422
l2 norm of gradients: 0.18300645699307583
l2 norm of weights: 5.66383993210646

Iteration Number: 8346
Loss: 28.815176883135248
l2 norm of gradients: 0.18299177884398904
l2 norm of weights: 5.663773674328636

Iteration Number: 8347
Loss: 28.813497868758684
l2 norm of gradients: 0.18297710340331189
l2 norm of weights: 5.663707421450261

Iteration Number: 8348
Loss: 28.811819123420097
l2 norm of gradients: 0.18296243067007237
l2 norm of weights: 5.66364117347066

Iteration Number: 8349
Loss: 28.810140647048843
l2 norm of gradients: 0.182947760643299
l2 norm of weights: 5.6635749303891645

Iteration Number: 8350
Loss: 28.808462439571173
l2 norm of gradients: 0.18293309332202115
l2 norm of weights: 5.663508692205102

Iteration Number: 8351
Loss: 28.806784500917335
l2 norm of gradients: 0.18291842870526842
l2 norm of weights: 5.6634424589178005

Iteration Number: 8352
Loss: 28.80510683101556
l2 norm of gradients: 0.18290376679207113
l2 norm of weights: 5.66337623052659

Iteration Number: 8353
Loss: 28.80342942979187
l2 norm of gradients: 0.18288910758146007
l2 norm of weights: 5.6633100070308

```

-----
Iteration Number: 8354
Loss: 28.801752297180133
l2 norm of gradients: 0.18287445107246672
l2 norm of weights: 5.663243788429758
-----
Iteration Number: 8355
Loss: 28.800075433101632
l2 norm of gradients: 0.18285979726412302
l2 norm of weights: 5.663177574722795
-----
Iteration Number: 8356
Loss: 28.798398837493703
l2 norm of gradients: 0.18284514615546135
l2 norm of weights: 5.663111365909239
-----
Iteration Number: 8357
Loss: 28.79672251027565
l2 norm of gradients: 0.1828304977455149
l2 norm of weights: 5.663045161988422
-----
Iteration Number: 8358
Loss: 28.79504645138352
l2 norm of gradients: 0.18281585203331727
l2 norm of weights: 5.662978962959672
-----
Iteration Number: 8359
Loss: 28.793370660742884
l2 norm of gradients: 0.18280120901790262
l2 norm of weights: 5.662912768822318
-----
Iteration Number: 8360
Loss: 28.791695138281323
l2 norm of gradients: 0.18278656869830562
l2 norm of weights: 5.662846579575693
-----
Iteration Number: 8361
Loss: 28.790019883930768
l2 norm of gradients: 0.18277193107356166
l2 norm of weights: 5.6627803952191265
-----
Iteration Number: 8362
Loss: 28.788344897615968
l2 norm of gradients: 0.1827572961427065
l2 norm of weights: 5.662714215751948
-----
Iteration Number: 8363
Loss: 28.786670179267034

```

```

12 norm of gradients: 0.18274266390477653
12 norm of weights: 5.662648041173489
-----
Iteration Number: 8364
Loss: 28.784995728816583
12 norm of gradients: 0.18272803435880872
12 norm of weights: 5.66258187148308
-----
Iteration Number: 8365
Loss: 28.783321546189615
12 norm of gradients: 0.18271340750384055
12 norm of weights: 5.662515706680052
-----
Iteration Number: 8366
Loss: 28.781647631315693
12 norm of gradients: 0.18269878333891007
12 norm of weights: 5.662449546763737
-----
Iteration Number: 8367
Loss: 28.779973984123156
12 norm of gradients: 0.1826841618630559
12 norm of weights: 5.662383391733466
-----
Iteration Number: 8368
Loss: 28.778300604543407
12 norm of gradients: 0.1826695430753171
12 norm of weights: 5.662317241588569
-----
Iteration Number: 8369
Loss: 28.776627492500904
12 norm of gradients: 0.1826549269747335
12 norm of weights: 5.6622510963283785
-----
Iteration Number: 8370
Loss: 28.774954647928233
12 norm of gradients: 0.18264031356034527
12 norm of weights: 5.6621849559522275
-----
Iteration Number: 8371
Loss: 28.773282070754252
12 norm of gradients: 0.18262570283119325
12 norm of weights: 5.662118820459447
-----
Iteration Number: 8372
Loss: 28.771609760908902
12 norm of gradients: 0.1826110947863188
12 norm of weights: 5.662052689849368
-----

```

Iteration Number: 8373
Loss: 28.769937718317227
l2 norm of gradients: 0.18259648942476375
l2 norm of weights: 5.661986564121324

Iteration Number: 8374
Loss: 28.76826594291041
l2 norm of gradients: 0.1825818867455707
l2 norm of weights: 5.6619204432746475

Iteration Number: 8375
Loss: 28.76659443461833
l2 norm of gradients: 0.18256728674778253
l2 norm of weights: 5.66185432730867

Iteration Number: 8376
Loss: 28.764923193371395
l2 norm of gradients: 0.18255268943044287
l2 norm of weights: 5.6617882162227255

Iteration Number: 8377
Loss: 28.76325221909525
l2 norm of gradients: 0.18253809479259572
l2 norm of weights: 5.661722110016147

Iteration Number: 8378
Loss: 28.76158151172306
l2 norm of gradients: 0.18252350283328586
l2 norm of weights: 5.661656008688266

Iteration Number: 8379
Loss: 28.75991107118001
l2 norm of gradients: 0.1825089135515584
l2 norm of weights: 5.6615899122384175

Iteration Number: 8380
Loss: 28.758240897396657
l2 norm of gradients: 0.18249432694645915
l2 norm of weights: 5.661523820665933

Iteration Number: 8381
Loss: 28.756570990306138
l2 norm of gradients: 0.1824797430170344
l2 norm of weights: 5.661457733970147

Iteration Number: 8382
Loss: 28.754901349832355
l2 norm of gradients: 0.18246516176233094


```

12 norm of weights: 5.661391652150395
-----
Iteration Number: 8383
Loss: 28.753231975906974
12 norm of gradients: 0.18245058318139623
12 norm of weights: 5.661325575206007
-----
Iteration Number: 8384
Loss: 28.751562868460002
12 norm of gradients: 0.18243600727327816
12 norm of weights: 5.6612595031363195
-----
Iteration Number: 8385
Loss: 28.749894027421206
12 norm of gradients: 0.18242143403702524
12 norm of weights: 5.661193435940666
-----
Iteration Number: 8386
Loss: 28.748225452718486
12 norm of gradients: 0.18240686347168653
12 norm of weights: 5.661127373618381
-----
Iteration Number: 8387
Loss: 28.746557144281056
12 norm of gradients: 0.18239229557631154
12 norm of weights: 5.661061316168799
-----
Iteration Number: 8388
Loss: 28.744889102040073
12 norm of gradients: 0.18237773034995045
12 norm of weights: 5.660995263591255
-----
Iteration Number: 8389
Loss: 28.743221325926893
12 norm of gradients: 0.18236316779165393
12 norm of weights: 5.6609292158850835
-----
Iteration Number: 8390
Loss: 28.741553815864396
12 norm of gradients: 0.1823486079004732
12 norm of weights: 5.660863173049618
-----
Iteration Number: 8391
Loss: 28.739886571789594
12 norm of gradients: 0.18233405067545994
12 norm of weights: 5.660797135084196
-----
Iteration Number: 8392

```

```

Loss: 28.738219593627278
l2 norm of gradients: 0.1823194961156665
l2 norm of weights: 5.6607311019881505
-----
Iteration Number: 8393
Loss: 28.73655288131
l2 norm of gradients: 0.1823049442201458
l2 norm of weights: 5.660665073760819
-----
Iteration Number: 8394
Loss: 28.734886434763258
l2 norm of gradients: 0.1822903949879511
l2 norm of weights: 5.660599050401535
-----
Iteration Number: 8395
Loss: 28.733220253922514
l2 norm of gradients: 0.1822758484181365
l2 norm of weights: 5.660533031909636
-----
Iteration Number: 8396
Loss: 28.73155433871371
l2 norm of gradients: 0.18226130450975633
l2 norm of weights: 5.660467018284456
-----
Iteration Number: 8397
Loss: 28.729888689068375
l2 norm of gradients: 0.1822467632618657
l2 norm of weights: 5.660401009525334
-----
Iteration Number: 8398
Loss: 28.72822330491495
l2 norm of gradients: 0.18223222467352013
l2 norm of weights: 5.660335005631603
-----
Iteration Number: 8399
Loss: 28.72655818618275
l2 norm of gradients: 0.18221768874377575
l2 norm of weights: 5.660269006602601
-----
Iteration Number: 8400
Loss: 28.72489333280499
l2 norm of gradients: 0.18220315547168914
l2 norm of weights: 5.6602030124376626
-----
Iteration Number: 8401
Loss: 28.723228744709612
l2 norm of gradients: 0.18218862485631765
l2 norm of weights: 5.660137023136127

```

```

-----
Iteration Number: 8402
Loss: 28.721564421823448
l2 norm of gradients: 0.18217409689671882
l2 norm of weights: 5.66007103869733
-----
Iteration Number: 8403
Loss: 28.719900364082058
l2 norm of gradients: 0.18215957159195106
l2 norm of weights: 5.660005059120609
-----
Iteration Number: 8404
Loss: 28.718236571409438
l2 norm of gradients: 0.18214504894107314
l2 norm of weights: 5.6599390844053
-----
Iteration Number: 8405
Loss: 28.716573043741217
l2 norm of gradients: 0.18213052894314444
l2 norm of weights: 5.6598731145507415
-----
Iteration Number: 8406
Loss: 28.71490978100355
l2 norm of gradients: 0.1821160115972248
l2 norm of weights: 5.659807149556269
-----
Iteration Number: 8407
Loss: 28.713246783130288
l2 norm of gradients: 0.1821014969023747
l2 norm of weights: 5.659741189421223
-----
Iteration Number: 8408
Loss: 28.711584050046728
l2 norm of gradients: 0.182086984857655
l2 norm of weights: 5.659675234144939
-----
Iteration Number: 8409
Loss: 28.70992158168758
l2 norm of gradients: 0.18207247546212743
l2 norm of weights: 5.659609283726756
-----
Iteration Number: 8410
Loss: 28.70825937797885
l2 norm of gradients: 0.18205796871485394
l2 norm of weights: 5.659543338166011
-----
Iteration Number: 8411
Loss: 28.706597438855464

```

```

12 norm of gradients: 0.18204346461489707
12 norm of weights: 5.659477397462043
-----
Iteration Number: 8412
Loss: 28.704935764242602
12 norm of gradients: 0.18202896316131995
12 norm of weights: 5.65941146161419
-----
Iteration Number: 8413
Loss: 28.70327435407434
12 norm of gradients: 0.18201446435318627
12 norm of weights: 5.659345530621791
-----
Iteration Number: 8414
Loss: 28.701613208277124
12 norm of gradients: 0.18199996818956027
12 norm of weights: 5.659279604484184
-----
Iteration Number: 8415
Loss: 28.699952326785997
12 norm of gradients: 0.18198547466950668
12 norm of weights: 5.659213683200709
-----
Iteration Number: 8416
Loss: 28.698291709527684
12 norm of gradients: 0.18197098379209076
12 norm of weights: 5.659147766770703
-----
Iteration Number: 8417
Loss: 28.69663135643428
12 norm of gradients: 0.18195649555637833
12 norm of weights: 5.659081855193507
-----
Iteration Number: 8418
Loss: 28.69497126743625
12 norm of gradients: 0.18194200996143572
12 norm of weights: 5.659015948468459
-----
Iteration Number: 8419
Loss: 28.693311442461294
12 norm of gradients: 0.1819275270063299
12 norm of weights: 5.658950046594899
-----
Iteration Number: 8420
Loss: 28.69165188144547
12 norm of gradients: 0.18191304669012817
12 norm of weights: 5.658884149572168
-----

```

Iteration Number: 8421
Loss: 28.689992584311305
l2 norm of gradients: 0.1818985690118986
l2 norm of weights: 5.658818257399603

Iteration Number: 8422
Loss: 28.68833355099678
l2 norm of gradients: 0.18188409397070965
l2 norm of weights: 5.658752370076545

Iteration Number: 8423
Loss: 28.68667478142794
l2 norm of gradients: 0.1818696215656303
l2 norm of weights: 5.658686487602336

Iteration Number: 8424
Loss: 28.685016275537542
l2 norm of gradients: 0.18185515179573017
l2 norm of weights: 5.658620609976313

Iteration Number: 8425
Loss: 28.683358033255498
l2 norm of gradients: 0.18184068466007938
l2 norm of weights: 5.658554737197819

Iteration Number: 8426
Loss: 28.681700054512
l2 norm of gradients: 0.18182622015774846
l2 norm of weights: 5.6584888692661925

Iteration Number: 8427
Loss: 28.680042339238465
l2 norm of gradients: 0.18181175828780868
l2 norm of weights: 5.658423006180775

Iteration Number: 8428
Loss: 28.67838488736394
l2 norm of gradients: 0.18179729904933173
l2 norm of weights: 5.658357147940909

Iteration Number: 8429
Loss: 28.676727698822827
l2 norm of gradients: 0.1817828424413898
l2 norm of weights: 5.658291294545933

Iteration Number: 8430
Loss: 28.675070773539648
l2 norm of gradients: 0.18176838846305562

```

12 norm of weights: 5.658225445995188
-----
Iteration Number: 8431
Loss: 28.673414111450583
12 norm of gradients: 0.18175393711340262
12 norm of weights: 5.658159602288018
-----
Iteration Number: 8432
Loss: 28.6717577124851
12 norm of gradients: 0.1817394883915045
12 norm of weights: 5.6580937634237625
-----
Iteration Number: 8433
Loss: 28.67010157657268
12 norm of gradients: 0.18172504229643568
12 norm of weights: 5.658027929401764
-----
Iteration Number: 8434
Loss: 28.668445703647105
12 norm of gradients: 0.18171059882727106
12 norm of weights: 5.657962100221363
-----
Iteration Number: 8435
Loss: 28.666790093635978
12 norm of gradients: 0.181696157983086
12 norm of weights: 5.657896275881901
-----
Iteration Number: 8436
Loss: 28.665134746470684
12 norm of gradients: 0.18168171976295658
12 norm of weights: 5.657830456382723
-----
Iteration Number: 8437
Loss: 28.66347966208417
12 norm of gradients: 0.18166728416595915
12 norm of weights: 5.657764641723169
-----
Iteration Number: 8438
Loss: 28.66182484040612
12 norm of gradients: 0.1816528511911708
12 norm of weights: 5.657698831902582
-----
Iteration Number: 8439
Loss: 28.660170281366163
12 norm of gradients: 0.1816384208376691
12 norm of weights: 5.657633026920304
-----
Iteration Number: 8440

```

Loss: 28.658515984897818
l2 norm of gradients: 0.18162399310453203
l2 norm of weights: 5.657567226775678

Iteration Number: 8441
Loss: 28.656861950929144
l2 norm of gradients: 0.1816095679908383
l2 norm of weights: 5.657501431468047

Iteration Number: 8442
Loss: 28.655208179394698
l2 norm of gradients: 0.18159514549566697
l2 norm of weights: 5.657435640996753

Iteration Number: 8443
Loss: 28.65355467022363
l2 norm of gradients: 0.18158072561809774
l2 norm of weights: 5.657369855361141

Iteration Number: 8444
Loss: 28.651901423346636
l2 norm of gradients: 0.18156630835721077
l2 norm of weights: 5.657304074560553

Iteration Number: 8445
Loss: 28.650248438698846
l2 norm of gradients: 0.18155189371208683
l2 norm of weights: 5.657238298594333

Iteration Number: 8446
Loss: 28.64859571620464
l2 norm of gradients: 0.1815374816818071
l2 norm of weights: 5.657172527461824

Iteration Number: 8447
Loss: 28.646943255800092
l2 norm of gradients: 0.18152307226545344
l2 norm of weights: 5.65710676116237

Iteration Number: 8448
Loss: 28.64529105741418
l2 norm of gradients: 0.18150866546210812
l2 norm of weights: 5.657040999695314

Iteration Number: 8449
Loss: 28.643639120979554
l2 norm of gradients: 0.1814942612708539
l2 norm of weights: 5.656975243060002

```

-----
Iteration Number: 8450
Loss: 28.641987446429972
l2 norm of gradients: 0.1814798596907742
l2 norm of weights: 5.656909491255778
-----
Iteration Number: 8451
Loss: 28.640336033690634
l2 norm of gradients: 0.1814654607209529
l2 norm of weights: 5.656843744281985
-----
Iteration Number: 8452
Loss: 28.638684882695998
l2 norm of gradients: 0.1814510643604744
l2 norm of weights: 5.656778002137967
-----
Iteration Number: 8453
Loss: 28.63703399337893
l2 norm of gradients: 0.18143667060842358
l2 norm of weights: 5.656712264823072
-----
Iteration Number: 8454
Loss: 28.63538336566726
l2 norm of gradients: 0.181422279463886
l2 norm of weights: 5.656646532336642
-----
Iteration Number: 8455
Loss: 28.63373299949942
l2 norm of gradients: 0.18140789092594758
l2 norm of weights: 5.656580804678022
-----
Iteration Number: 8456
Loss: 28.63208289479729
l2 norm of gradients: 0.18139350499369483
l2 norm of weights: 5.656515081846558
-----
Iteration Number: 8457
Loss: 28.630433051498166
l2 norm of gradients: 0.1813791216662148
l2 norm of weights: 5.656449363841595
-----
Iteration Number: 8458
Loss: 28.62878346953451
l2 norm of gradients: 0.18136474094259503
l2 norm of weights: 5.6563836506624785
-----
Iteration Number: 8459
Loss: 28.62713414883468

```



```

12 norm of gradients: 0.18135036282192357
12 norm of weights: 5.656317942308554
-----
Iteration Number: 8460
Loss: 28.625485089332233
12 norm of gradients: 0.18133598730328912
12 norm of weights: 5.656252238779167
-----
Iteration Number: 8461
Loss: 28.623836290955712
12 norm of gradients: 0.1813216143857807
12 norm of weights: 5.656186540073664
-----
Iteration Number: 8462
Loss: 28.62218775364275
12 norm of gradients: 0.18130724406848803
12 norm of weights: 5.65612084619139
-----
Iteration Number: 8463
Loss: 28.620539477319163
12 norm of gradients: 0.1812928763505012
12 norm of weights: 5.656055157131693
-----
Iteration Number: 8464
Loss: 28.61889146191689
12 norm of gradients: 0.18127851123091104
12 norm of weights: 5.655989472893918
-----
Iteration Number: 8465
Loss: 28.617243707374183
12 norm of gradients: 0.18126414870880864
12 norm of weights: 5.655923793477411
-----
Iteration Number: 8466
Loss: 28.615596213615493
12 norm of gradients: 0.1812497887832858
12 norm of weights: 5.655858118881521
-----
Iteration Number: 8467
Loss: 28.613948980575437
12 norm of gradients: 0.18123543145343476
12 norm of weights: 5.655792449105592
-----
Iteration Number: 8468
Loss: 28.612302008183462
12 norm of gradients: 0.1812210767183483
12 norm of weights: 5.655726784148972
-----

```

Iteration Number: 8469
Loss: 28.610655296377715
l2 norm of gradients: 0.18120672457711973
l2 norm of weights: 5.655661124011009

Iteration Number: 8470
Loss: 28.609008845084336
l2 norm of gradients: 0.18119237502884283
l2 norm of weights: 5.6555954686910495

Iteration Number: 8471
Loss: 28.607362654235274
l2 norm of gradients: 0.181178028072612
l2 norm of weights: 5.655529818188441

Iteration Number: 8472
Loss: 28.605716723767745
l2 norm of gradients: 0.1811636837075221
l2 norm of weights: 5.655464172502531

Iteration Number: 8473
Loss: 28.60407105360484
l2 norm of gradients: 0.1811493419326684
l2 norm of weights: 5.655398531632667

Iteration Number: 8474
Loss: 28.602425643687397
l2 norm of gradients: 0.18113500274714694
l2 norm of weights: 5.655332895578197

Iteration Number: 8475
Loss: 28.600780493943848
l2 norm of gradients: 0.1811206661500541
l2 norm of weights: 5.655267264338469

Iteration Number: 8476
Loss: 28.599135604302916
l2 norm of gradients: 0.18110633214048677
l2 norm of weights: 5.655201637912832

Iteration Number: 8477
Loss: 28.597490974701643
l2 norm of gradients: 0.18109200071754245
l2 norm of weights: 5.655136016300633

Iteration Number: 8478
Loss: 28.59584660506917
l2 norm of gradients: 0.1810776718803191

```

12 norm of weights: 5.655070399501221
-----
Iteration Number: 8479
Loss: 28.59420249534082
12 norm of gradients: 0.18106334562791515
12 norm of weights: 5.655004787513945
-----
Iteration Number: 8480
Loss: 28.592558645445255
12 norm of gradients: 0.18104902195942973
12 norm of weights: 5.654939180338153
-----
Iteration Number: 8481
Loss: 28.59091505531523
12 norm of gradients: 0.18103470087396223
12 norm of weights: 5.654873577973195
-----
Iteration Number: 8482
Loss: 28.589271724884373
12 norm of gradients: 0.18102038237061277
12 norm of weights: 5.654807980418419
-----
Iteration Number: 8483
Loss: 28.58762865408193
12 norm of gradients: 0.1810060664484819
12 norm of weights: 5.654742387673174
-----
Iteration Number: 8484
Loss: 28.5859858428449
12 norm of gradients: 0.1809917531066707
12 norm of weights: 5.654676799736811
-----
Iteration Number: 8485
Loss: 28.584343291103192
12 norm of gradients: 0.18097744234428076
12 norm of weights: 5.6546112166086795
-----
Iteration Number: 8486
Loss: 28.58270099878753
12 norm of gradients: 0.18096313416041412
12 norm of weights: 5.654545638288127
-----
Iteration Number: 8487
Loss: 28.581058965832217
12 norm of gradients: 0.18094882855417352
12 norm of weights: 5.654480064774505
-----
Iteration Number: 8488

```

Loss: 28.57941719216796
l2 norm of gradients: 0.180934525524662
l2 norm of weights: 5.6544144960671625

Iteration Number: 8489
Loss: 28.57777567773084
l2 norm of gradients: 0.18092022507098315
l2 norm of weights: 5.654348932165451

Iteration Number: 8490
Loss: 28.576134422449407
l2 norm of gradients: 0.18090592719224127
l2 norm of weights: 5.654283373068719

Iteration Number: 8491
Loss: 28.574493426255827
l2 norm of gradients: 0.18089163188754095
l2 norm of weights: 5.65421781877632

Iteration Number: 8492
Loss: 28.57285268908478
l2 norm of gradients: 0.18087733915598747
l2 norm of weights: 5.654152269287601

Iteration Number: 8493
Loss: 28.57121221086745
l2 norm of gradients: 0.18086304899668637
l2 norm of weights: 5.654086724601913

Iteration Number: 8494
Loss: 28.569571991539338
l2 norm of gradients: 0.18084876140874404
l2 norm of weights: 5.65402118471861

Iteration Number: 8495
Loss: 28.56793203102797
l2 norm of gradients: 0.1808344763912671
l2 norm of weights: 5.653955649637041

Iteration Number: 8496
Loss: 28.566292329269547
l2 norm of gradients: 0.18082019394336282
l2 norm of weights: 5.653890119356556

Iteration Number: 8497
Loss: 28.564652886197173
l2 norm of gradients: 0.18080591406413904
l2 norm of weights: 5.653824593876509

```

-----
Iteration Number: 8498
Loss: 28.563013701740072
l2 norm of gradients: 0.18079163675270385
l2 norm of weights: 5.65375907319625
-----
Iteration Number: 8499
Loss: 28.561374775833393
l2 norm of gradients: 0.1807773620081662
l2 norm of weights: 5.653693557315131
-----
Iteration Number: 8500
Loss: 28.55973610840865
l2 norm of gradients: 0.18076308982963527
l2 norm of weights: 5.653628046232503
-----
Iteration Number: 8501
Loss: 28.55809769939946
l2 norm of gradients: 0.18074882021622082
l2 norm of weights: 5.653562539947718
-----
Iteration Number: 8502
Loss: 28.55645954873879
l2 norm of gradients: 0.18073455316703324
l2 norm of weights: 5.65349703846013
-----
Iteration Number: 8503
Loss: 28.55482165635723
l2 norm of gradients: 0.18072028868118337
l2 norm of weights: 5.653431541769089
-----
Iteration Number: 8504
Loss: 28.553184022189555
l2 norm of gradients: 0.18070602675778244
l2 norm of weights: 5.653366049873947
-----
Iteration Number: 8505
Loss: 28.55154664616987
l2 norm of gradients: 0.1806917673959424
l2 norm of weights: 5.653300562774059
-----
Iteration Number: 8506
Loss: 28.5499095282289
l2 norm of gradients: 0.18067751059477552
l2 norm of weights: 5.653235080468776
-----
Iteration Number: 8507
Loss: 28.548272668299564

```

```

12 norm of gradients: 0.18066325635339467
12 norm of weights: 5.653169602957452
-----
Iteration Number: 8508
Loss: 28.546636066314598
12 norm of gradients: 0.18064900467091322
12 norm of weights: 5.653104130239438
-----
Iteration Number: 8509
Loss: 28.54499972220909
12 norm of gradients: 0.18063475554644506
12 norm of weights: 5.653038662314089
-----
Iteration Number: 8510
Loss: 28.543363635911998
12 norm of gradients: 0.1806205089791045
12 norm of weights: 5.6529731991807575
-----
Iteration Number: 8511
Loss: 28.541727807360804
12 norm of gradients: 0.18060626496800655
12 norm of weights: 5.652907740838797
-----
Iteration Number: 8512
Loss: 28.54009223648554
12 norm of gradients: 0.18059202351226653
12 norm of weights: 5.652842287287561
-----
Iteration Number: 8513
Loss: 28.538456923221162
12 norm of gradients: 0.18057778461100027
12 norm of weights: 5.652776838526404
-----
Iteration Number: 8514
Loss: 28.536821867499278
12 norm of gradients: 0.1805635482633243
12 norm of weights: 5.652711394554679
-----
Iteration Number: 8515
Loss: 28.535187069252142
12 norm of gradients: 0.1805493144683555
12 norm of weights: 5.65264595537174
-----
Iteration Number: 8516
Loss: 28.533552528416067
12 norm of gradients: 0.1805350832252113
12 norm of weights: 5.652580520976941
-----

```

Iteration Number: 8517
Loss: 28.531918244922444
l2 norm of gradients: 0.18052085453300956
l2 norm of weights: 5.6525150913696365

Iteration Number: 8518
Loss: 28.530284218700626
l2 norm of gradients: 0.1805066283908688
l2 norm of weights: 5.652449666549182

Iteration Number: 8519
Loss: 28.52865044969009
l2 norm of gradients: 0.18049240479790793
l2 norm of weights: 5.652384246514931

Iteration Number: 8520
Loss: 28.527016937822875
l2 norm of gradients: 0.18047818375324626
l2 norm of weights: 5.652318831266238

Iteration Number: 8521
Loss: 28.525383683029165
l2 norm of gradients: 0.18046396525600397
l2 norm of weights: 5.6522534208024595

Iteration Number: 8522
Loss: 28.523750685242206
l2 norm of gradients: 0.1804497493053014
l2 norm of weights: 5.652188015122949

Iteration Number: 8523
Loss: 28.522117944400957
l2 norm of gradients: 0.1804355359002595
l2 norm of weights: 5.652122614227063

Iteration Number: 8524
Loss: 28.520485460432482
l2 norm of gradients: 0.18042132503999972
l2 norm of weights: 5.652057218114155

Iteration Number: 8525
Loss: 28.518853233272793
l2 norm of gradients: 0.18040711672364398
l2 norm of weights: 5.651991826783583

Iteration Number: 8526
Loss: 28.517221262853663
l2 norm of gradients: 0.18039291095031487

```

12 norm of weights: 5.651926440234701
-----
Iteration Number: 8527
Loss: 28.515589549111095
12 norm of gradients: 0.18037870771913528
12 norm of weights: 5.651861058466864
-----
Iteration Number: 8528
Loss: 28.513958091977155
12 norm of gradients: 0.18036450702922865
12 norm of weights: 5.65179568147943
-----
Iteration Number: 8529
Loss: 28.512326891384337
12 norm of gradients: 0.18035030887971903
12 norm of weights: 5.651730309271754
-----
Iteration Number: 8530
Loss: 28.510695947268655
12 norm of gradients: 0.1803361132697308
12 norm of weights: 5.651664941843192
-----
Iteration Number: 8531
Loss: 28.50906525956095
12 norm of gradients: 0.180321920198389
12 norm of weights: 5.651599579193101
-----
Iteration Number: 8532
Loss: 28.50743482819615
12 norm of gradients: 0.18030772966481912
12 norm of weights: 5.651534221320838
-----
Iteration Number: 8533
Loss: 28.505804653106264
12 norm of gradients: 0.18029354166814707
12 norm of weights: 5.651468868225758
-----
Iteration Number: 8534
Loss: 28.504174734229288
12 norm of gradients: 0.1802793562074994
12 norm of weights: 5.65140351990722
-----
Iteration Number: 8535
Loss: 28.502545071495327
12 norm of gradients: 0.18026517328200303
12 norm of weights: 5.651338176364579
-----
Iteration Number: 8536

```


Loss: 28.500915664837727
l2 norm of gradients: 0.18025099289078547
l2 norm of weights: 5.651272837597194

Iteration Number: 8537
Loss: 28.499286514189624
l2 norm of gradients: 0.18023681503297476
l2 norm of weights: 5.65120750360442

Iteration Number: 8538
Loss: 28.497657619489154
l2 norm of gradients: 0.18022263970769917
l2 norm of weights: 5.651142174385616

Iteration Number: 8539
Loss: 28.49602898066346
l2 norm of gradients: 0.1802084669140879
l2 norm of weights: 5.65107684994014

Iteration Number: 8540
Loss: 28.494400597652575
l2 norm of gradients: 0.18019429665127024
l2 norm of weights: 5.651011530267349

Iteration Number: 8541
Loss: 28.492772470386612
l2 norm of gradients: 0.18018012891837634
l2 norm of weights: 5.650946215366601

Iteration Number: 8542
Loss: 28.491144598801228
l2 norm of gradients: 0.18016596371453658
l2 norm of weights: 5.650880905237254

Iteration Number: 8543
Loss: 28.489516982826913
l2 norm of gradients: 0.18015180103888187
l2 norm of weights: 5.6508155998786656

Iteration Number: 8544
Loss: 28.487889622401546
l2 norm of gradients: 0.18013764089054377
l2 norm of weights: 5.650750299290196

Iteration Number: 8545
Loss: 28.486262517458574
l2 norm of gradients: 0.18012348326865427
l2 norm of weights: 5.6506850034712

```
-----
Iteration Number: 8546
Loss: 28.48463566793038
l2 norm of gradients: 0.1801093281723457
l2 norm of weights: 5.650619712421041
-----
Iteration Number: 8547
Loss: 28.48300907375083
l2 norm of gradients: 0.18009517560075106
l2 norm of weights: 5.650554426139074
-----
Iteration Number: 8548
Loss: 28.481382734853774
l2 norm of gradients: 0.18008102555300387
l2 norm of weights: 5.650489144624659
-----
Iteration Number: 8549
Loss: 28.479756651175265
l2 norm of gradients: 0.180066878028238
l2 norm of weights: 5.650423867877157
-----
Iteration Number: 8550
Loss: 28.478130822646012
l2 norm of gradients: 0.18005273302558794
l2 norm of weights: 5.6503585958959235
-----
Iteration Number: 8551
Loss: 28.47650524920451
l2 norm of gradients: 0.18003859054418858
l2 norm of weights: 5.65029332868032
-----
Iteration Number: 8552
Loss: 28.47487993078083
l2 norm of gradients: 0.18002445058317545
l2 norm of weights: 5.650228066229706
-----
Iteration Number: 8553
Loss: 28.47325486730926
l2 norm of gradients: 0.18001031314168436
l2 norm of weights: 5.650162808543442
-----
Iteration Number: 8554
Loss: 28.4716300587275
l2 norm of gradients: 0.17999617821885178
l2 norm of weights: 5.650097555620886
-----
Iteration Number: 8555
Loss: 28.47000550496747
```

```

12 norm of gradients: 0.1799820458138146
12 norm of weights: 5.650032307461399
-----
Iteration Number: 8556
Loss: 28.468381205961123
12 norm of gradients: 0.17996791592571035
12 norm of weights: 5.649967064064341
-----
Iteration Number: 8557
Loss: 28.466757161646196
12 norm of gradients: 0.17995378855367672
12 norm of weights: 5.649901825429071
-----
Iteration Number: 8558
Loss: 28.465133371956984
12 norm of gradients: 0.1799396636968523
12 norm of weights: 5.649836591554951
-----
Iteration Number: 8559
Loss: 28.463509836826812
12 norm of gradients: 0.17992554135437586
12 norm of weights: 5.649771362441341
-----
Iteration Number: 8560
Loss: 28.46188655618736
12 norm of gradients: 0.17991142152538683
12 norm of weights: 5.649706138087601
-----
Iteration Number: 8561
Loss: 28.460263529975812
12 norm of gradients: 0.1798973042090251
12 norm of weights: 5.649640918493093
-----
Iteration Number: 8562
Loss: 28.458640758125586
12 norm of gradients: 0.17988318940443096
12 norm of weights: 5.6495757036571765
-----
Iteration Number: 8563
Loss: 28.457018240572353
12 norm of gradients: 0.17986907711074535
12 norm of weights: 5.649510493579215
-----
Iteration Number: 8564
Loss: 28.455395977249392
12 norm of gradients: 0.1798549673271096
12 norm of weights: 5.649445288258567
-----

```

Iteration Number: 8565
Loss: 28.453773968091063
l2 norm of gradients: 0.17984086005266547
l2 norm of weights: 5.649380087694595

Iteration Number: 8566
Loss: 28.452152213032495
l2 norm of gradients: 0.17982675528655537
l2 norm of weights: 5.649314891886662

Iteration Number: 8567
Loss: 28.450530712005907
l2 norm of gradients: 0.17981265302792207
l2 norm of weights: 5.649249700834127

Iteration Number: 8568
Loss: 28.448909464949203
l2 norm of gradients: 0.17979855327590893
l2 norm of weights: 5.649184514536354

Iteration Number: 8569
Loss: 28.44728847179269
l2 norm of gradients: 0.17978445602965967
l2 norm of weights: 5.649119332992704

Iteration Number: 8570
Loss: 28.445667732477425
l2 norm of gradients: 0.17977036128831864
l2 norm of weights: 5.649054156202539

Iteration Number: 8571
Loss: 28.44404724693128
l2 norm of gradients: 0.17975626905103068
l2 norm of weights: 5.648988984165222

Iteration Number: 8572
Loss: 28.442427015093678
l2 norm of gradients: 0.17974217931694086
l2 norm of weights: 5.648923816880115

Iteration Number: 8573
Loss: 28.440807036896476
l2 norm of gradients: 0.17972809208519513
l2 norm of weights: 5.64885865434658

Iteration Number: 8574
Loss: 28.43918731227395
l2 norm of gradients: 0.17971400735493961

```
l2 norm of weights: 5.6487934965639806
-----
Iteration Number: 8575
Loss: 28.43756784116268
l2 norm of gradients: 0.17969992512532107
l2 norm of weights: 5.64872834353168
-----
Iteration Number: 8576
Loss: 28.435948623496763
l2 norm of gradients: 0.17968584539548674
l2 norm of weights: 5.648663195249039
-----
Iteration Number: 8577
Loss: 28.434329659211304
l2 norm of gradients: 0.17967176816458427
l2 norm of weights: 5.648598051715424
-----
Iteration Number: 8578
Loss: 28.4327109482404
l2 norm of gradients: 0.1796576934317619
l2 norm of weights: 5.648532912930196
-----
Iteration Number: 8579
Loss: 28.431092490518346
l2 norm of gradients: 0.17964362119616828
l2 norm of weights: 5.6484677788927184
-----
Iteration Number: 8580
Loss: 28.429474285981527
l2 norm of gradients: 0.17962955145695259
l2 norm of weights: 5.648402649602357
-----
Iteration Number: 8581
Loss: 28.427856334561632
l2 norm of gradients: 0.17961548421326448
l2 norm of weights: 5.648337525058473
-----
Iteration Number: 8582
Loss: 28.426238636197976
l2 norm of gradients: 0.17960141946425406
l2 norm of weights: 5.648272405260431
-----
Iteration Number: 8583
Loss: 28.424621190824514
l2 norm of gradients: 0.17958735720907198
l2 norm of weights: 5.648207290207595
-----
Iteration Number: 8584
```

Loss: 28.423003998372664
l2 norm of gradients: 0.17957329744686928
l2 norm of weights: 5.64814217989933

Iteration Number: 8585
Loss: 28.42138705878022
l2 norm of gradients: 0.17955924017679764
l2 norm of weights: 5.648077074335

Iteration Number: 8586
Loss: 28.419770371981834
l2 norm of gradients: 0.1795451853980091
l2 norm of weights: 5.648011973513968

Iteration Number: 8587
Loss: 28.418153937913413
l2 norm of gradients: 0.17953113310965615
l2 norm of weights: 5.647946877435602

Iteration Number: 8588
Loss: 28.416537756507367
l2 norm of gradients: 0.1795170833108919
l2 norm of weights: 5.647881786099263

Iteration Number: 8589
Loss: 28.414921827701107
l2 norm of gradients: 0.17950303600086984
l2 norm of weights: 5.647816699504317

Iteration Number: 8590
Loss: 28.413306151426823
l2 norm of gradients: 0.179488991178744
l2 norm of weights: 5.647751617650131

Iteration Number: 8591
Loss: 28.411690727624666
l2 norm of gradients: 0.17947494884366888
l2 norm of weights: 5.647686540536068

Iteration Number: 8592
Loss: 28.410075556225436
l2 norm of gradients: 0.17946090899479944
l2 norm of weights: 5.647621468161494

Iteration Number: 8593
Loss: 28.408460637165984
l2 norm of gradients: 0.17944687163129114
l2 norm of weights: 5.647556400525774

```

-----
Iteration Number: 8594
Loss: 28.40684597037942
l2 norm of gradients: 0.17943283675229985
l2 norm of weights: 5.647491337628274
-----
Iteration Number: 8595
Loss: 28.405231555804843
l2 norm of gradients: 0.17941880435698204
l2 norm of weights: 5.647426279468361
-----
Iteration Number: 8596
Loss: 28.403617393373583
l2 norm of gradients: 0.17940477444449465
l2 norm of weights: 5.647361226045398
-----
Iteration Number: 8597
Loss: 28.402003483022593
l2 norm of gradients: 0.17939074701399496
l2 norm of weights: 5.647296177358754
-----
Iteration Number: 8598
Loss: 28.40038982468795
l2 norm of gradients: 0.1793767220646409
l2 norm of weights: 5.647231133407794
-----
Iteration Number: 8599
Loss: 28.398776418305413
l2 norm of gradients: 0.17936269959559079
l2 norm of weights: 5.647166094191884
-----
Iteration Number: 8600
Loss: 28.397163263807535
l2 norm of gradients: 0.17934867960600342
l2 norm of weights: 5.647101059710391
-----
Iteration Number: 8601
Loss: 28.39555036113279
l2 norm of gradients: 0.17933466209503812
l2 norm of weights: 5.64703602996268
-----
Iteration Number: 8602
Loss: 28.393937710212345
l2 norm of gradients: 0.17932064706185472
l2 norm of weights: 5.646971004948121
-----
Iteration Number: 8603
Loss: 28.392325310983615

```

```

12 norm of gradients: 0.17930663450561338
12 norm of weights: 5.646905984666079
-----
Iteration Number: 8604
Loss: 28.390713163387158
12 norm of gradients: 0.1792926244254749
12 norm of weights: 5.646840969115921
-----
Iteration Number: 8605
Loss: 28.389101267350583
12 norm of gradients: 0.17927861682060045
12 norm of weights: 5.646775958297013
-----
Iteration Number: 8606
Loss: 28.387489622813515
12 norm of gradients: 0.17926461169015176
12 norm of weights: 5.646710952208726
-----
Iteration Number: 8607
Loss: 28.38587822971073
12 norm of gradients: 0.17925060903329096
12 norm of weights: 5.646645950850423
-----
Iteration Number: 8608
Loss: 28.384267087977324
12 norm of gradients: 0.17923660884918072
12 norm of weights: 5.646580954221475
-----
Iteration Number: 8609
Loss: 28.382656197551462
12 norm of gradients: 0.17922261113698423
12 norm of weights: 5.646515962321249
-----
Iteration Number: 8610
Loss: 28.381045558365777
12 norm of gradients: 0.1792086158958649
12 norm of weights: 5.646450975149112
-----
Iteration Number: 8611
Loss: 28.37943517035446
12 norm of gradients: 0.179194623124987
12 norm of weights: 5.646385992704433
-----
Iteration Number: 8612
Loss: 28.377825033456922
12 norm of gradients: 0.179180632823515
12 norm of weights: 5.646321014986579
-----

```


Iteration Number: 8613
Loss: 28.37621514760957
l2 norm of gradients: 0.17916664499061397
l2 norm of weights: 5.64625604199492

Iteration Number: 8614
Loss: 28.37460551274178
l2 norm of gradients: 0.17915265962544935
l2 norm of weights: 5.646191073728823

Iteration Number: 8615
Loss: 28.37299612879661
l2 norm of gradients: 0.17913867672718714
l2 norm of weights: 5.646126110187658

Iteration Number: 8616
Loss: 28.371386995705528
l2 norm of gradients: 0.1791246962949939
l2 norm of weights: 5.646061151370793

Iteration Number: 8617
Loss: 28.36977811340363
l2 norm of gradients: 0.17911071832803638
l2 norm of weights: 5.645996197277595

Iteration Number: 8618
Loss: 28.368169481829405
l2 norm of gradients: 0.1790967428254821
l2 norm of weights: 5.645931247907438

Iteration Number: 8619
Loss: 28.366561100918318
l2 norm of gradients: 0.1790827697864989
l2 norm of weights: 5.6458663032596865

Iteration Number: 8620
Loss: 28.36495297060785
l2 norm of gradients: 0.17906879921025512
l2 norm of weights: 5.645801363333713

Iteration Number: 8621
Loss: 28.36334509082822
l2 norm of gradients: 0.17905483109591963
l2 norm of weights: 5.645736428128885

Iteration Number: 8622
Loss: 28.36173746151907
l2 norm of gradients: 0.17904086544266168

```

12 norm of weights: 5.645671497644572
-----
Iteration Number: 8623
Loss: 28.360130082617935
12 norm of gradients: 0.17902690224965107
12 norm of weights: 5.645606571880145
-----
Iteration Number: 8624
Loss: 28.35852295405575
12 norm of gradients: 0.17901294151605804
12 norm of weights: 5.645541650834974
-----
Iteration Number: 8625
Loss: 28.35691607577433
12 norm of gradients: 0.17899898324105332
12 norm of weights: 5.645476734508429
-----
Iteration Number: 8626
Loss: 28.355309447704773
12 norm of gradients: 0.17898502742380804
12 norm of weights: 5.645411822899878
-----
Iteration Number: 8627
Loss: 28.353703069786043
12 norm of gradients: 0.17897107406349388
12 norm of weights: 5.645346916008695
-----
Iteration Number: 8628
Loss: 28.352096941954112
12 norm of gradients: 0.178957123159283
12 norm of weights: 5.645282013834247
-----
Iteration Number: 8629
Loss: 28.350491064144272
12 norm of gradients: 0.1789431747103479
12 norm of weights: 5.645217116375908
-----
Iteration Number: 8630
Loss: 28.348885436293358
12 norm of gradients: 0.1789292287158618
12 norm of weights: 5.645152223633046
-----
Iteration Number: 8631
Loss: 28.34728005833615
12 norm of gradients: 0.1789152851749981
12 norm of weights: 5.645087335605033
-----
Iteration Number: 8632

```

```

Loss: 28.345674930210166
12 norm of gradients: 0.17890134408693092
12 norm of weights: 5.64502245229124
-----
Iteration Number: 8633
Loss: 28.34407005184827
12 norm of gradients: 0.17888740545083467
12 norm of weights: 5.6449575736910385
-----
Iteration Number: 8634
Loss: 28.34246542319034
12 norm of gradients: 0.17887346926588432
12 norm of weights: 5.644892699803799
-----
Iteration Number: 8635
Loss: 28.34086104417414
12 norm of gradients: 0.1788595355312553
12 norm of weights: 5.644827830628895
-----
Iteration Number: 8636
Loss: 28.33925691472931
12 norm of gradients: 0.1788456042461235
12 norm of weights: 5.644762966165696
-----
Iteration Number: 8637
Loss: 28.337653034799764
12 norm of gradients: 0.17883167540966521
12 norm of weights: 5.644698106413574
-----
Iteration Number: 8638
Loss: 28.336049404316885
12 norm of gradients: 0.1788177490210573
12 norm of weights: 5.644633251371902
-----
Iteration Number: 8639
Loss: 28.334446023217694
12 norm of gradients: 0.1788038250794771
12 norm of weights: 5.6445684010400505
-----
Iteration Number: 8640
Loss: 28.332842891438702
12 norm of gradients: 0.17878990358410224
12 norm of weights: 5.644503555417394
-----
Iteration Number: 8641
Loss: 28.33124000891898
12 norm of gradients: 0.17877598453411112
12 norm of weights: 5.644438714503304

```

```

-----
Iteration Number: 8642
Loss: 28.329637375591503
l2 norm of gradients: 0.17876206792868224
l2 norm of weights: 5.644373878297151
-----
Iteration Number: 8643
Loss: 28.328034991393324
l2 norm of gradients: 0.17874815376699493
l2 norm of weights: 5.64430904679831
-----
Iteration Number: 8644
Loss: 28.32643285626111
l2 norm of gradients: 0.17873424204822871
l2 norm of weights: 5.6442442200061524
-----
Iteration Number: 8645
Loss: 28.324830970135007
l2 norm of gradients: 0.17872033277156368
l2 norm of weights: 5.644179397920053
-----
Iteration Number: 8646
Loss: 28.323229332946596
l2 norm of gradients: 0.17870642593618044
l2 norm of weights: 5.644114580539382
-----
Iteration Number: 8647
Loss: 28.32162794463258
l2 norm of gradients: 0.17869252154125995
l2 norm of weights: 5.644049767863515
-----
Iteration Number: 8648
Loss: 28.320026805130226
l2 norm of gradients: 0.17867861958598374
l2 norm of weights: 5.643984959891824
-----
Iteration Number: 8649
Loss: 28.318425914379805
l2 norm of gradients: 0.1786647200695337
l2 norm of weights: 5.643920156623683
-----
Iteration Number: 8650
Loss: 28.316825272314187
l2 norm of gradients: 0.1786508229910923
l2 norm of weights: 5.643855358058467
-----
Iteration Number: 8651
Loss: 28.315224878870996

```

```

12 norm of gradients: 0.1786369283498424
12 norm of weights: 5.643790564195546
-----
Iteration Number: 8652
Loss: 28.313624733986153
12 norm of gradients: 0.17862303614496733
12 norm of weights: 5.643725775034298
-----
Iteration Number: 8653
Loss: 28.312024837596994
12 norm of gradients: 0.1786091463756509
12 norm of weights: 5.643660990574094
-----
Iteration Number: 8654
Loss: 28.31042518964058
12 norm of gradients: 0.1785952590410774
12 norm of weights: 5.643596210814311
-----
Iteration Number: 8655
Loss: 28.308825790053234
12 norm of gradients: 0.1785813741404315
12 norm of weights: 5.643531435754321
-----
Iteration Number: 8656
Loss: 28.307226638771688
12 norm of gradients: 0.17856749167289845
12 norm of weights: 5.6434666653934995
-----
Iteration Number: 8657
Loss: 28.305627735730816
12 norm of gradients: 0.17855361163766387
12 norm of weights: 5.643401899731221
-----
Iteration Number: 8658
Loss: 28.304029080871338
12 norm of gradients: 0.17853973403391385
12 norm of weights: 5.64333713876686
-----
Iteration Number: 8659
Loss: 28.302430674128594
12 norm of gradients: 0.17852585886083505
12 norm of weights: 5.6432723824997915
-----
Iteration Number: 8660
Loss: 28.30083251543701
12 norm of gradients: 0.17851198611761449
12 norm of weights: 5.64320763092939
-----

```

Iteration Number: 8661
Loss: 28.299234604736817
l2 norm of gradients: 0.1784981158034396
l2 norm of weights: 5.643142884055032

Iteration Number: 8662
Loss: 28.29763694196287
l2 norm of gradients: 0.17848424791749837
l2 norm of weights: 5.643078141876092

Iteration Number: 8663
Loss: 28.29603952705475
l2 norm of gradients: 0.17847038245897914
l2 norm of weights: 5.643013404391946

Iteration Number: 8664
Loss: 28.294442359945315
l2 norm of gradients: 0.17845651942707094
l2 norm of weights: 5.642948671601968

Iteration Number: 8665
Loss: 28.292845440574705
l2 norm of gradients: 0.17844265882096302
l2 norm of weights: 5.642883943505537

Iteration Number: 8666
Loss: 28.2912487688786
l2 norm of gradients: 0.1784288006398452
l2 norm of weights: 5.642819220102026

Iteration Number: 8667
Loss: 28.289652344793392
l2 norm of gradients: 0.17841494488290768
l2 norm of weights: 5.642754501390811

Iteration Number: 8668
Loss: 28.288056168257494
l2 norm of gradients: 0.17840109154934133
l2 norm of weights: 5.642689787371269

Iteration Number: 8669
Loss: 28.286460239207265
l2 norm of gradients: 0.17838724063833714
l2 norm of weights: 5.6426250780427765

Iteration Number: 8670
Loss: 28.284864557580025
l2 norm of gradients: 0.17837339214908682

```

12 norm of weights: 5.64256037340471
-----
Iteration Number: 8671
Loss: 28.283269123313723
12 norm of gradients: 0.17835954608078242
12 norm of weights: 5.6424956734564455
-----
Iteration Number: 8672
Loss: 28.281673936344003
12 norm of gradients: 0.17834570243261655
12 norm of weights: 5.6424309781973605
-----
Iteration Number: 8673
Loss: 28.280078996608875
12 norm of gradients: 0.17833186120378217
12 norm of weights: 5.642366287626831
-----
Iteration Number: 8674
Loss: 28.27848430404615
12 norm of gradients: 0.17831802239347272
12 norm of weights: 5.642301601744234
-----
Iteration Number: 8675
Loss: 28.276889858591243
12 norm of gradients: 0.17830418600088216
12 norm of weights: 5.642236920548948
-----
Iteration Number: 8676
Loss: 28.275295660181484
12 norm of gradients: 0.1782903520252048
12 norm of weights: 5.642172244040348
-----
Iteration Number: 8677
Loss: 28.273701708759475
12 norm of gradients: 0.1782765204656355
12 norm of weights: 5.642107572217814
-----
Iteration Number: 8678
Loss: 28.272108004254434
12 norm of gradients: 0.1782626913213696
12 norm of weights: 5.642042905080722
-----
Iteration Number: 8679
Loss: 28.270514546608457
12 norm of gradients: 0.17824886459160275
12 norm of weights: 5.64197824262845
-----
Iteration Number: 8680

```

```

Loss: 28.268921335757337
l2 norm of gradients: 0.17823504027553116
l2 norm of weights: 5.641913584860376
-----
Iteration Number: 8681
Loss: 28.267328371640634
l2 norm of gradients: 0.17822121837235155
l2 norm of weights: 5.641848931775877
-----
Iteration Number: 8682
Loss: 28.2657356541926
l2 norm of gradients: 0.1782073988812609
l2 norm of weights: 5.641784283374333
-----
Iteration Number: 8683
Loss: 28.264143183353212
l2 norm of gradients: 0.1781935818014568
l2 norm of weights: 5.641719639655122
-----
Iteration Number: 8684
Loss: 28.262550959057872
l2 norm of gradients: 0.17817976713213735
l2 norm of weights: 5.64165500061762
-----
Iteration Number: 8685
Loss: 28.26095898124678
l2 norm of gradients: 0.1781659548725009
l2 norm of weights: 5.641590366261208
-----
Iteration Number: 8686
Loss: 28.259367249854442
l2 norm of gradients: 0.1781521450217464
l2 norm of weights: 5.641525736585264
-----
Iteration Number: 8687
Loss: 28.257775764819268
l2 norm of gradients: 0.17813833757907319
l2 norm of weights: 5.641461111589166
-----
Iteration Number: 8688
Loss: 28.2561845260806
l2 norm of gradients: 0.17812453254368116
l2 norm of weights: 5.641396491272295
-----
Iteration Number: 8689
Loss: 28.25459353357421
l2 norm of gradients: 0.17811072991477048
l2 norm of weights: 5.641331875634028

```



```

-----
Iteration Number: 8690
Loss: 28.253002787238486
l2 norm of gradients: 0.17809692969154195
l2 norm of weights: 5.641267264673745
-----
Iteration Number: 8691
Loss: 28.251412287009384
l2 norm of gradients: 0.17808313187319674
l2 norm of weights: 5.641202658390825
-----
Iteration Number: 8692
Loss: 28.24982203282689
l2 norm of gradients: 0.1780693364589364
l2 norm of weights: 5.641138056784649
-----
Iteration Number: 8693
Loss: 28.24823202462719
l2 norm of gradients: 0.17805554344796307
l2 norm of weights: 5.641073459854595
-----
Iteration Number: 8694
Loss: 28.246642262349233
l2 norm of gradients: 0.17804175283947926
l2 norm of weights: 5.641008867600044
-----
Iteration Number: 8695
Loss: 28.245052745928284
l2 norm of gradients: 0.17802796463268794
l2 norm of weights: 5.640944280020374
-----
Iteration Number: 8696
Loss: 28.24346347530583
l2 norm of gradients: 0.17801417882679252
l2 norm of weights: 5.6408796971149675
-----
Iteration Number: 8697
Loss: 28.241874450415285
l2 norm of gradients: 0.1780003954209969
l2 norm of weights: 5.640815118883204
-----
Iteration Number: 8698
Loss: 28.240285671198496
l2 norm of gradients: 0.1779866144145054
l2 norm of weights: 5.640750545324463
-----
Iteration Number: 8699
Loss: 28.238697137590954

```

```

12 norm of gradients: 0.1779728358065228
12 norm of weights: 5.640685976438125
-----
Iteration Number: 8700
Loss: 28.23710884953177
12 norm of gradients: 0.1779590595962542
12 norm of weights: 5.640621412223572
-----
Iteration Number: 8701
Loss: 28.235520806956348
12 norm of gradients: 0.17794528578290542
12 norm of weights: 5.6405568526801835
-----
Iteration Number: 8702
Loss: 28.233933009805344
12 norm of gradients: 0.17793151436568252
12 norm of weights: 5.640492297807341
-----
Iteration Number: 8703
Loss: 28.232345458016354
12 norm of gradients: 0.1779177453437921
12 norm of weights: 5.6404277476044244
-----
Iteration Number: 8704
Loss: 28.230758151525766
12 norm of gradients: 0.1779039787164411
12 norm of weights: 5.640363202070818
-----
Iteration Number: 8705
Loss: 28.22917109027283
12 norm of gradients: 0.17789021448283698
12 norm of weights: 5.640298661205899
-----
Iteration Number: 8706
Loss: 28.22758427419459
12 norm of gradients: 0.1778764526421877
12 norm of weights: 5.640234125009052
-----
Iteration Number: 8707
Loss: 28.22599770322938
12 norm of gradients: 0.1778626931937016
12 norm of weights: 5.640169593479658
-----
Iteration Number: 8708
Loss: 28.22441137731468
12 norm of gradients: 0.17784893613658742
12 norm of weights: 5.640105066617097
-----

```

Iteration Number: 8709
Loss: 28.222825296391058
l2 norm of gradients: 0.17783518147005445
l2 norm of weights: 5.640040544420753

Iteration Number: 8710
Loss: 28.22123946039363
l2 norm of gradients: 0.17782142919331237
l2 norm of weights: 5.6399760268900065

Iteration Number: 8711
Loss: 28.2196538692621
l2 norm of gradients: 0.1778076793055713
l2 norm of weights: 5.639911514024241

Iteration Number: 8712
Loss: 28.218068522932658
l2 norm of gradients: 0.17779393180604178
l2 norm of weights: 5.639847005822837

Iteration Number: 8713
Loss: 28.216483421346375
l2 norm of gradients: 0.1777801866939349
l2 norm of weights: 5.639782502285179

Iteration Number: 8714
Loss: 28.21489856444077
l2 norm of gradients: 0.17776644396846208
l2 norm of weights: 5.639718003410649

Iteration Number: 8715
Loss: 28.213313952152046
l2 norm of gradients: 0.17775270362883522
l2 norm of weights: 5.639653509198628

Iteration Number: 8716
Loss: 28.211729584420958
l2 norm of gradients: 0.17773896567426667
l2 norm of weights: 5.639589019648501

Iteration Number: 8717
Loss: 28.210145461183618
l2 norm of gradients: 0.1777252301039693
l2 norm of weights: 5.63952453475965

Iteration Number: 8718
Loss: 28.208561582379943
l2 norm of gradients: 0.17771149691715626

```
12 norm of weights: 5.639460054531458
-----
Iteration Number: 8719
Loss: 28.206977947945976
12 norm of gradients: 0.17769776611304125
12 norm of weights: 5.63939557896331
-----
Iteration Number: 8720
Loss: 28.205394557822927
12 norm of gradients: 0.17768403769083838
12 norm of weights: 5.639331108054586
-----
Iteration Number: 8721
Loss: 28.2038114119465
12 norm of gradients: 0.17767031164976224
12 norm of weights: 5.639266641804673
-----
Iteration Number: 8722
Loss: 28.202228510257996
12 norm of gradients: 0.17765658798902784
12 norm of weights: 5.639202180212952
-----
Iteration Number: 8723
Loss: 28.20064585269262
12 norm of gradients: 0.17764286670785057
12 norm of weights: 5.639137723278809
-----
Iteration Number: 8724
Loss: 28.199063439191672
12 norm of gradients: 0.1776291478054464
12 norm of weights: 5.639073271001627
-----
Iteration Number: 8725
Loss: 28.197481269689764
12 norm of gradients: 0.17761543128103152
12 norm of weights: 5.63900882338079
-----
Iteration Number: 8726
Loss: 28.19589934412936
12 norm of gradients: 0.17760171713382286
12 norm of weights: 5.6389443804156825
-----
Iteration Number: 8727
Loss: 28.19431766244917
12 norm of gradients: 0.1775880053630375
12 norm of weights: 5.638879942105688
-----
Iteration Number: 8728
```

```

Loss: 28.19273622458385
l2 norm of gradients: 0.17757429596789315
l2 norm of weights: 5.638815508450192
-----
Iteration Number: 8729
Loss: 28.191155030473862
l2 norm of gradients: 0.1775605889476079
l2 norm of weights: 5.638751079448579
-----
Iteration Number: 8730
Loss: 28.18957408005923
l2 norm of gradients: 0.17754688430140023
l2 norm of weights: 5.6386866551002335
-----
Iteration Number: 8731
Loss: 28.18799337327554
l2 norm of gradients: 0.17753318202848914
l2 norm of weights: 5.63862223540454
-----
Iteration Number: 8732
Loss: 28.18641291006495
l2 norm of gradients: 0.17751948212809404
l2 norm of weights: 5.6385578203608855
-----
Iteration Number: 8733
Loss: 28.184832690362835
l2 norm of gradients: 0.17750578459943472
l2 norm of weights: 5.638493409968652
-----
Iteration Number: 8734
Loss: 28.18325271410991
l2 norm of gradients: 0.17749208944173148
l2 norm of weights: 5.638429004227228
-----
Iteration Number: 8735
Loss: 28.181672981242954
l2 norm of gradients: 0.17747839665420503
l2 norm of weights: 5.638364603135997
-----
Iteration Number: 8736
Loss: 28.180093491705012
l2 norm of gradients: 0.17746470623607652
l2 norm of weights: 5.638300206694345
-----
Iteration Number: 8737
Loss: 28.1785142454286
l2 norm of gradients: 0.17745101818656756
l2 norm of weights: 5.638235814901659

```

```

-----
Iteration Number: 8738
Loss: 28.176935242357466
l2 norm of gradients: 0.17743733250490015
l2 norm of weights: 5.638171427757324
-----
Iteration Number: 8739
Loss: 28.17535648242828
l2 norm of gradients: 0.17742364919029674
l2 norm of weights: 5.638107045260724
-----
Iteration Number: 8740
Loss: 28.173777965581028
l2 norm of gradients: 0.1774099682419803
l2 norm of weights: 5.6380426674112485
-----
Iteration Number: 8741
Loss: 28.172199691752816
l2 norm of gradients: 0.17739628965917403
l2 norm of weights: 5.637978294208281
-----
Iteration Number: 8742
Loss: 28.170621660882706
l2 norm of gradients: 0.17738261344110176
l2 norm of weights: 5.637913925651211
-----
Iteration Number: 8743
Loss: 28.169043872908905
l2 norm of gradients: 0.1773689395869877
l2 norm of weights: 5.637849561739422
-----
Iteration Number: 8744
Loss: 28.16746632777499
l2 norm of gradients: 0.17735526809605656
l2 norm of weights: 5.637785202472302
-----
Iteration Number: 8745
Loss: 28.16588902541379
l2 norm of gradients: 0.1773415989675333
l2 norm of weights: 5.637720847849239
-----
Iteration Number: 8746
Loss: 28.164311965767162
l2 norm of gradients: 0.1773279322006434
l2 norm of weights: 5.637656497869617
-----
Iteration Number: 8747
Loss: 28.162735148775663

```

```

12 norm of gradients: 0.17731426779461293
12 norm of weights: 5.637592152532826
-----
Iteration Number: 8748
Loss: 28.161158574375932
12 norm of gradients: 0.17730060574866813
12 norm of weights: 5.637527811838252
-----
Iteration Number: 8749
Loss: 28.15958224250504
12 norm of gradients: 0.1772869460620359
12 norm of weights: 5.6374634757852835
-----
Iteration Number: 8750
Loss: 28.158006153107227
12 norm of gradients: 0.1772732887339434
12 norm of weights: 5.637399144373306
-----
Iteration Number: 8751
Loss: 28.156430306118843
12 norm of gradients: 0.17725963376361834
12 norm of weights: 5.637334817601708
-----
Iteration Number: 8752
Loss: 28.154854701476324
12 norm of gradients: 0.1772459811502888
12 norm of weights: 5.6372704954698785
-----
Iteration Number: 8753
Loss: 28.15327933912381
12 norm of gradients: 0.17723233089318327
12 norm of weights: 5.637206177977204
-----
Iteration Number: 8754
Loss: 28.151704218998727
12 norm of gradients: 0.17721868299153076
12 norm of weights: 5.6371418651230725
-----
Iteration Number: 8755
Loss: 28.15012934103788
12 norm of gradients: 0.17720503744456076
12 norm of weights: 5.637077556906873
-----
Iteration Number: 8756
Loss: 28.148554705180906
12 norm of gradients: 0.1771913942515029
12 norm of weights: 5.637013253327995
-----

```

Iteration Number: 8757
Loss: 28.14698031136904
l2 norm of gradients: 0.1771777534115876
l2 norm of weights: 5.636948954385823

Iteration Number: 8758
Loss: 28.145406159541917
l2 norm of gradients: 0.17716411492404543
l2 norm of weights: 5.63688466007975

Iteration Number: 8759
Loss: 28.143832249637814
l2 norm of gradients: 0.17715047878810755
l2 norm of weights: 5.636820370409162

Iteration Number: 8760
Loss: 28.142258581595623
l2 norm of gradients: 0.1771368450030055
l2 norm of weights: 5.63675608537345

Iteration Number: 8761
Loss: 28.140685155351157
l2 norm of gradients: 0.17712321356797128
l2 norm of weights: 5.636691804972001

Iteration Number: 8762
Loss: 28.13911197085108
l2 norm of gradients: 0.17710958448223724
l2 norm of weights: 5.636627529204205

Iteration Number: 8763
Loss: 28.137539028030417
l2 norm of gradients: 0.17709595774503625
l2 norm of weights: 5.636563258069452

Iteration Number: 8764
Loss: 28.135966326829532
l2 norm of gradients: 0.1770823333556015
l2 norm of weights: 5.63649899156713

Iteration Number: 8765
Loss: 28.13439386718681
l2 norm of gradients: 0.1770687113131668
l2 norm of weights: 5.636434729696629

Iteration Number: 8766
Loss: 28.132821649042867
l2 norm of gradients: 0.17705509161696614


```
l2 norm of weights: 5.6363704724573385
-----
Iteration Number: 8767
Loss: 28.131249672334068
l2 norm of gradients: 0.17704147426623415
l2 norm of weights: 5.63630621984865
-----
Iteration Number: 8768
Loss: 28.129677937004047
l2 norm of gradients: 0.17702785926020573
l2 norm of weights: 5.636241971869953
-----
Iteration Number: 8769
Loss: 28.128106442991946
l2 norm of gradients: 0.17701424659811626
l2 norm of weights: 5.636177728520635
-----
Iteration Number: 8770
Loss: 28.12653519023366
l2 norm of gradients: 0.17700063627920168
l2 norm of weights: 5.636113489800088
-----
Iteration Number: 8771
Loss: 28.124964178672254
l2 norm of gradients: 0.1769870283026981
l2 norm of weights: 5.636049255707705
-----
Iteration Number: 8772
Loss: 28.123393408245796
l2 norm of gradients: 0.17697342266784227
l2 norm of weights: 5.635985026242872
-----
Iteration Number: 8773
Loss: 28.121822878893962
l2 norm of gradients: 0.17695981937387126
l2 norm of weights: 5.635920801404981
-----
Iteration Number: 8774
Loss: 28.120252590557374
l2 norm of gradients: 0.1769462184200226
l2 norm of weights: 5.635856581193425
-----
Iteration Number: 8775
Loss: 28.118682543173392
l2 norm of gradients: 0.1769326198055342
l2 norm of weights: 5.6357923656075934
-----
Iteration Number: 8776
```

Loss: 28.11711273668258
l2 norm of gradients: 0.1769190235296445
l2 norm of weights: 5.635728154646877

Iteration Number: 8777
Loss: 28.11554317102464
l2 norm of gradients: 0.1769054295915922
l2 norm of weights: 5.6356639483106665

Iteration Number: 8778
Loss: 28.113973846141484
l2 norm of gradients: 0.17689183799061664
l2 norm of weights: 5.635599746598355

Iteration Number: 8779
Loss: 28.11240476197151
l2 norm of gradients: 0.17687824872595742
l2 norm of weights: 5.635535549509331

Iteration Number: 8780
Loss: 28.110835918451958
l2 norm of gradients: 0.17686466179685445
l2 norm of weights: 5.635471357042991

Iteration Number: 8781
Loss: 28.109267315525194
l2 norm of gradients: 0.17685107720254842
l2 norm of weights: 5.6354071691987215

Iteration Number: 8782
Loss: 28.107698953130967
l2 norm of gradients: 0.17683749494228018
l2 norm of weights: 5.635342985975917

Iteration Number: 8783
Loss: 28.10613083120727
l2 norm of gradients: 0.176823915015291
l2 norm of weights: 5.635278807373969

Iteration Number: 8784
Loss: 28.104562949695993
l2 norm of gradients: 0.1768103374208227
l2 norm of weights: 5.635214633392271

Iteration Number: 8785
Loss: 28.10299530853681
l2 norm of gradients: 0.1767967621581174
l2 norm of weights: 5.635150464030212

```

-----
Iteration Number: 8786
Loss: 28.10142790766779
l2 norm of gradients: 0.17678318922641778
l2 norm of weights: 5.635086299287187
-----
Iteration Number: 8787
Loss: 28.09986074703251
l2 norm of gradients: 0.1767696186249668
l2 norm of weights: 5.635022139162588
-----
Iteration Number: 8788
Loss: 28.09829382656726
l2 norm of gradients: 0.17675605035300787
l2 norm of weights: 5.634957983655809
-----
Iteration Number: 8789
Loss: 28.096727146211848
l2 norm of gradients: 0.17674248440978488
l2 norm of weights: 5.634893832766241
-----
Iteration Number: 8790
Loss: 28.095160705907347
l2 norm of gradients: 0.1767289207945421
l2 norm of weights: 5.634829686493276
-----
Iteration Number: 8791
Loss: 28.093594505595526
l2 norm of gradients: 0.17671535950652428
l2 norm of weights: 5.634765544836309
-----
Iteration Number: 8792
Loss: 28.09202854521346
l2 norm of gradients: 0.17670180054497645
l2 norm of weights: 5.634701407794733
-----
Iteration Number: 8793
Loss: 28.090462824703557
l2 norm of gradients: 0.17668824390914423
l2 norm of weights: 5.6346372753679415
-----
Iteration Number: 8794
Loss: 28.088897344004668
l2 norm of gradients: 0.17667468959827343
l2 norm of weights: 5.634573147555328
-----
Iteration Number: 8795
Loss: 28.087332103057634

```

```

12 norm of gradients: 0.1766611376116106
12 norm of weights: 5.634509024356284
-----
Iteration Number: 8796
Loss: 28.085767101799057
12 norm of gradients: 0.1766475879484024
12 norm of weights: 5.634444905770205
-----
Iteration Number: 8797
Loss: 28.084202340175054
12 norm of gradients: 0.1766340406078961
12 norm of weights: 5.634380791796485
-----
Iteration Number: 8798
Loss: 28.082637818122382
12 norm of gradients: 0.17662049558933932
12 norm of weights: 5.634316682434518
-----
Iteration Number: 8799
Loss: 28.08107353558152
12 norm of gradients: 0.17660695289198006
12 norm of weights: 5.634252577683698
-----
Iteration Number: 8800
Loss: 28.07950949249347
12 norm of gradients: 0.17659341251506683
12 norm of weights: 5.634188477543419
-----
Iteration Number: 8801
Loss: 28.077945688794923
12 norm of gradients: 0.17657987445784845
12 norm of weights: 5.634124382013076
-----
Iteration Number: 8802
Loss: 28.07638212443155
12 norm of gradients: 0.17656633871957428
12 norm of weights: 5.634060291092063
-----
Iteration Number: 8803
Loss: 28.07481879933947
12 norm of gradients: 0.17655280529949396
12 norm of weights: 5.633996204779774
-----
Iteration Number: 8804
Loss: 28.073255713463652
12 norm of gradients: 0.17653927419685767
12 norm of weights: 5.633932123075606
-----

```

Iteration Number: 8805
Loss: 28.071692866738996
l2 norm of gradients: 0.1765257454109159
l2 norm of weights: 5.633868045978953

Iteration Number: 8806
Loss: 28.07013025910711
l2 norm of gradients: 0.17651221894091962
l2 norm of weights: 5.633803973489209

Iteration Number: 8807
Loss: 28.06856789051067
l2 norm of gradients: 0.17649869478612015
l2 norm of weights: 5.633739905605769

Iteration Number: 8808
Loss: 28.06700576088885
l2 norm of gradients: 0.17648517294576938
l2 norm of weights: 5.6336758423280315

Iteration Number: 8809
Loss: 28.065443870180047
l2 norm of gradients: 0.1764716534191194
l2 norm of weights: 5.633611783655387

Iteration Number: 8810
Loss: 28.063882218329365
l2 norm of gradients: 0.17645813620542286
l2 norm of weights: 5.633547729587236

Iteration Number: 8811
Loss: 28.06232080527362
l2 norm of gradients: 0.1764446213039328
l2 norm of weights: 5.633483680122972

Iteration Number: 8812
Loss: 28.060759630953186
l2 norm of gradients: 0.17643110871390258
l2 norm of weights: 5.633419635261991

Iteration Number: 8813
Loss: 28.059198695311373
l2 norm of gradients: 0.17641759843458615
l2 norm of weights: 5.633355595003689

Iteration Number: 8814
Loss: 28.05763799828394
l2 norm of gradients: 0.17640409046523767

```

12 norm of weights: 5.633291559347462
-----
Iteration Number: 8815
Loss: 28.056077539816553
12 norm of gradients: 0.17639058480511186
12 norm of weights: 5.633227528292707
-----
Iteration Number: 8816
Loss: 28.05451731984871
12 norm of gradients: 0.1763770814534638
12 norm of weights: 5.63316350183882
-----
Iteration Number: 8817
Loss: 28.052957338317764
12 norm of gradients: 0.17636358040954897
12 norm of weights: 5.633099479985198
-----
Iteration Number: 8818
Loss: 28.051397595166694
12 norm of gradients: 0.17635008167262328
12 norm of weights: 5.633035462731237
-----
Iteration Number: 8819
Loss: 28.0498380903361
12 norm of gradients: 0.17633658524194307
12 norm of weights: 5.632971450076334
-----
Iteration Number: 8820
Loss: 28.04827882376667
12 norm of gradients: 0.176323091116765
12 norm of weights: 5.632907442019886
-----
Iteration Number: 8821
Loss: 28.04671979539891
12 norm of gradients: 0.17630959929634626
12 norm of weights: 5.63284343856129
-----
Iteration Number: 8822
Loss: 28.04516100517552
12 norm of gradients: 0.1762961097799444
12 norm of weights: 5.632779439699943
-----
Iteration Number: 8823
Loss: 28.043602453032555
12 norm of gradients: 0.17628262256681732
12 norm of weights: 5.6327154454352435
-----
Iteration Number: 8824

```

Loss: 28.04204413891564
l2 norm of gradients: 0.17626913765622343
l2 norm of weights: 5.632651455766589

Iteration Number: 8825
Loss: 28.040486062760174
l2 norm of gradients: 0.17625565504742144
l2 norm of weights: 5.632587470693375

Iteration Number: 8826
Loss: 28.038928224511555
l2 norm of gradients: 0.17624217473967066
l2 norm of weights: 5.632523490215001

Iteration Number: 8827
Loss: 28.037370624109577
l2 norm of gradients: 0.17622869673223052
l2 norm of weights: 5.632459514330865

Iteration Number: 8828
Loss: 28.03581326149522
l2 norm of gradients: 0.17621522102436113
l2 norm of weights: 5.632395543040364

Iteration Number: 8829
Loss: 28.03425613660746
l2 norm of gradients: 0.1762017476153228
l2 norm of weights: 5.632331576342896

Iteration Number: 8830
Loss: 28.032699249388077
l2 norm of gradients: 0.1761882765043765
l2 norm of weights: 5.632267614237862

Iteration Number: 8831
Loss: 28.031142599779503
l2 norm of gradients: 0.17617480769078323
l2 norm of weights: 5.6322036567246565

Iteration Number: 8832
Loss: 28.029586187721783
l2 norm of gradients: 0.17616134117380478
l2 norm of weights: 5.6321397038026815

Iteration Number: 8833
Loss: 28.028030013154186
l2 norm of gradients: 0.17614787695270312
l2 norm of weights: 5.632075755471333

```

-----
Iteration Number: 8834
Loss: 28.0264740760207
l2 norm of gradients: 0.17613441502674068
l2 norm of weights: 5.6320118117300115
-----
Iteration Number: 8835
Loss: 28.02491837625868
l2 norm of gradients: 0.1761209553951803
l2 norm of weights: 5.631947872578116
-----
Iteration Number: 8836
Loss: 28.023362913814577
l2 norm of gradients: 0.1761074980572852
l2 norm of weights: 5.6318839380150445
-----
Iteration Number: 8837
Loss: 28.021807688622765
l2 norm of gradients: 0.17609404301231915
l2 norm of weights: 5.631820008040196
-----
Iteration Number: 8838
Loss: 28.020252700629392
l2 norm of gradients: 0.17608059025954606
l2 norm of weights: 5.631756082652972
-----
Iteration Number: 8839
Loss: 28.018697949773745
l2 norm of gradients: 0.1760671397982305
l2 norm of weights: 5.6316921618527696
-----
Iteration Number: 8840
Loss: 28.017143435995663
l2 norm of gradients: 0.17605369162763726
l2 norm of weights: 5.631628245638989
-----
Iteration Number: 8841
Loss: 28.015589159240314
l2 norm of gradients: 0.17604024574703164
l2 norm of weights: 5.631564334011032
-----
Iteration Number: 8842
Loss: 28.014035119443708
l2 norm of gradients: 0.1760268021556793
l2 norm of weights: 5.631500426968295
-----
Iteration Number: 8843
Loss: 28.01248131654988

```



```

12 norm of gradients: 0.17601336085284633
12 norm of weights: 5.631436524510182
-----
Iteration Number: 8844
Loss: 28.010927750499288
12 norm of gradients: 0.17599992183779922
12 norm of weights: 5.631372626636089
-----
Iteration Number: 8845
Loss: 28.00937442123504
12 norm of gradients: 0.1759864851098048
12 norm of weights: 5.631308733345419
-----
Iteration Number: 8846
Loss: 28.007821328694117
12 norm of gradients: 0.17597305066813035
12 norm of weights: 5.631244844637572
-----
Iteration Number: 8847
Loss: 28.00626847282425
12 norm of gradients: 0.1759596185120436
12 norm of weights: 5.631180960511948
-----
Iteration Number: 8848
Loss: 28.004715853559492
12 norm of gradients: 0.17594618864081266
12 norm of weights: 5.631117080967948
-----
Iteration Number: 8849
Loss: 28.003163470845262
12 norm of gradients: 0.17593276105370595
12 norm of weights: 5.631053206004972
-----
Iteration Number: 8850
Loss: 28.001611324623894
12 norm of gradients: 0.17591933574999236
12 norm of weights: 5.630989335622422
-----
Iteration Number: 8851
Loss: 28.000059414834627
12 norm of gradients: 0.17590591272894118
12 norm of weights: 5.630925469819699
-----
Iteration Number: 8852
Loss: 27.99850774141747
12 norm of gradients: 0.17589249198982215
12 norm of weights: 5.630861608596203
-----

```

Iteration Number: 8853
Loss: 27.996956304316825
l2 norm of gradients: 0.1758790735319053
l2 norm of weights: 5.630797751951338

Iteration Number: 8854
Loss: 27.995405103472805
l2 norm of gradients: 0.17586565735446114
l2 norm of weights: 5.630733899884502

Iteration Number: 8855
Loss: 27.993854138827345
l2 norm of gradients: 0.1758522434567605
l2 norm of weights: 5.630670052395099

Iteration Number: 8856
Loss: 27.99230341032211
l2 norm of gradients: 0.17583883183807475
l2 norm of weights: 5.63060620948253

Iteration Number: 8857
Loss: 27.990752917896845
l2 norm of gradients: 0.1758254224976755
l2 norm of weights: 5.630542371146197

Iteration Number: 8858
Loss: 27.989202661494136
l2 norm of gradients: 0.17581201543483493
l2 norm of weights: 5.6304785373855015

Iteration Number: 8859
Loss: 27.987652641057043
l2 norm of gradients: 0.1757986106488254
l2 norm of weights: 5.630414708199846

Iteration Number: 8860
Loss: 27.98610285652439
l2 norm of gradients: 0.17578520813891985
l2 norm of weights: 5.630350883588633

Iteration Number: 8861
Loss: 27.984553307839278
l2 norm of gradients: 0.1757718079043916
l2 norm of weights: 5.630287063551265

Iteration Number: 8862
Loss: 27.98300399494427
l2 norm of gradients: 0.1757584099445142

```

12 norm of weights: 5.630223248087144
-----
Iteration Number: 8863
Loss: 27.981454917780724
12 norm of gradients: 0.1757450142585618
12 norm of weights: 5.630159437195672
-----
Iteration Number: 8864
Loss: 27.979906076286422
12 norm of gradients: 0.17573162084580882
12 norm of weights: 5.630095630876253
-----
Iteration Number: 8865
Loss: 27.97835747040685
12 norm of gradients: 0.1757182297055302
12 norm of weights: 5.63003182912829
-----
Iteration Number: 8866
Loss: 27.976809100083397
12 norm of gradients: 0.17570484083700114
12 norm of weights: 5.629968031951185
-----
Iteration Number: 8867
Loss: 27.97526096525834
12 norm of gradients: 0.17569145423949725
12 norm of weights: 5.629904239344342
-----
Iteration Number: 8868
Loss: 27.9737130658701
12 norm of gradients: 0.17567806991229462
12 norm of weights: 5.629840451307164
-----
Iteration Number: 8869
Loss: 27.97216540186307
12 norm of gradients: 0.17566468785466968
12 norm of weights: 5.629776667839054
-----
Iteration Number: 8870
Loss: 27.970617973179593
12 norm of gradients: 0.1756513080658993
12 norm of weights: 5.629712888939416
-----
Iteration Number: 8871
Loss: 27.969070779759836
12 norm of gradients: 0.1756379305452607
12 norm of weights: 5.629649114607655
-----
Iteration Number: 8872

```

Loss: 27.96752382154565
l2 norm of gradients: 0.17562455529203144
l2 norm of weights: 5.629585344843171

Iteration Number: 8873
Loss: 27.965977098479698
l2 norm of gradients: 0.1756111823054896
l2 norm of weights: 5.629521579645373

Iteration Number: 8874
Loss: 27.96443061050675
l2 norm of gradients: 0.17559781158491358
l2 norm of weights: 5.629457819013661

Iteration Number: 8875
Loss: 27.962884357562686
l2 norm of gradients: 0.17558444312958213
l2 norm of weights: 5.6293940629474415

Iteration Number: 8876
Loss: 27.961338339591997
l2 norm of gradients: 0.17557107693877452
l2 norm of weights: 5.629330311446117

Iteration Number: 8877
Loss: 27.959792556537288
l2 norm of gradients: 0.17555771301177034
l2 norm of weights: 5.629266564509095

Iteration Number: 8878
Loss: 27.958247008338866
l2 norm of gradients: 0.1755443513478495
l2 norm of weights: 5.6292028221357775

Iteration Number: 8879
Loss: 27.95670169494225
l2 norm of gradients: 0.17553099194629235
l2 norm of weights: 5.62913908432557

Iteration Number: 8880
Loss: 27.95515661628565
l2 norm of gradients: 0.17551763480637975
l2 norm of weights: 5.629075351077876

Iteration Number: 8881
Loss: 27.95361177231406
l2 norm of gradients: 0.1755042799273928
l2 norm of weights: 5.629011622392103

```

-----
Iteration Number: 8882
Loss: 27.95206716296718
l2 norm of gradients: 0.17549092730861307
l2 norm of weights: 5.628947898267656
-----
Iteration Number: 8883
Loss: 27.950522788188074
l2 norm of gradients: 0.17547757694932245
l2 norm of weights: 5.628884178703937
-----
Iteration Number: 8884
Loss: 27.948978647919215
l2 norm of gradients: 0.1754642288488033
l2 norm of weights: 5.628820463700356
-----
Iteration Number: 8885
Loss: 27.947434742101827
l2 norm of gradients: 0.17545088300633832
l2 norm of weights: 5.628756753256314
-----
Iteration Number: 8886
Loss: 27.94589107067828
l2 norm of gradients: 0.17543753942121063
l2 norm of weights: 5.62869304737122
-----
Iteration Number: 8887
Loss: 27.944347633590812
l2 norm of gradients: 0.17542419809270365
l2 norm of weights: 5.6286293460444785
-----
Iteration Number: 8888
Loss: 27.942804430780757
l2 norm of gradients: 0.1754108590201013
l2 norm of weights: 5.628565649275495
-----
Iteration Number: 8889
Loss: 27.941261462193694
l2 norm of gradients: 0.17539752220268792
l2 norm of weights: 5.628501957063677
-----
Iteration Number: 8890
Loss: 27.93971872776809
l2 norm of gradients: 0.17538418763974808
l2 norm of weights: 5.628438269408429
-----
Iteration Number: 8891
Loss: 27.93817622744541

```

```

12 norm of gradients: 0.1753708553305669
12 norm of weights: 5.6283745863091585
-----
Iteration Number: 8892
Loss: 27.936633961171815
12 norm of gradients: 0.17535752527442972
12 norm of weights: 5.628310907765271
-----
Iteration Number: 8893
Loss: 27.935091928888657
12 norm of gradients: 0.1753441974706224
12 norm of weights: 5.6282472337761735
-----
Iteration Number: 8894
Loss: 27.933550130534346
12 norm of gradients: 0.17533087191843114
12 norm of weights: 5.628183564341273
-----
Iteration Number: 8895
Loss: 27.932008566056272
12 norm of gradients: 0.17531754861714255
12 norm of weights: 5.628119899459977
-----
Iteration Number: 8896
Loss: 27.930467235395458
12 norm of gradients: 0.1753042275660436
12 norm of weights: 5.62805623913169
-----
Iteration Number: 8897
Loss: 27.928926138490564
12 norm of gradients: 0.17529090876442163
12 norm of weights: 5.627992583355822
-----
Iteration Number: 8898
Loss: 27.927385275288714
12 norm of gradients: 0.1752775922115645
12 norm of weights: 5.627928932131777
-----
Iteration Number: 8899
Loss: 27.925844645731242
12 norm of gradients: 0.17526427790676016
12 norm of weights: 5.627865285458966
-----
Iteration Number: 8900
Loss: 27.92430424976087
12 norm of gradients: 0.17525096584929725
12 norm of weights: 5.6278016433367934
-----

```

Iteration Number: 8901
Loss: 27.922764087316587
l2 norm of gradients: 0.1752376560384647
l2 norm of weights: 5.627738005764669

Iteration Number: 8902
Loss: 27.92122415834342
l2 norm of gradients: 0.17522434847355167
l2 norm of weights: 5.627674372741999

Iteration Number: 8903
Loss: 27.919684462784893
l2 norm of gradients: 0.17521104315384797
l2 norm of weights: 5.627610744268193

Iteration Number: 8904
Loss: 27.918145000580758
l2 norm of gradients: 0.17519774007864355
l2 norm of weights: 5.627547120342657

Iteration Number: 8905
Loss: 27.91660577167697
l2 norm of gradients: 0.17518443924722893
l2 norm of weights: 5.627483500964798

Iteration Number: 8906
Loss: 27.91506677601367
l2 norm of gradients: 0.17517114065889494
l2 norm of weights: 5.627419886134029

Iteration Number: 8907
Loss: 27.913528013533355
l2 norm of gradients: 0.1751578443129327
l2 norm of weights: 5.627356275849754

Iteration Number: 8908
Loss: 27.911989484179962
l2 norm of gradients: 0.17514455020863381
l2 norm of weights: 5.627292670111385

Iteration Number: 8909
Loss: 27.910451187895905
l2 norm of gradients: 0.17513125834529034
l2 norm of weights: 5.627229068918327

Iteration Number: 8910
Loss: 27.908913124622828
l2 norm of gradients: 0.1751179687221945

```

12 norm of weights: 5.627165472269991
-----
Iteration Number: 8911
Loss: 27.90737529430261
12 norm of gradients: 0.17510468133863913
12 norm of weights: 5.627101880165785
-----
Iteration Number: 8912
Loss: 27.9058376968835
12 norm of gradients: 0.1750913961939173
12 norm of weights: 5.627038292605119
-----
Iteration Number: 8913
Loss: 27.904300332299023
12 norm of gradients: 0.17507811328732253
12 norm of weights: 5.626974709587401
-----
Iteration Number: 8914
Loss: 27.902763200500715
12 norm of gradients: 0.17506483261814867
12 norm of weights: 5.626911131112041
-----
Iteration Number: 8915
Loss: 27.901226301425844
12 norm of gradients: 0.17505155418569
12 norm of weights: 5.626847557178449
-----
Iteration Number: 8916
Loss: 27.899689635019698
12 norm of gradients: 0.1750382779892411
12 norm of weights: 5.626783987786034
-----
Iteration Number: 8917
Loss: 27.898153201224737
12 norm of gradients: 0.17502500402809706
12 norm of weights: 5.626720422934205
-----
Iteration Number: 8918
Loss: 27.896616999981426
12 norm of gradients: 0.17501173230155323
12 norm of weights: 5.626656862622372
-----
Iteration Number: 8919
Loss: 27.89508103123722
12 norm of gradients: 0.17499846280890538
12 norm of weights: 5.626593306849946
-----
Iteration Number: 8920

```


Loss: 27.893545294929886
l2 norm of gradients: 0.17498519554944966
l2 norm of weights: 5.626529755616335

Iteration Number: 8921
Loss: 27.892009791005748
l2 norm of gradients: 0.1749719305224826
l2 norm of weights: 5.626466208920951

Iteration Number: 8922
Loss: 27.89047451940534
l2 norm of gradients: 0.17495866772730115
l2 norm of weights: 5.626402666763205

Iteration Number: 8923
Loss: 27.888939480074992
l2 norm of gradients: 0.17494540716320253
l2 norm of weights: 5.626339129142505

Iteration Number: 8924
Loss: 27.88740467295466
l2 norm of gradients: 0.1749321488294845
l2 norm of weights: 5.626275596058264

Iteration Number: 8925
Loss: 27.885870097989006
l2 norm of gradients: 0.17491889272544492
l2 norm of weights: 5.62621206750989

Iteration Number: 8926
Loss: 27.88433575512013
l2 norm of gradients: 0.17490563885038235
l2 norm of weights: 5.626148543496797

Iteration Number: 8927
Loss: 27.882801644289493
l2 norm of gradients: 0.1748923872035956
l2 norm of weights: 5.626085024018393

Iteration Number: 8928
Loss: 27.881267765443667
l2 norm of gradients: 0.17487913778438374
l2 norm of weights: 5.626021509074091

Iteration Number: 8929
Loss: 27.879734118522457
l2 norm of gradients: 0.17486589059204635
l2 norm of weights: 5.625957998663302

```

-----
Iteration Number: 8930
Loss: 27.878200703471293
l2 norm of gradients: 0.17485264562588343
l2 norm of weights: 5.625894492785436
-----
Iteration Number: 8931
Loss: 27.8766675202312
l2 norm of gradients: 0.17483940288519512
l2 norm of weights: 5.625830991439907
-----
Iteration Number: 8932
Loss: 27.875134568748805
l2 norm of gradients: 0.17482616236928222
l2 norm of weights: 5.625767494626124
-----
Iteration Number: 8933
Loss: 27.873601848962178
l2 norm of gradients: 0.1748129240774457
l2 norm of weights: 5.625704002343499
-----
Iteration Number: 8934
Loss: 27.87206936081974
l2 norm of gradients: 0.17479968800898704
l2 norm of weights: 5.625640514591446
-----
Iteration Number: 8935
Loss: 27.8705371042603
l2 norm of gradients: 0.17478645416320798
l2 norm of weights: 5.6255770313693745
-----
Iteration Number: 8936
Loss: 27.86900507922909
l2 norm of gradients: 0.1747732225394107
l2 norm of weights: 5.625513552676699
-----
Iteration Number: 8937
Loss: 27.867473285669508
l2 norm of gradients: 0.17475999313689772
l2 norm of weights: 5.625450078512829
-----
Iteration Number: 8938
Loss: 27.865941723525083
l2 norm of gradients: 0.174746765954972
l2 norm of weights: 5.62538660887718
-----
Iteration Number: 8939
Loss: 27.8644103927371

```

```

12 norm of gradients: 0.17473354099293675
12 norm of weights: 5.625323143769162
-----
Iteration Number: 8940
Loss: 27.8628792932504
12 norm of gradients: 0.17472031825009568
12 norm of weights: 5.6252596831881885
-----
Iteration Number: 8941
Loss: 27.86134842500973
12 norm of gradients: 0.17470709772575282
12 norm of weights: 5.625196227133672
-----
Iteration Number: 8942
Loss: 27.859817787955457
12 norm of gradients: 0.17469387941921252
12 norm of weights: 5.625132775605025
-----
Iteration Number: 8943
Loss: 27.858287382032895
12 norm of gradients: 0.17468066332977963
12 norm of weights: 5.625069328601663
-----
Iteration Number: 8944
Loss: 27.856757207182795
12 norm of gradients: 0.17466744945675922
12 norm of weights: 5.625005886122996
-----
Iteration Number: 8945
Loss: 27.855227263351967
12 norm of gradients: 0.17465423779945682
12 norm of weights: 5.624942448168439
-----
Iteration Number: 8946
Loss: 27.853697550482984
12 norm of gradients: 0.17464102835717832
12 norm of weights: 5.624879014737405
-----
Iteration Number: 8947
Loss: 27.852168068517777
12 norm of gradients: 0.17462782112923
12 norm of weights: 5.624815585829307
-----
Iteration Number: 8948
Loss: 27.850638817400547
12 norm of gradients: 0.17461461611491844
12 norm of weights: 5.624752161443559
-----

```

Iteration Number: 8949
Loss: 27.849109797075833
l2 norm of gradients: 0.17460141331355064
l2 norm of weights: 5.624688741579575

Iteration Number: 8950
Loss: 27.847581007485974
l2 norm of gradients: 0.17458821272443395
l2 norm of weights: 5.624625326236769

Iteration Number: 8951
Loss: 27.84605244857454
l2 norm of gradients: 0.17457501434687614
l2 norm of weights: 5.624561915414555

Iteration Number: 8952
Loss: 27.84452412028619
l2 norm of gradients: 0.17456181818018526
l2 norm of weights: 5.624498509112347

Iteration Number: 8953
Loss: 27.842996022562104
l2 norm of gradients: 0.17454862422366985
l2 norm of weights: 5.6244351073295595

Iteration Number: 8954
Loss: 27.841468155348
l2 norm of gradients: 0.17453543247663864
l2 norm of weights: 5.624371710065605

Iteration Number: 8955
Loss: 27.839940518587174
l2 norm of gradients: 0.1745222429384009
l2 norm of weights: 5.624308317319901

Iteration Number: 8956
Loss: 27.838413112221957
l2 norm of gradients: 0.17450905560826618
l2 norm of weights: 5.624244929091861

Iteration Number: 8957
Loss: 27.836885936198872
l2 norm of gradients: 0.1744958704855444
l2 norm of weights: 5.6241815453809005

Iteration Number: 8958
Loss: 27.83535899045996
l2 norm of gradients: 0.17448268756954588

```

12 norm of weights: 5.624118166186433
-----
Iteration Number: 8959
Loss: 27.83383227494656
12 norm of gradients: 0.1744695068595813
12 norm of weights: 5.624054791507873
-----
Iteration Number: 8960
Loss: 27.832305789605222
12 norm of gradients: 0.1744563283549617
12 norm of weights: 5.623991421344638
-----
Iteration Number: 8961
Loss: 27.830779534378287
12 norm of gradients: 0.17444315205499844
12 norm of weights: 5.623928055696141
-----
Iteration Number: 8962
Loss: 27.829253509210723
12 norm of gradients: 0.17442997795900334
12 norm of weights: 5.6238646945618
-----
Iteration Number: 8963
Loss: 27.827727714047292
12 norm of gradients: 0.17441680606628845
12 norm of weights: 5.623801337941029
-----
Iteration Number: 8964
Loss: 27.826202148829125
12 norm of gradients: 0.17440363637616632
12 norm of weights: 5.623737985833243
-----
Iteration Number: 8965
Loss: 27.824676813500147
12 norm of gradients: 0.17439046888794982
12 norm of weights: 5.623674638237859
-----
Iteration Number: 8966
Loss: 27.823151708006677
12 norm of gradients: 0.1743773036009521
12 norm of weights: 5.6236112951542925
-----
Iteration Number: 8967
Loss: 27.82162683228898
12 norm of gradients: 0.17436414051448682
12 norm of weights: 5.62354795658196
-----
Iteration Number: 8968

```

```

Loss: 27.820102186296772
l2 norm of gradients: 0.17435097962786794
l2 norm of weights: 5.623484622520277
-----
Iteration Number: 8969
Loss: 27.81857776996798
l2 norm of gradients: 0.1743378209404097
l2 norm of weights: 5.6234212929686604
-----
Iteration Number: 8970
Loss: 27.817053583249336
l2 norm of gradients: 0.17432466445142683
l2 norm of weights: 5.623357967926526
-----
Iteration Number: 8971
Loss: 27.815529626081602
l2 norm of gradients: 0.17431151016023436
l2 norm of weights: 5.623294647393291
-----
Iteration Number: 8972
Loss: 27.81400589841546
l2 norm of gradients: 0.17429835806614769
l2 norm of weights: 5.623231331368372
-----
Iteration Number: 8973
Loss: 27.812482400189335
l2 norm of gradients: 0.1742852081684825
l2 norm of weights: 5.623168019851185
-----
Iteration Number: 8974
Loss: 27.81095913134819
l2 norm of gradients: 0.17427206046655508
l2 norm of weights: 5.623104712841149
-----
Iteration Number: 8975
Loss: 27.809436091836695
l2 norm of gradients: 0.17425891495968177
l2 norm of weights: 5.623041410337678
-----
Iteration Number: 8976
Loss: 27.807913281600005
l2 norm of gradients: 0.17424577164717953
l2 norm of weights: 5.622978112340193
-----
Iteration Number: 8977
Loss: 27.80639070057929
l2 norm of gradients: 0.17423263052836546
l2 norm of weights: 5.622914818848109

```

```

-----
Iteration Number: 8978
Loss: 27.804868348721552
l2 norm of gradients: 0.17421949160255717
l2 norm of weights: 5.622851529860843
-----
Iteration Number: 8979
Loss: 27.803346225969864
l2 norm of gradients: 0.17420635486907263
l2 norm of weights: 5.622788245377813
-----
Iteration Number: 8980
Loss: 27.801824332266122
l2 norm of gradients: 0.17419322032723003
l2 norm of weights: 5.6227249653984375
-----
Iteration Number: 8981
Loss: 27.800302667558004
l2 norm of gradients: 0.17418008797634812
l2 norm of weights: 5.6226616899221336
-----
Iteration Number: 8982
Loss: 27.798781231787594
l2 norm of gradients: 0.17416695781574584
l2 norm of weights: 5.62259841894832
-----
Iteration Number: 8983
Loss: 27.797260024902307
l2 norm of gradients: 0.17415382984474256
l2 norm of weights: 5.622535152476414
-----
Iteration Number: 8984
Loss: 27.795739046841053
l2 norm of gradients: 0.17414070406265802
l2 norm of weights: 5.622471890505834
-----
Iteration Number: 8985
Loss: 27.79421829755185
l2 norm of gradients: 0.1741275804688123
l2 norm of weights: 5.622408633035999
-----
Iteration Number: 8986
Loss: 27.792697776975803
l2 norm of gradients: 0.17411445906252582
l2 norm of weights: 5.622345380066327
-----
Iteration Number: 8987
Loss: 27.791177485062267

```

```

12 norm of gradients: 0.17410133984311932
12 norm of weights: 5.622282131596236
-----
Iteration Number: 8988
Loss: 27.789657421751127
12 norm of gradients: 0.17408822280991412
12 norm of weights: 5.622218887625146
-----
Iteration Number: 8989
Loss: 27.788137586988015
12 norm of gradients: 0.17407510796223155
12 norm of weights: 5.622155648152475
-----
Iteration Number: 8990
Loss: 27.78661798071772
12 norm of gradients: 0.17406199529939356
12 norm of weights: 5.622092413177643
-----
Iteration Number: 8991
Loss: 27.78509860288324
12 norm of gradients: 0.17404888482072234
12 norm of weights: 5.622029182700066
-----
Iteration Number: 8992
Loss: 27.783579453430946
12 norm of gradients: 0.17403577652554053
12 norm of weights: 5.621965956719167
-----
Iteration Number: 8993
Loss: 27.782060532303458
12 norm of gradients: 0.174022670413171
12 norm of weights: 5.6219027352343645
-----
Iteration Number: 8994
Loss: 27.780541839445693
12 norm of gradients: 0.17400956648293706
12 norm of weights: 5.621839518245076
-----
Iteration Number: 8995
Loss: 27.77902337480399
12 norm of gradients: 0.17399646473416233
12 norm of weights: 5.621776305750722
-----
Iteration Number: 8996
Loss: 27.777505138320468
12 norm of gradients: 0.17398336516617086
12 norm of weights: 5.621713097750724
-----

```


Iteration Number: 8997
Loss: 27.77598712994028
l2 norm of gradients: 0.17397026777828692
l2 norm of weights: 5.6216498942444995

Iteration Number: 8998
Loss: 27.774469349607806
l2 norm of gradients: 0.1739571725698353
l2 norm of weights: 5.6215866952314695

Iteration Number: 8999
Loss: 27.77295179726775
l2 norm of gradients: 0.17394407954014104
l2 norm of weights: 5.621523500711055

Iteration Number: 9000
Loss: 27.771434472864
l2 norm of gradients: 0.1739309886885295
l2 norm of weights: 5.621460310682674

Iteration Number: 9001
Loss: 27.769917376341898
l2 norm of gradients: 0.17391790001432644
l2 norm of weights: 5.621397125145749

Iteration Number: 9002
Loss: 27.768400507648558
l2 norm of gradients: 0.1739048135168581
l2 norm of weights: 5.621333944099699

Iteration Number: 9003
Loss: 27.76688386672265
l2 norm of gradients: 0.17389172919545082
l2 norm of weights: 5.621270767543946

Iteration Number: 9004
Loss: 27.76536745351431
l2 norm of gradients: 0.1738786470494315
l2 norm of weights: 5.621207595477909

Iteration Number: 9005
Loss: 27.763851267965673
l2 norm of gradients: 0.1738655670781273
l2 norm of weights: 5.621144427901011

Iteration Number: 9006
Loss: 27.762335310020223
l2 norm of gradients: 0.17385248928086572

```

12 norm of weights: 5.621081264812671
-----
Iteration Number: 9007
Loss: 27.760819579624325
12 norm of gradients: 0.17383941365697467
12 norm of weights: 5.621018106212311
-----
Iteration Number: 9008
Loss: 27.759304076721993
12 norm of gradients: 0.17382634020578233
12 norm of weights: 5.6209549520993525
-----
Iteration Number: 9009
Loss: 27.757788801262368
12 norm of gradients: 0.17381326892661733
12 norm of weights: 5.620891802473216
-----
Iteration Number: 9010
Loss: 27.75627375318103
12 norm of gradients: 0.17380019981880854
12 norm of weights: 5.620828657333324
-----
Iteration Number: 9011
Loss: 27.754758932430907
12 norm of gradients: 0.17378713288168532
12 norm of weights: 5.620765516679097
-----
Iteration Number: 9012
Loss: 27.753244338953188
12 norm of gradients: 0.17377406811457727
12 norm of weights: 5.620702380509958
-----
Iteration Number: 9013
Loss: 27.75172997269323
12 norm of gradients: 0.17376100551681425
12 norm of weights: 5.620639248825328
-----
Iteration Number: 9014
Loss: 27.75021583359734
12 norm of gradients: 0.17374794508772676
12 norm of weights: 5.620576121624629
-----
Iteration Number: 9015
Loss: 27.74870192160804
12 norm of gradients: 0.17373488682664545
12 norm of weights: 5.620512998907284
-----
Iteration Number: 9016

```

Loss: 27.747188236669476
l2 norm of gradients: 0.17372183073290126
l2 norm of weights: 5.620449880672713

Iteration Number: 9017
Loss: 27.745674778731736
l2 norm of gradients: 0.1737087768058256
l2 norm of weights: 5.620386766920341

Iteration Number: 9018
Loss: 27.744161547734258
l2 norm of gradients: 0.17369572504475017
l2 norm of weights: 5.62032365764959

Iteration Number: 9019
Loss: 27.74264854362422
l2 norm of gradients: 0.17368267544900712
l2 norm of weights: 5.620260552859881

Iteration Number: 9020
Loss: 27.741135766345764
l2 norm of gradients: 0.17366962801792876
l2 norm of weights: 5.620197452550638

Iteration Number: 9021
Loss: 27.739623215844944
l2 norm of gradients: 0.17365658275084792
l2 norm of weights: 5.620134356721284

Iteration Number: 9022
Loss: 27.73811089206535
l2 norm of gradients: 0.17364353964709764
l2 norm of weights: 5.620071265371241

Iteration Number: 9023
Loss: 27.736598794953654
l2 norm of gradients: 0.17363049870601144
l2 norm of weights: 5.6200081784999325

Iteration Number: 9024
Loss: 27.735086924455025
l2 norm of gradients: 0.17361745992692307
l2 norm of weights: 5.6199450961067825

Iteration Number: 9025
Loss: 27.73357528051233
l2 norm of gradients: 0.17360442330916673
l2 norm of weights: 5.619882018191214

```

-----
Iteration Number: 9026
Loss: 27.732063863073588
l2 norm of gradients: 0.17359138885207687
l2 norm of weights: 5.61981894475265
-----
Iteration Number: 9027
Loss: 27.73055267207959
l2 norm of gradients: 0.17357835655498835
l2 norm of weights: 5.6197558757905135
-----
Iteration Number: 9028
Loss: 27.72904170747874
l2 norm of gradients: 0.1735653264172363
l2 norm of weights: 5.619692811304231
-----
Iteration Number: 9029
Loss: 27.727530969217046
l2 norm of gradients: 0.17355229843815626
l2 norm of weights: 5.619629751293223
-----
Iteration Number: 9030
Loss: 27.726020457236718
l2 norm of gradients: 0.17353927261708416
l2 norm of weights: 5.619566695756915
-----
Iteration Number: 9031
Loss: 27.724510171486056
l2 norm of gradients: 0.1735262489533562
l2 norm of weights: 5.619503644694731
-----
Iteration Number: 9032
Loss: 27.723000111906245
l2 norm of gradients: 0.17351322744630884
l2 norm of weights: 5.619440598106096
-----
Iteration Number: 9033
Loss: 27.72149027844585
l2 norm of gradients: 0.17350020809527905
l2 norm of weights: 5.619377555990433
-----
Iteration Number: 9034
Loss: 27.71998067104815
l2 norm of gradients: 0.17348719089960407
l2 norm of weights: 5.619314518347168
-----
Iteration Number: 9035
Loss: 27.71847128966077

```

```

12 norm of gradients: 0.1734741758586215
12 norm of weights: 5.619251485175724
-----
Iteration Number: 9036
Loss: 27.71696213422723
12 norm of gradients: 0.1734611629716692
12 norm of weights: 5.619188456475526
-----
Iteration Number: 9037
Loss: 27.715453204692665
12 norm of gradients: 0.17344815223808546
12 norm of weights: 5.619125432245999
-----
Iteration Number: 9038
Loss: 27.71394450100266
12 norm of gradients: 0.17343514365720897
12 norm of weights: 5.61906241248657
-----
Iteration Number: 9039
Loss: 27.712436023103123
12 norm of gradients: 0.17342213722837857
12 norm of weights: 5.618999397196661
-----
Iteration Number: 9040
Loss: 27.710927770937946
12 norm of gradients: 0.1734091329509337
12 norm of weights: 5.618936386375699
-----
Iteration Number: 9041
Loss: 27.70941974445626
12 norm of gradients: 0.1733961308242137
12 norm of weights: 5.618873380023109
-----
Iteration Number: 9042
Loss: 27.707911943598457
12 norm of gradients: 0.17338313084755885
12 norm of weights: 5.618810378138315
-----
Iteration Number: 9043
Loss: 27.706404368313045
12 norm of gradients: 0.17337013302030935
12 norm of weights: 5.6187473807207455
-----
Iteration Number: 9044
Loss: 27.70489701854542
12 norm of gradients: 0.1733571373418058
12 norm of weights: 5.618684387769824
-----

```

Iteration Number: 9045
Loss: 27.7033898942392
l2 norm of gradients: 0.17334414381138927
l2 norm of weights: 5.618621399284978

Iteration Number: 9046
Loss: 27.701882995340142
l2 norm of gradients: 0.17333115242840103
l2 norm of weights: 5.618558415265631

Iteration Number: 9047
Loss: 27.70037632179621
l2 norm of gradients: 0.1733181631921828
l2 norm of weights: 5.618495435711212

Iteration Number: 9048
Loss: 27.698869873550322
l2 norm of gradients: 0.17330517610207652
l2 norm of weights: 5.618432460621145

Iteration Number: 9049
Loss: 27.69736365054835
l2 norm of gradients: 0.1732921911574246
l2 norm of weights: 5.6183694899948575

Iteration Number: 9050
Loss: 27.69585765273797
l2 norm of gradients: 0.1732792083575697
l2 norm of weights: 5.618306523831775

Iteration Number: 9051
Loss: 27.69435188006301
l2 norm of gradients: 0.17326622770185482
l2 norm of weights: 5.618243562131325

Iteration Number: 9052
Loss: 27.69284633246854
l2 norm of gradients: 0.17325324918962337
l2 norm of weights: 5.618180604892935

Iteration Number: 9053
Loss: 27.6913410099015
l2 norm of gradients: 0.17324027282021895
l2 norm of weights: 5.6181176521160285

Iteration Number: 9054
Loss: 27.689835912307643
l2 norm of gradients: 0.1732272985929857

```

12 norm of weights: 5.618054703800036
-----
Iteration Number: 9055
Loss: 27.688331039630615
12 norm of gradients: 0.17321432650726792
12 norm of weights: 5.617991759944384
-----
Iteration Number: 9056
Loss: 27.686826391817952
12 norm of gradients: 0.1732013565624103
12 norm of weights: 5.617928820548498
-----
Iteration Number: 9057
Loss: 27.685321968816364
12 norm of gradients: 0.17318838875775797
12 norm of weights: 5.617865885611806
-----
Iteration Number: 9058
Loss: 27.68381777056781
12 norm of gradients: 0.17317542309265618
12 norm of weights: 5.617802955133736
-----
Iteration Number: 9059
Loss: 27.68231379702203
12 norm of gradients: 0.1731624595664507
12 norm of weights: 5.6177400291137145
-----
Iteration Number: 9060
Loss: 27.68081004812071
12 norm of gradients: 0.17314949817848763
12 norm of weights: 5.617677107551171
-----
Iteration Number: 9061
Loss: 27.679306523812315
12 norm of gradients: 0.17313653892811323
12 norm of weights: 5.617614190445532
-----
Iteration Number: 9062
Loss: 27.67780322404407
12 norm of gradients: 0.17312358181467427
12 norm of weights: 5.617551277796226
-----
Iteration Number: 9063
Loss: 27.6763001487593
12 norm of gradients: 0.17311062683751788
12 norm of weights: 5.617488369602681
-----
Iteration Number: 9064

```

Loss: 27.67479729790379
l2 norm of gradients: 0.1730976739959913
l2 norm of weights: 5.617425465864324

Iteration Number: 9065
Loss: 27.673294671425086
l2 norm of gradients: 0.1730847232894423
l2 norm of weights: 5.617362566580584

Iteration Number: 9066
Loss: 27.67179226926641
l2 norm of gradients: 0.17307177471721893
l2 norm of weights: 5.617299671750891

Iteration Number: 9067
Loss: 27.670290091378327
l2 norm of gradients: 0.17305882827866956
l2 norm of weights: 5.617236781374673

Iteration Number: 9068
Loss: 27.66878813770108
l2 norm of gradients: 0.1730458839731429
l2 norm of weights: 5.617173895451356

Iteration Number: 9069
Loss: 27.66728640818384
l2 norm of gradients: 0.17303294179998804
l2 norm of weights: 5.617111013980372

Iteration Number: 9070
Loss: 27.665784902772042
l2 norm of gradients: 0.17302000175855428
l2 norm of weights: 5.617048136961148

Iteration Number: 9071
Loss: 27.66428362141193
l2 norm of gradients: 0.17300706384819134
l2 norm of weights: 5.616985264393113

Iteration Number: 9072
Loss: 27.66278256405093
l2 norm of gradients: 0.17299412806824926
l2 norm of weights: 5.616922396275698

Iteration Number: 9073
Loss: 27.66128173063157
l2 norm of gradients: 0.17298119441807844
l2 norm of weights: 5.6168595326083315


```

-----
Iteration Number: 9074
Loss: 27.65978112109968
l2 norm of gradients: 0.17296826289702955
l2 norm of weights: 5.616796673390442
-----
Iteration Number: 9075
Loss: 27.65828073540608
l2 norm of gradients: 0.1729553335044536
l2 norm of weights: 5.6167338186214595
-----
Iteration Number: 9076
Loss: 27.656780573493148
l2 norm of gradients: 0.17294240623970197
l2 norm of weights: 5.616670968300814
-----
Iteration Number: 9077
Loss: 27.655280635308397
l2 norm of gradients: 0.1729294811021264
l2 norm of weights: 5.616608122427935
-----
Iteration Number: 9078
Loss: 27.65378092079657
l2 norm of gradients: 0.17291655809107878
l2 norm of weights: 5.616545281002252
-----
Iteration Number: 9079
Loss: 27.652281429904335
l2 norm of gradients: 0.17290363720591156
l2 norm of weights: 5.616482444023196
-----
Iteration Number: 9080
Loss: 27.650782162579034
l2 norm of gradients: 0.17289071844597734
l2 norm of weights: 5.616419611490198
-----
Iteration Number: 9081
Loss: 27.64928311876534
l2 norm of gradients: 0.1728778018106292
l2 norm of weights: 5.6163567834026855
-----
Iteration Number: 9082
Loss: 27.647784298411892
l2 norm of gradients: 0.1728648872992204
l2 norm of weights: 5.61629395976009
-----
Iteration Number: 9083
Loss: 27.646285701459522

```

```

12 norm of gradients: 0.17285197491110463
12 norm of weights: 5.616231140561843
-----
Iteration Number: 9084
Loss: 27.644787327859632
12 norm of gradients: 0.17283906464563586
12 norm of weights: 5.616168325807374
-----
Iteration Number: 9085
Loss: 27.643289177558295
12 norm of gradients: 0.17282615650216834
12 norm of weights: 5.616105515496114
-----
Iteration Number: 9086
Loss: 27.64179125049862
12 norm of gradients: 0.17281325048005686
12 norm of weights: 5.616042709627495
-----
Iteration Number: 9087
Loss: 27.640293546629294
12 norm of gradients: 0.17280034657865626
12 norm of weights: 5.615979908200948
-----
Iteration Number: 9088
Loss: 27.638796065895345
12 norm of gradients: 0.17278744479732186
12 norm of weights: 5.615917111215901
-----
Iteration Number: 9089
Loss: 27.63729880824513
12 norm of gradients: 0.1727745451354093
12 norm of weights: 5.615854318671789
-----
Iteration Number: 9090
Loss: 27.635801773621537
12 norm of gradients: 0.17276164759227453
12 norm of weights: 5.615791530568041
-----
Iteration Number: 9091
Loss: 27.634304961974077
12 norm of gradients: 0.17274875216727373
12 norm of weights: 5.61572874690409
-----
Iteration Number: 9092
Loss: 27.632808373246924
12 norm of gradients: 0.1727358588597636
12 norm of weights: 5.615665967679367
-----

```

Iteration Number: 9093
Loss: 27.631312007388384
l2 norm of gradients: 0.172722967669101
l2 norm of weights: 5.615603192893303

Iteration Number: 9094
Loss: 27.629815864343392
l2 norm of gradients: 0.17271007859464316
l2 norm of weights: 5.615540422545331

Iteration Number: 9095
Loss: 27.6283199440611
l2 norm of gradients: 0.17269719163574765
l2 norm of weights: 5.615477656634883

Iteration Number: 9096
Loss: 27.62682424648372
l2 norm of gradients: 0.17268430679177238
l2 norm of weights: 5.61541489516139

Iteration Number: 9097
Loss: 27.625328771560056
l2 norm of gradients: 0.17267142406207556
l2 norm of weights: 5.615352138124285

Iteration Number: 9098
Loss: 27.62383351923875
l2 norm of gradients: 0.1726585434460157
l2 norm of weights: 5.615289385523001

Iteration Number: 9099
Loss: 27.622338489462578
l2 norm of gradients: 0.17264566494295167
l2 norm of weights: 5.615226637356969

Iteration Number: 9100
Loss: 27.620843682179565
l2 norm of gradients: 0.17263278855224265
l2 norm of weights: 5.615163893625622

Iteration Number: 9101
Loss: 27.61934909733573
l2 norm of gradients: 0.1726199142732481
l2 norm of weights: 5.615101154328394

Iteration Number: 9102
Loss: 27.617854734878694
l2 norm of gradients: 0.172607042105328

```

12 norm of weights: 5.615038419464716
-----
Iteration Number: 9103
Loss: 27.61636059475456
12 norm of gradients: 0.17259417204784225
12 norm of weights: 5.614975689034022
-----
Iteration Number: 9104
Loss: 27.614866676910655
12 norm of gradients: 0.17258130410015154
12 norm of weights: 5.614912963035745
-----
Iteration Number: 9105
Loss: 27.613372981292407
12 norm of gradients: 0.17256843826161652
12 norm of weights: 5.614850241469318
-----
Iteration Number: 9106
Loss: 27.611879507846794
12 norm of gradients: 0.17255557453159834
12 norm of weights: 5.614787524334175
-----
Iteration Number: 9107
Loss: 27.610386256519725
12 norm of gradients: 0.17254271290945844
12 norm of weights: 5.614724811629748
-----
Iteration Number: 9108
Loss: 27.608893227261103
12 norm of gradients: 0.17252985339455854
12 norm of weights: 5.614662103355472
-----
Iteration Number: 9109
Loss: 27.60740042001309
12 norm of gradients: 0.17251699598626072
12 norm of weights: 5.614599399510779
-----
Iteration Number: 9110
Loss: 27.605907834725564
12 norm of gradients: 0.17250414068392733
12 norm of weights: 5.614536700095105
-----
Iteration Number: 9111
Loss: 27.604415471344744
12 norm of gradients: 0.17249128748692116
12 norm of weights: 5.6144740051078825
-----
Iteration Number: 9112

```

Loss: 27.6029233298177
l2 norm of gradients: 0.17247843639460514
l2 norm of weights: 5.614411314548546

Iteration Number: 9113
Loss: 27.601431410090157
l2 norm of gradients: 0.17246558740634269
l2 norm of weights: 5.614348628416529

Iteration Number: 9114
Loss: 27.599939712109524
l2 norm of gradients: 0.17245274052149742
l2 norm of weights: 5.6142859467112665

Iteration Number: 9115
Loss: 27.598448235822012
l2 norm of gradients: 0.17243989573943336
l2 norm of weights: 5.614223269432194

Iteration Number: 9116
Loss: 27.59695698117643
l2 norm of gradients: 0.17242705305951478
l2 norm of weights: 5.614160596578744

Iteration Number: 9117
Loss: 27.59546594811652
l2 norm of gradients: 0.17241421248110625
l2 norm of weights: 5.614097928150351

Iteration Number: 9118
Loss: 27.59397513659235
l2 norm of gradients: 0.17240137400357278
l2 norm of weights: 5.614035264146452

Iteration Number: 9119
Loss: 27.592484546548185
l2 norm of gradients: 0.1723885376262795
l2 norm of weights: 5.61397260456648

Iteration Number: 9120
Loss: 27.590994177931474
l2 norm of gradients: 0.17237570334859212
l2 norm of weights: 5.61390994940987

Iteration Number: 9121
Loss: 27.589504030689422
l2 norm of gradients: 0.17236287116987645
l2 norm of weights: 5.613847298676059

```
-----
Iteration Number: 9122
Loss: 27.588014104771652
l2 norm of gradients: 0.17235004108949867
l2 norm of weights: 5.6137846523644805
-----
Iteration Number: 9123
Loss: 27.586524400122723
l2 norm of gradients: 0.1723372131068253
l2 norm of weights: 5.61372201047457
-----
Iteration Number: 9124
Loss: 27.58503491668865
l2 norm of gradients: 0.1723243872212232
l2 norm of weights: 5.613659373005763
-----
Iteration Number: 9125
Loss: 27.583545654417044
l2 norm of gradients: 0.17231156343205944
l2 norm of weights: 5.613596739957497
-----
Iteration Number: 9126
Loss: 27.582056613256533
l2 norm of gradients: 0.17229874173870152
l2 norm of weights: 5.613534111329205
-----
Iteration Number: 9127
Loss: 27.5805677931534
l2 norm of gradients: 0.17228592214051722
l2 norm of weights: 5.613471487120325
-----
Iteration Number: 9128
Loss: 27.579079194054664
l2 norm of gradients: 0.17227310463687465
l2 norm of weights: 5.613408867330292
-----
Iteration Number: 9129
Loss: 27.577590815907875
l2 norm of gradients: 0.1722602892271421
l2 norm of weights: 5.613346251958541
-----
Iteration Number: 9130
Loss: 27.57610265865742
l2 norm of gradients: 0.1722474759106884
l2 norm of weights: 5.613283641004511
-----
Iteration Number: 9131
Loss: 27.574614722252793
```

```

12 norm of gradients: 0.17223466468688248
12 norm of weights: 5.613221034467636
-----
Iteration Number: 9132
Loss: 27.573127006642626
12 norm of gradients: 0.17222185555509378
12 norm of weights: 5.613158432347353
-----
Iteration Number: 9133
Loss: 27.57163951177119
12 norm of gradients: 0.17220904851469182
12 norm of weights: 5.613095834643099
-----
Iteration Number: 9134
Loss: 27.570152237587322
12 norm of gradients: 0.17219624356504665
12 norm of weights: 5.613033241354311
-----
Iteration Number: 9135
Loss: 27.568665184036348
12 norm of gradients: 0.17218344070552852
12 norm of weights: 5.612970652480426
-----
Iteration Number: 9136
Loss: 27.56717835106817
12 norm of gradients: 0.172170639935508
12 norm of weights: 5.612908068020879
-----
Iteration Number: 9137
Loss: 27.56569173863002
12 norm of gradients: 0.172157841254356
12 norm of weights: 5.612845487975109
-----
Iteration Number: 9138
Loss: 27.56420534666568
12 norm of gradients: 0.17214504466144373
12 norm of weights: 5.612782912342552
-----
Iteration Number: 9139
Loss: 27.56271917512701
12 norm of gradients: 0.17213225015614272
12 norm of weights: 5.612720341122646
-----
Iteration Number: 9140
Loss: 27.56123322395777
12 norm of gradients: 0.17211945773782475
12 norm of weights: 5.61265777431483
-----

```

Iteration Number: 9141
Loss: 27.5597474931071
l2 norm of gradients: 0.17210666740586195
l2 norm of weights: 5.612595211918538

Iteration Number: 9142
Loss: 27.558261982522367
l2 norm of gradients: 0.17209387915962684
l2 norm of weights: 5.612532653933209

Iteration Number: 9143
Loss: 27.556776692148375
l2 norm of gradients: 0.17208109299849214
l2 norm of weights: 5.612470100358283

Iteration Number: 9144
Loss: 27.555291621936615
l2 norm of gradients: 0.1720683089218309
l2 norm of weights: 5.612407551193195

Iteration Number: 9145
Loss: 27.553806771831777
l2 norm of gradients: 0.17205552692901652
l2 norm of weights: 5.612345006437385

Iteration Number: 9146
Loss: 27.55232214177933
l2 norm of gradients: 0.17204274701942265
l2 norm of weights: 5.6122824660902895

Iteration Number: 9147
Loss: 27.550837731732162
l2 norm of gradients: 0.17202996919242328
l2 norm of weights: 5.612219930151349

Iteration Number: 9148
Loss: 27.5493535416353
l2 norm of gradients: 0.17201719344739275
l2 norm of weights: 5.61215739862

Iteration Number: 9149
Loss: 27.547869571433104
l2 norm of gradients: 0.17200441978370565
l2 norm of weights: 5.612094871495681

Iteration Number: 9150
Loss: 27.54638582107879
l2 norm of gradients: 0.17199164820073687


```
12 norm of weights: 5.6120323487778325
-----
Iteration Number: 9151
Loss: 27.54490229051547
12 norm of gradients: 0.17197887869786166
12 norm of weights: 5.611969830465891
-----
Iteration Number: 9152
Loss: 27.543418979692124
12 norm of gradients: 0.1719661112744555
12 norm of weights: 5.6119073165592965
-----
Iteration Number: 9153
Loss: 27.541935888556154
12 norm of gradients: 0.17195334592989428
12 norm of weights: 5.611844807057488
-----
Iteration Number: 9154
Loss: 27.540453017056823
12 norm of gradients: 0.1719405826635541
12 norm of weights: 5.611782301959905
-----
Iteration Number: 9155
Loss: 27.538970365138486
12 norm of gradients: 0.17192782147481148
12 norm of weights: 5.611719801265984
-----
Iteration Number: 9156
Loss: 27.53748793275085
12 norm of gradients: 0.17191506236304305
12 norm of weights: 5.611657304975169
-----
Iteration Number: 9157
Loss: 27.53600571984222
12 norm of gradients: 0.17190230532762596
12 norm of weights: 5.611594813086896
-----
Iteration Number: 9158
Loss: 27.534523726358625
12 norm of gradients: 0.17188955036793752
12 norm of weights: 5.611532325600605
-----
Iteration Number: 9159
Loss: 27.5330419522475
12 norm of gradients: 0.17187679748335538
12 norm of weights: 5.611469842515737
-----
Iteration Number: 9160
```

Loss: 27.53156039745659
l2 norm of gradients: 0.1718640466732576
l2 norm of weights: 5.611407363831731

Iteration Number: 9161
Loss: 27.530079061935865
l2 norm of gradients: 0.17185129793702236
l2 norm of weights: 5.611344889548026

Iteration Number: 9162
Loss: 27.528597945631585
l2 norm of gradients: 0.1718385512740283
l2 norm of weights: 5.6112824196640645

Iteration Number: 9163
Loss: 27.52711704849098
l2 norm of gradients: 0.17182580668365433
l2 norm of weights: 5.611219954179285

Iteration Number: 9164
Loss: 27.525636370462863
l2 norm of gradients: 0.1718130641652795
l2 norm of weights: 5.611157493093128

Iteration Number: 9165
Loss: 27.524155911494113
l2 norm of gradients: 0.17180032371828338
l2 norm of weights: 5.611095036405034

Iteration Number: 9166
Loss: 27.52267567153232
l2 norm of gradients: 0.1717875853420458
l2 norm of weights: 5.611032584114443

Iteration Number: 9167
Loss: 27.52119565052609
l2 norm of gradients: 0.17177484903594675
l2 norm of weights: 5.610970136220796

Iteration Number: 9168
Loss: 27.51971584842308
l2 norm of gradients: 0.17176211479936673
l2 norm of weights: 5.6109076927235355

Iteration Number: 9169
Loss: 27.51823626517134
l2 norm of gradients: 0.1717493826316864
l2 norm of weights: 5.6108452536221

```

-----
Iteration Number: 9170
Loss: 27.51675690071789
l2 norm of gradients: 0.1717366525322867
l2 norm of weights: 5.610782818915931
-----
Iteration Number: 9171
Loss: 27.515277755012864
l2 norm of gradients: 0.171723924500549
l2 norm of weights: 5.610720388604471
-----
Iteration Number: 9172
Loss: 27.513798828000244
l2 norm of gradients: 0.17171119853585487
l2 norm of weights: 5.610657962687159
-----
Iteration Number: 9173
Loss: 27.512320119632516
l2 norm of gradients: 0.17169847463758617
l2 norm of weights: 5.6105955411634385
-----
Iteration Number: 9174
Loss: 27.5108416298545
l2 norm of gradients: 0.17168575280512519
l2 norm of weights: 5.6105331240327505
-----
Iteration Number: 9175
Loss: 27.50936335861487
l2 norm of gradients: 0.17167303303785436
l2 norm of weights: 5.610470711294536
-----
Iteration Number: 9176
Loss: 27.50788530586209
l2 norm of gradients: 0.1716603153351565
l2 norm of weights: 5.610408302948236
-----
Iteration Number: 9177
Loss: 27.50640747154425
l2 norm of gradients: 0.1716475996964147
l2 norm of weights: 5.610345898993295
-----
Iteration Number: 9178
Loss: 27.504929855608943
l2 norm of gradients: 0.17163488612101235
l2 norm of weights: 5.610283499429152
-----
Iteration Number: 9179
Loss: 27.50345245800361

```

```

12 norm of gradients: 0.17162217460833318
12 norm of weights: 5.610221104255252
-----
Iteration Number: 9180
Loss: 27.501975278677513
12 norm of gradients: 0.17160946515776113
12 norm of weights: 5.6101587134710345
-----
Iteration Number: 9181
Loss: 27.500498317578586
12 norm of gradients: 0.17159675776868055
12 norm of weights: 5.610096327075943
-----
Iteration Number: 9182
Loss: 27.499021574653135
12 norm of gradients: 0.171584052440476
12 norm of weights: 5.6100339450694205
-----
Iteration Number: 9183
Loss: 27.497545049851052
12 norm of gradients: 0.1715713491725323
12 norm of weights: 5.609971567450908
-----
Iteration Number: 9184
Loss: 27.496068743122066
12 norm of gradients: 0.17155864796423478
12 norm of weights: 5.60990919421985
-----
Iteration Number: 9185
Loss: 27.4945926544101
12 norm of gradients: 0.17154594881496885
12 norm of weights: 5.609846825375688
-----
Iteration Number: 9186
Loss: 27.493116783666906
12 norm of gradients: 0.1715332517241202
12 norm of weights: 5.609784460917867
-----
Iteration Number: 9187
Loss: 27.49164113083982
12 norm of gradients: 0.17152055669107505
12 norm of weights: 5.6097221008458265
-----
Iteration Number: 9188
Loss: 27.490165695877188
12 norm of gradients: 0.17150786371521964
12 norm of weights: 5.6096597451590124
-----

```

Iteration Number: 9189
Loss: 27.488690478725463
l2 norm of gradients: 0.17149517279594073
l2 norm of weights: 5.609597393856867

Iteration Number: 9190
Loss: 27.48721547933321
l2 norm of gradients: 0.17148248393262527
l2 norm of weights: 5.609535046938834

Iteration Number: 9191
Loss: 27.48574069765162
l2 norm of gradients: 0.17146979712466048
l2 norm of weights: 5.609472704404357

Iteration Number: 9192
Loss: 27.484266133626992
l2 norm of gradients: 0.17145711237143393
l2 norm of weights: 5.609410366252879

Iteration Number: 9193
Loss: 27.482791787206732
l2 norm of gradients: 0.1714444296723335
l2 norm of weights: 5.609348032483845

Iteration Number: 9194
Loss: 27.481317658341418
l2 norm of gradients: 0.17143174902674724
l2 norm of weights: 5.609285703096697

Iteration Number: 9195
Loss: 27.47984374697579
l2 norm of gradients: 0.17141907043406368
l2 norm of weights: 5.6092233780908805

Iteration Number: 9196
Loss: 27.478370053062754
l2 norm of gradients: 0.17140639389367146
l2 norm of weights: 5.6091610574658395

Iteration Number: 9197
Loss: 27.476896576547396
l2 norm of gradients: 0.1713937194049597
l2 norm of weights: 5.609098741221017

Iteration Number: 9198
Loss: 27.475423317379303
l2 norm of gradients: 0.17138104696731757

```
12 norm of weights: 5.609036429355859
-----
Iteration Number: 9199
Loss: 27.473950275506972
12 norm of gradients: 0.17136837658013485
12 norm of weights: 5.608974121869809
-----
Iteration Number: 9200
Loss: 27.472477450877815
12 norm of gradients: 0.17135570824280133
12 norm of weights: 5.608911818762311
-----
Iteration Number: 9201
Loss: 27.471004843442973
12 norm of gradients: 0.17134304195470723
12 norm of weights: 5.608849520032812
-----
Iteration Number: 9202
Loss: 27.469532453147142
12 norm of gradients: 0.171330377715243
12 norm of weights: 5.6087872256807545
-----
Iteration Number: 9203
Loss: 27.468060279940975
12 norm of gradients: 0.17131771552379943
12 norm of weights: 5.608724935705584
-----
Iteration Number: 9204
Loss: 27.46658832377386
12 norm of gradients: 0.17130505537976765
12 norm of weights: 5.608662650106745
-----
Iteration Number: 9205
Loss: 27.465116584593464
12 norm of gradients: 0.17129239728253898
12 norm of weights: 5.6086003688836845
-----
Iteration Number: 9206
Loss: 27.46364506234731
12 norm of gradients: 0.171279741231505
12 norm of weights: 5.608538092035847
-----
Iteration Number: 9207
Loss: 27.462173756983404
12 norm of gradients: 0.1712670872260577
12 norm of weights: 5.608475819562677
-----
Iteration Number: 9208
```

Loss: 27.460702668454758
l2 norm of gradients: 0.17125443526558937
l2 norm of weights: 5.60841355146362

Iteration Number: 9209
Loss: 27.459231796703545
l2 norm of gradients: 0.17124178534949247
l2 norm of weights: 5.608351287738123

Iteration Number: 9210
Loss: 27.457761141684216
l2 norm of gradients: 0.17122913747715976
l2 norm of weights: 5.608289028385631

Iteration Number: 9211
Loss: 27.45629070334177
l2 norm of gradients: 0.17121649164798441
l2 norm of weights: 5.6082267734055895

Iteration Number: 9212
Loss: 27.45482048162619
l2 norm of gradients: 0.17120384786135973
l2 norm of weights: 5.608164522797445

Iteration Number: 9213
Loss: 27.453350476485937
l2 norm of gradients: 0.17119120611667954
l2 norm of weights: 5.608102276560643

Iteration Number: 9214
Loss: 27.45188068786954
l2 norm of gradients: 0.17117856641333767
l2 norm of weights: 5.608040034694632

Iteration Number: 9215
Loss: 27.45041111572721
l2 norm of gradients: 0.1711659287507284
l2 norm of weights: 5.607977797198855

Iteration Number: 9216
Loss: 27.448941760005013
l2 norm of gradients: 0.17115329312824623
l2 norm of weights: 5.60791556407276

Iteration Number: 9217
Loss: 27.447472620653645
l2 norm of gradients: 0.17114065954528615
l2 norm of weights: 5.607853335315794

```

-----
Iteration Number: 9218
Loss: 27.446003697621645
l2 norm of gradients: 0.1711280280012431
l2 norm of weights: 5.607791110927404
-----
Iteration Number: 9219
Loss: 27.444534990856233
l2 norm of gradients: 0.1711153984955125
l2 norm of weights: 5.607728890907035
-----
Iteration Number: 9220
Loss: 27.4430665003087
l2 norm of gradients: 0.17110277102749014
l2 norm of weights: 5.607666675254136
-----
Iteration Number: 9221
Loss: 27.441598225925418
l2 norm of gradients: 0.1710901455965719
l2 norm of weights: 5.607604463968153
-----
Iteration Number: 9222
Loss: 27.440130167658594
l2 norm of gradients: 0.17107752220215408
l2 norm of weights: 5.607542257048532
-----
Iteration Number: 9223
Loss: 27.438662325452345
l2 norm of gradients: 0.17106490084363318
l2 norm of weights: 5.607480054494723
-----
Iteration Number: 9224
Loss: 27.437194699260232
l2 norm of gradients: 0.1710522815204061
l2 norm of weights: 5.607417856306172
-----
Iteration Number: 9225
Loss: 27.435727289028108
l2 norm of gradients: 0.17103966423186992
l2 norm of weights: 5.607355662482325
-----
Iteration Number: 9226
Loss: 27.4342600947058
l2 norm of gradients: 0.17102704897742207
l2 norm of weights: 5.607293473022633
-----
Iteration Number: 9227
Loss: 27.43279311624185

```



```

12 norm of gradients: 0.17101443575646022
12 norm of weights: 5.607231287926541
-----
Iteration Number: 9228
Loss: 27.431326353586897
12 norm of gradients: 0.1710018245683823
12 norm of weights: 5.6071691071934975
-----
Iteration Number: 9229
Loss: 27.42985980668776
12 norm of gradients: 0.1709892154125866
12 norm of weights: 5.607106930822951
-----
Iteration Number: 9230
Loss: 27.428393475494584
12 norm of gradients: 0.17097660828847167
12 norm of weights: 5.607044758814349
-----
Iteration Number: 9231
Loss: 27.42692735995512
12 norm of gradients: 0.1709640031954363
12 norm of weights: 5.6069825911671405
-----
Iteration Number: 9232
Loss: 27.425461460022017
12 norm of gradients: 0.17095140013287963
12 norm of weights: 5.606920427880773
-----
Iteration Number: 9233
Loss: 27.423995775639426
12 norm of gradients: 0.17093879910020102
12 norm of weights: 5.606858268954696
-----
Iteration Number: 9234
Loss: 27.42253030675951
12 norm of gradients: 0.17092620009680012
12 norm of weights: 5.606796114388357
-----
Iteration Number: 9235
Loss: 27.421065053329897
12 norm of gradients: 0.17091360312207696
12 norm of weights: 5.606733964181205
-----
Iteration Number: 9236
Loss: 27.4196000153013
12 norm of gradients: 0.17090100817543172
12 norm of weights: 5.606671818332689
-----

```

Iteration Number: 9237
Loss: 27.41813519261885
l2 norm of gradients: 0.1708884152562649
l2 norm of weights: 5.606609676842258

Iteration Number: 9238
Loss: 27.416670585237693
l2 norm of gradients: 0.17087582436397733
l2 norm of weights: 5.606547539709362

Iteration Number: 9239
Loss: 27.41520619310331
l2 norm of gradients: 0.17086323549797003
l2 norm of weights: 5.606485406933448

Iteration Number: 9240
Loss: 27.41374201616539
l2 norm of gradients: 0.1708506486576445
l2 norm of weights: 5.606423278513966

Iteration Number: 9241
Loss: 27.412278054375275
l2 norm of gradients: 0.1708380638424022
l2 norm of weights: 5.606361154450366

Iteration Number: 9242
Loss: 27.41081430767833
l2 norm of gradients: 0.1708254810516452
l2 norm of weights: 5.606299034742097

Iteration Number: 9243
Loss: 27.409350776025345
l2 norm of gradients: 0.17081290028477558
l2 norm of weights: 5.60623691938861

Iteration Number: 9244
Loss: 27.407887459365924
l2 norm of gradients: 0.17080032154119593
l2 norm of weights: 5.606174808389352

Iteration Number: 9245
Loss: 27.40642435764818
l2 norm of gradients: 0.17078774482030895
l2 norm of weights: 5.6061127017437755

Iteration Number: 9246
Loss: 27.404961470824183
l2 norm of gradients: 0.1707751701215177

```
l2 norm of weights: 5.606050599451329
-----
Iteration Number: 9247
Loss: 27.403498798841024
l2 norm of gradients: 0.17076259744422542
l2 norm of weights: 5.605988501511463
-----
Iteration Number: 9248
Loss: 27.4020363416479
l2 norm of gradients: 0.17075002678783585
l2 norm of weights: 5.605926407923628
-----
Iteration Number: 9249
Loss: 27.400574099194465
l2 norm of gradients: 0.17073745815175276
l2 norm of weights: 5.605864318687274
-----
Iteration Number: 9250
Loss: 27.399112071430668
l2 norm of gradients: 0.1707248915353803
l2 norm of weights: 5.605802233801851
-----
Iteration Number: 9251
Loss: 27.397650258307216
l2 norm of gradients: 0.170712326938123
l2 norm of weights: 5.605740153266812
-----
Iteration Number: 9252
Loss: 27.396188659769845
l2 norm of gradients: 0.17069976435938547
l2 norm of weights: 5.6056780770816035
-----
Iteration Number: 9253
Loss: 27.39472727577041
l2 norm of gradients: 0.1706872037985727
l2 norm of weights: 5.605616005245679
-----
Iteration Number: 9254
Loss: 27.393266106258636
l2 norm of gradients: 0.17067464525509
l2 norm of weights: 5.605553937758489
-----
Iteration Number: 9255
Loss: 27.391805151180677
l2 norm of gradients: 0.17066208872834285
l2 norm of weights: 5.605491874619485
-----
Iteration Number: 9256
```

```

Loss: 27.39034441049126
l2 norm of gradients: 0.17064953421773713
l2 norm of weights: 5.605429815828117
-----
Iteration Number: 9257
Loss: 27.388883884136682
l2 norm of gradients: 0.17063698172267885
l2 norm of weights: 5.605367761383837
-----
Iteration Number: 9258
Loss: 27.387423572066474
l2 norm of gradients: 0.17062443124257448
l2 norm of weights: 5.605305711286096
-----
Iteration Number: 9259
Loss: 27.385963474229904
l2 norm of gradients: 0.17061188277683062
l2 norm of weights: 5.6052436655343465
-----
Iteration Number: 9260
Loss: 27.38450359057766
l2 norm of gradients: 0.17059933632485416
l2 norm of weights: 5.605181624128039
-----
Iteration Number: 9261
Loss: 27.383043921058366
l2 norm of gradients: 0.1705867918860523
l2 norm of weights: 5.605119587066625
-----
Iteration Number: 9262
Loss: 27.381584465621813
l2 norm of gradients: 0.17057424945983252
l2 norm of weights: 5.605057554349557
-----
Iteration Number: 9263
Loss: 27.380125224217394
l2 norm of gradients: 0.17056170904560253
l2 norm of weights: 5.604995525976287
-----
Iteration Number: 9264
Loss: 27.37866619679724
l2 norm of gradients: 0.1705491706427704
l2 norm of weights: 5.604933501946267
-----
Iteration Number: 9265
Loss: 27.3772073833071
l2 norm of gradients: 0.17053663425074436
l2 norm of weights: 5.6048714822589485

```

```

-----
Iteration Number: 9266
Loss: 27.375748783697613
l2 norm of gradients: 0.17052409986893302
l2 norm of weights: 5.6048094669137845
-----
Iteration Number: 9267
Loss: 27.374290397920365
l2 norm of gradients: 0.17051156749674523
l2 norm of weights: 5.6047474559102275
-----
Iteration Number: 9268
Loss: 27.372832225924235
l2 norm of gradients: 0.17049903713359
l2 norm of weights: 5.60468544924773
-----
Iteration Number: 9269
Loss: 27.37137426765701
l2 norm of gradients: 0.1704865087788768
l2 norm of weights: 5.604623446925744
-----
Iteration Number: 9270
Loss: 27.36991652307114
l2 norm of gradients: 0.17047398243201528
l2 norm of weights: 5.6045614489437225
-----
Iteration Number: 9271
Loss: 27.36845899211382
l2 norm of gradients: 0.1704614580924153
l2 norm of weights: 5.604499455301119
-----
Iteration Number: 9272
Loss: 27.367001674736276
l2 norm of gradients: 0.1704489357594872
l2 norm of weights: 5.6044374659973855
-----
Iteration Number: 9273
Loss: 27.365544570888936
l2 norm of gradients: 0.17043641543264126
l2 norm of weights: 5.604375481031976
-----
Iteration Number: 9274
Loss: 27.364087680519095
l2 norm of gradients: 0.1704238971112883
l2 norm of weights: 5.604313500404343
-----
Iteration Number: 9275
Loss: 27.36263100358059

```

```

12 norm of gradients: 0.17041138079483942
12 norm of weights: 5.604251524113941
-----
Iteration Number: 9276
Loss: 27.361174540019633
12 norm of gradients: 0.17039886648270575
12 norm of weights: 5.604189552160222
-----
Iteration Number: 9277
Loss: 27.3597182897865
12 norm of gradients: 0.17038635417429893
12 norm of weights: 5.604127584542641
-----
Iteration Number: 9278
Loss: 27.358262252832006
12 norm of gradients: 0.1703738438690308
12 norm of weights: 5.604065621260651
-----
Iteration Number: 9279
Loss: 27.356806429107124
12 norm of gradients: 0.17036133556631342
12 norm of weights: 5.604003662313705
-----
Iteration Number: 9280
Loss: 27.35535081855909
12 norm of gradients: 0.1703488292655591
12 norm of weights: 5.603941707701258
-----
Iteration Number: 9281
Loss: 27.35389542113812
12 norm of gradients: 0.17033632496618054
12 norm of weights: 5.603879757422764
-----
Iteration Number: 9282
Loss: 27.352440236796287
12 norm of gradients: 0.17032382266759064
12 norm of weights: 5.603817811477676
-----
Iteration Number: 9283
Loss: 27.35098526548245
12 norm of gradients: 0.17031132236920252
12 norm of weights: 5.60375586986545
-----
Iteration Number: 9284
Loss: 27.349530507148195
12 norm of gradients: 0.17029882407042965
12 norm of weights: 5.603693932585539
-----

```

Iteration Number: 9285
Loss: 27.348075961738164
l2 norm of gradients: 0.17028632777068575
l2 norm of weights: 5.603631999637399

Iteration Number: 9286
Loss: 27.34662162920886
l2 norm of gradients: 0.17027383346938477
l2 norm of weights: 5.603570071020482

Iteration Number: 9287
Loss: 27.34516750950595
l2 norm of gradients: 0.17026134116594088
l2 norm of weights: 5.603508146734245

Iteration Number: 9288
Loss: 27.343713602580863
l2 norm of gradients: 0.17024885085976874
l2 norm of weights: 5.603446226778142

Iteration Number: 9289
Loss: 27.342259908384808
l2 norm of gradients: 0.170236362550283
l2 norm of weights: 5.603384311151628

Iteration Number: 9290
Loss: 27.340806426865775
l2 norm of gradients: 0.17022387623689875
l2 norm of weights: 5.603322399854158

Iteration Number: 9291
Loss: 27.33935315797558
l2 norm of gradients: 0.1702113919190313
l2 norm of weights: 5.603260492885188

Iteration Number: 9292
Loss: 27.337900101663642
l2 norm of gradients: 0.1701989095960962
l2 norm of weights: 5.6031985902441726

Iteration Number: 9293
Loss: 27.336447257878635
l2 norm of gradients: 0.17018642926750932
l2 norm of weights: 5.603136691930568

Iteration Number: 9294
Loss: 27.334994626572243
l2 norm of gradients: 0.1701739509326867

```

12 norm of weights: 5.603074797943828
-----
Iteration Number: 9295
Loss: 27.333542207696805
12 norm of gradients: 0.1701614745910448
12 norm of weights: 5.603012908283409
-----
Iteration Number: 9296
Loss: 27.332090001197198
12 norm of gradients: 0.17014900024200022
12 norm of weights: 5.602951022948767
-----
Iteration Number: 9297
Loss: 27.330638007026458
12 norm of gradients: 0.17013652788496988
12 norm of weights: 5.602889141939358
-----
Iteration Number: 9298
Loss: 27.329186225138006
12 norm of gradients: 0.17012405751937087
12 norm of weights: 5.602827265254638
-----
Iteration Number: 9299
Loss: 27.327734655476036
12 norm of gradients: 0.17011158914462068
12 norm of weights: 5.602765392894064
-----
Iteration Number: 9300
Loss: 27.32628329799436
12 norm of gradients: 0.170099122760137
12 norm of weights: 5.602703524857089
-----
Iteration Number: 9301
Loss: 27.324832152644156
12 norm of gradients: 0.17008665836533773
12 norm of weights: 5.602641661143173
-----
Iteration Number: 9302
Loss: 27.32338121937289
12 norm of gradients: 0.1700741959596412
12 norm of weights: 5.60257980175177
-----
Iteration Number: 9303
Loss: 27.321930498130907
12 norm of gradients: 0.1700617355424658
12 norm of weights: 5.6025179466823385
-----
Iteration Number: 9304

```


Loss: 27.320479988869927
l2 norm of gradients: 0.1700492771132303
l2 norm of weights: 5.602456095934334

Iteration Number: 9305
Loss: 27.319029691542216
l2 norm of gradients: 0.1700368206713537
l2 norm of weights: 5.602394249507213

Iteration Number: 9306
Loss: 27.3175796060936
l2 norm of gradients: 0.1700243662162553
l2 norm of weights: 5.602332407400433

Iteration Number: 9307
Loss: 27.316129732477396
l2 norm of gradients: 0.17001191374735464
l2 norm of weights: 5.602270569613451

Iteration Number: 9308
Loss: 27.31468007064486
l2 norm of gradients: 0.16999946326407148
l2 norm of weights: 5.602208736145725

Iteration Number: 9309
Loss: 27.313230620543553
l2 norm of gradients: 0.16998701476582587
l2 norm of weights: 5.60214690699671

Iteration Number: 9310
Loss: 27.311781382125226
l2 norm of gradients: 0.16997456825203816
l2 norm of weights: 5.602085082165865

Iteration Number: 9311
Loss: 27.31033235533888
l2 norm of gradients: 0.1699621237221289
l2 norm of weights: 5.602023261652647

Iteration Number: 9312
Loss: 27.30888354013749
l2 norm of gradients: 0.169949681175519
l2 norm of weights: 5.6019614454565145

Iteration Number: 9313
Loss: 27.307434936472625
l2 norm of gradients: 0.1699372406116295
l2 norm of weights: 5.601899633576924

```

-----
Iteration Number: 9314
Loss: 27.30598654428918
l2 norm of gradients: 0.16992480202988167
l2 norm of weights: 5.601837826013333
-----
Iteration Number: 9315
Loss: 27.304538363544125
l2 norm of gradients: 0.16991236542969732
l2 norm of weights: 5.6017760227652005
-----
Iteration Number: 9316
Loss: 27.30309039418328
l2 norm of gradients: 0.16989993081049817
l2 norm of weights: 5.601714223831984
-----
Iteration Number: 9317
Loss: 27.301642636158228
l2 norm of gradients: 0.1698874981717064
l2 norm of weights: 5.601652429213143
-----
Iteration Number: 9318
Loss: 27.30019508942176
l2 norm of gradients: 0.16987506751274442
l2 norm of weights: 5.601590638908134
-----
Iteration Number: 9319
Loss: 27.298747753919663
l2 norm of gradients: 0.16986263883303493
l2 norm of weights: 5.601528852916416
-----
Iteration Number: 9320
Loss: 27.29730062960944
l2 norm of gradients: 0.16985021213200077
l2 norm of weights: 5.601467071237448
-----
Iteration Number: 9321
Loss: 27.295853716436078
l2 norm of gradients: 0.16983778740906516
l2 norm of weights: 5.6014052938706875
-----
Iteration Number: 9322
Loss: 27.294407014352586
l2 norm of gradients: 0.16982536466365153
l2 norm of weights: 5.601343520815594
-----
Iteration Number: 9323
Loss: 27.292960523306764

```

```

12 norm of gradients: 0.1698129438951835
12 norm of weights: 5.6012817520716265
-----
Iteration Number: 9324
Loss: 27.291514243254596
12 norm of gradients: 0.1698005251030851
12 norm of weights: 5.601219987638244
-----
Iteration Number: 9325
Loss: 27.29006817414181
12 norm of gradients: 0.16978810828678048
12 norm of weights: 5.6011582275149046
-----
Iteration Number: 9326
Loss: 27.288622315922535
12 norm of gradients: 0.16977569344569413
12 norm of weights: 5.6010964717010685
-----
Iteration Number: 9327
Loss: 27.287176668546447
12 norm of gradients: 0.16976328057925077
12 norm of weights: 5.6010347201961945
-----
Iteration Number: 9328
Loss: 27.285731231961318
12 norm of gradients: 0.1697508696868753
12 norm of weights: 5.600972972999743
-----
Iteration Number: 9329
Loss: 27.28428600612209
12 norm of gradients: 0.16973846076799304
12 norm of weights: 5.600911230111172
-----
Iteration Number: 9330
Loss: 27.282840990977146
12 norm of gradients: 0.16972605382202943
12 norm of weights: 5.600849491529942
-----
Iteration Number: 9331
Loss: 27.28139618647686
12 norm of gradients: 0.16971364884841017
12 norm of weights: 5.600787757255513
-----
Iteration Number: 9332
Loss: 27.279951592575237
12 norm of gradients: 0.16970124584656132
12 norm of weights: 5.600726027287345
-----

```

Iteration Number: 9333
Loss: 27.27850720921871
l2 norm of gradients: 0.16968884481590915
l2 norm of weights: 5.600664301624896

Iteration Number: 9334
Loss: 27.277063036362108
l2 norm of gradients: 0.16967644575588003
l2 norm of weights: 5.600602580267628

Iteration Number: 9335
Loss: 27.2756190739542
l2 norm of gradients: 0.16966404866590093
l2 norm of weights: 5.600540863215002

Iteration Number: 9336
Loss: 27.274175321945936
l2 norm of gradients: 0.16965165354539863
l2 norm of weights: 5.600479150466477

Iteration Number: 9337
Loss: 27.272731780288034
l2 norm of gradients: 0.16963926039380053
l2 norm of weights: 5.600417442021513

Iteration Number: 9338
Loss: 27.27128844893272
l2 norm of gradients: 0.16962686921053413
l2 norm of weights: 5.60035573787957

Iteration Number: 9339
Loss: 27.26984532782906
l2 norm of gradients: 0.1696144799950272
l2 norm of weights: 5.6002940380401105

Iteration Number: 9340
Loss: 27.268402416930314
l2 norm of gradients: 0.16960209274670773
l2 norm of weights: 5.600232342502595

Iteration Number: 9341
Loss: 27.266959716185717
l2 norm of gradients: 0.16958970746500404
l2 norm of weights: 5.600170651266485

Iteration Number: 9342
Loss: 27.265517225546038
l2 norm of gradients: 0.16957732414934465

```

12 norm of weights: 5.6001089643312385
-----
Iteration Number: 9343
Loss: 27.264074944962424
12 norm of gradients: 0.16956494279915835
12 norm of weights: 5.60004728169632
-----
Iteration Number: 9344
Loss: 27.262632874387034
12 norm of gradients: 0.16955256341387417
12 norm of weights: 5.599985603361189
-----
Iteration Number: 9345
Loss: 27.261191013770283
12 norm of gradients: 0.1695401859929214
12 norm of weights: 5.599923929325307
-----
Iteration Number: 9346
Loss: 27.259749363062447
12 norm of gradients: 0.16952781053572952
12 norm of weights: 5.599862259588135
-----
Iteration Number: 9347
Loss: 27.258307922213522
12 norm of gradients: 0.16951543704172842
12 norm of weights: 5.599800594149136
-----
Iteration Number: 9348
Loss: 27.256866691178352
12 norm of gradients: 0.16950306551034808
12 norm of weights: 5.599738933007771
-----
Iteration Number: 9349
Loss: 27.25542566990569
12 norm of gradients: 0.1694906959410188
12 norm of weights: 5.599677276163501
-----
Iteration Number: 9350
Loss: 27.25398485834714
12 norm of gradients: 0.16947832833317109
12 norm of weights: 5.599615623615789
-----
Iteration Number: 9351
Loss: 27.252544256453106
12 norm of gradients: 0.16946596268623576
12 norm of weights: 5.599553975364096
-----
Iteration Number: 9352

```

Loss: 27.251103864174365
l2 norm of gradients: 0.16945359899964388
l2 norm of weights: 5.599492331407886

Iteration Number: 9353
Loss: 27.24966368146396
l2 norm of gradients: 0.16944123727282673
l2 norm of weights: 5.599430691746619

Iteration Number: 9354
Loss: 27.248223708272473
l2 norm of gradients: 0.16942887750521582
l2 norm of weights: 5.599369056379759

Iteration Number: 9355
Loss: 27.246783944548433
l2 norm of gradients: 0.16941651969624297
l2 norm of weights: 5.599307425306767

Iteration Number: 9356
Loss: 27.245344390248086
l2 norm of gradients: 0.16940416384534018
l2 norm of weights: 5.599245798527107

Iteration Number: 9357
Loss: 27.24390504531924
l2 norm of gradients: 0.1693918099519398
l2 norm of weights: 5.599184176040241

Iteration Number: 9358
Loss: 27.242465909712276
l2 norm of gradients: 0.16937945801547433
l2 norm of weights: 5.599122557845633

Iteration Number: 9359
Loss: 27.24102698338044
l2 norm of gradients: 0.16936710803537652
l2 norm of weights: 5.599060943942744

Iteration Number: 9360
Loss: 27.23958826627347
l2 norm of gradients: 0.16935476001107938
l2 norm of weights: 5.598999334331038

Iteration Number: 9361
Loss: 27.23814975834396
l2 norm of gradients: 0.16934241394201632
l2 norm of weights: 5.598937729009978

```

-----
Iteration Number: 9362
Loss: 27.236711459543777
l2 norm of gradients: 0.16933006982762072
l2 norm of weights: 5.598876127979028
-----
Iteration Number: 9363
Loss: 27.235273369822675
l2 norm of gradients: 0.16931772766732642
l2 norm of weights: 5.59881453123765
-----
Iteration Number: 9364
Loss: 27.233835489134755
l2 norm of gradients: 0.1693053874605674
l2 norm of weights: 5.598752938785308
-----
Iteration Number: 9365
Loss: 27.23239781742778
l2 norm of gradients: 0.16929304920677793
l2 norm of weights: 5.598691350621467
-----
Iteration Number: 9366
Loss: 27.23096035465511
l2 norm of gradients: 0.16928071290539257
l2 norm of weights: 5.598629766745589
-----
Iteration Number: 9367
Loss: 27.229523100765903
l2 norm of gradients: 0.16926837855584603
l2 norm of weights: 5.598568187157138
-----
Iteration Number: 9368
Loss: 27.22808605571598
l2 norm of gradients: 0.16925604615757334
l2 norm of weights: 5.598506611855578
-----
Iteration Number: 9369
Loss: 27.226649219452955
l2 norm of gradients: 0.16924371571000968
l2 norm of weights: 5.598445040840375
-----
Iteration Number: 9370
Loss: 27.225212591930692
l2 norm of gradients: 0.1692313872125906
l2 norm of weights: 5.59838347411099
-----
Iteration Number: 9371
Loss: 27.223776173097747

```

```

12 norm of gradients: 0.1692190606647518
12 norm of weights: 5.59832191166689
-----
Iteration Number: 9372
Loss: 27.222339962907537
12 norm of gradients: 0.1692067360659293
12 norm of weights: 5.598260353507537
-----
Iteration Number: 9373
Loss: 27.220903961313425
12 norm of gradients: 0.16919441341555932
12 norm of weights: 5.598198799632398
-----
Iteration Number: 9374
Loss: 27.219468168262445
12 norm of gradients: 0.16918209271307832
12 norm of weights: 5.598137250040935
-----
Iteration Number: 9375
Loss: 27.218032583710322
12 norm of gradients: 0.16916977395792296
12 norm of weights: 5.598075704732615
-----
Iteration Number: 9376
Loss: 27.21659720760544
12 norm of gradients: 0.1691574571495303
12 norm of weights: 5.598014163706901
-----
Iteration Number: 9377
Loss: 27.215162039901692
12 norm of gradients: 0.16914514228733743
12 norm of weights: 5.59795262696326
-----
Iteration Number: 9378
Loss: 27.21372708054807
12 norm of gradients: 0.16913282937078186
12 norm of weights: 5.597891094501156
-----
Iteration Number: 9379
Loss: 27.21229232950068
12 norm of gradients: 0.16912051839930126
12 norm of weights: 5.597829566320053
-----
Iteration Number: 9380
Loss: 27.210857786706182
12 norm of gradients: 0.16910820937233356
12 norm of weights: 5.5977680424194185
-----

```


Iteration Number: 9381
Loss: 27.2094234521187
l2 norm of gradients: 0.1690959022893169
l2 norm of weights: 5.597706522798717

Iteration Number: 9382
Loss: 27.207989325689752
l2 norm of gradients: 0.16908359714968973
l2 norm of weights: 5.5976450074574124

Iteration Number: 9383
Loss: 27.206555407370217
l2 norm of gradients: 0.16907129395289067
l2 norm of weights: 5.597583496394972

Iteration Number: 9384
Loss: 27.205121697113345
l2 norm of gradients: 0.16905899269835864
l2 norm of weights: 5.597521989610862

Iteration Number: 9385
Loss: 27.203688194868864
l2 norm of gradients: 0.16904669338553274
l2 norm of weights: 5.597460487104548

Iteration Number: 9386
Loss: 27.202254900590916
l2 norm of gradients: 0.1690343960138524
l2 norm of weights: 5.5973989888754945

Iteration Number: 9387
Loss: 27.200821814228103
l2 norm of gradients: 0.16902210058275713
l2 norm of weights: 5.597337494923169

Iteration Number: 9388
Loss: 27.199388935735666
l2 norm of gradients: 0.1690098070916869
l2 norm of weights: 5.5972760052470365

Iteration Number: 9389
Loss: 27.197956265062423
l2 norm of gradients: 0.1689975155400818
l2 norm of weights: 5.597214519846566

Iteration Number: 9390
Loss: 27.19652380216187
l2 norm of gradients: 0.16898522592738208

```

12 norm of weights: 5.5971530387212205
-----
Iteration Number: 9391
Loss: 27.19509154698397
12 norm of gradients: 0.16897293825302837
12 norm of weights: 5.597091561870468
-----
Iteration Number: 9392
Loss: 27.19365949948174
12 norm of gradients: 0.1689606525164615
12 norm of weights: 5.597030089293775
-----
Iteration Number: 9393
Loss: 27.192227659608754
12 norm of gradients: 0.16894836871712246
12 norm of weights: 5.5969686209906095
-----
Iteration Number: 9394
Loss: 27.19079602731509
12 norm of gradients: 0.1689360868544526
12 norm of weights: 5.5969071569604365
-----
Iteration Number: 9395
Loss: 27.189364602552722
12 norm of gradients: 0.16892380692789344
12 norm of weights: 5.596845697202723
-----
Iteration Number: 9396
Loss: 27.187933385272302
12 norm of gradients: 0.16891152893688674
12 norm of weights: 5.5967842417169384
-----
Iteration Number: 9397
Loss: 27.18650237542746
12 norm of gradients: 0.16889925288087454
12 norm of weights: 5.596722790502547
-----
Iteration Number: 9398
Loss: 27.18507157296896
12 norm of gradients: 0.16888697875929903
12 norm of weights: 5.596661343559018
-----
Iteration Number: 9399
Loss: 27.183640977851088
12 norm of gradients: 0.16887470657160272
12 norm of weights: 5.596599900885818
-----
Iteration Number: 9400

```

Loss: 27.18221059002472
l2 norm of gradients: 0.16886243631722833
l2 norm of weights: 5.596538462482415

Iteration Number: 9401
Loss: 27.18078040943926
l2 norm of gradients: 0.16885016799561875
l2 norm of weights: 5.596477028348276

Iteration Number: 9402
Loss: 27.17935043604882
l2 norm of gradients: 0.1688379016062173
l2 norm of weights: 5.596415598482869

Iteration Number: 9403
Loss: 27.177920669807094
l2 norm of gradients: 0.16882563714846732
l2 norm of weights: 5.596354172885662

Iteration Number: 9404
Loss: 27.176491110663054
l2 norm of gradients: 0.16881337462181253
l2 norm of weights: 5.596292751556123

Iteration Number: 9405
Loss: 27.175061758568756
l2 norm of gradients: 0.16880111402569678
l2 norm of weights: 5.59623133449372

Iteration Number: 9406
Loss: 27.173632613480542
l2 norm of gradients: 0.1687888553595642
l2 norm of weights: 5.596169921697921

Iteration Number: 9407
Loss: 27.172203675343
l2 norm of gradients: 0.1687765986228592
l2 norm of weights: 5.596108513168196

Iteration Number: 9408
Loss: 27.17077494411692
l2 norm of gradients: 0.1687643438150264
l2 norm of weights: 5.596047108904011

Iteration Number: 9409
Loss: 27.169346419749242
l2 norm of gradients: 0.1687520909355106
l2 norm of weights: 5.595985708904835

```

-----
Iteration Number: 9410
Loss: 27.167918102189475
l2 norm of gradients: 0.1687398399837569
l2 norm of weights: 5.595924313170137
-----
Iteration Number: 9411
Loss: 27.166489991396134
l2 norm of gradients: 0.16872759095921058
l2 norm of weights: 5.5958629216993865
-----
Iteration Number: 9412
Loss: 27.165062087319527
l2 norm of gradients: 0.1687153438613172
l2 norm of weights: 5.595801534492051
-----
Iteration Number: 9413
Loss: 27.16363438991024
l2 norm of gradients: 0.1687030986895226
l2 norm of weights: 5.595740151547601
-----
Iteration Number: 9414
Loss: 27.162206899119912
l2 norm of gradients: 0.1686908554432727
l2 norm of weights: 5.595678772865504
-----
Iteration Number: 9415
Loss: 27.16077961490143
l2 norm of gradients: 0.16867861412201376
l2 norm of weights: 5.595617398445231
-----
Iteration Number: 9416
Loss: 27.159352537208754
l2 norm of gradients: 0.1686663747251924
l2 norm of weights: 5.59555602828625
-----
Iteration Number: 9417
Loss: 27.15792566599295
l2 norm of gradients: 0.16865413725225512
l2 norm of weights: 5.59549466238803
-----
Iteration Number: 9418
Loss: 27.15649900120618
l2 norm of gradients: 0.16864190170264898
l2 norm of weights: 5.5954333007500425
-----
Iteration Number: 9419
Loss: 27.155072542798045

```

```

12 norm of gradients: 0.16862966807582114
12 norm of weights: 5.595371943371755
-----
Iteration Number: 9420
Loss: 27.153646290725828
12 norm of gradients: 0.16861743637121898
12 norm of weights: 5.595310590252638
-----
Iteration Number: 9421
Loss: 27.152220244939524
12 norm of gradients: 0.1686052065882902
12 norm of weights: 5.595249241392162
-----
Iteration Number: 9422
Loss: 27.150794405389995
12 norm of gradients: 0.16859297872648263
12 norm of weights: 5.595187896789796
-----
Iteration Number: 9423
Loss: 27.149368772033583
12 norm of gradients: 0.1685807527852444
12 norm of weights: 5.5951265564450114
-----
Iteration Number: 9424
Loss: 27.147943344817303
12 norm of gradients: 0.1685685287640238
12 norm of weights: 5.595065220357276
-----
Iteration Number: 9425
Loss: 27.146518123697206
12 norm of gradients: 0.1685563066622694
12 norm of weights: 5.595003888526064
-----
Iteration Number: 9426
Loss: 27.145093108626195
12 norm of gradients: 0.16854408647943003
12 norm of weights: 5.594942560950841
-----
Iteration Number: 9427
Loss: 27.14366829955318
12 norm of gradients: 0.16853186821495475
12 norm of weights: 5.594881237631081
-----
Iteration Number: 9428
Loss: 27.14224369643289
12 norm of gradients: 0.16851965186829274
12 norm of weights: 5.5948199185662535
-----

```

Iteration Number: 9429
Loss: 27.14081929921747
l2 norm of gradients: 0.16850743743889351
l2 norm of weights: 5.59475860375583

Iteration Number: 9430
Loss: 27.13939510786023
l2 norm of gradients: 0.1684952249262068
l2 norm of weights: 5.59469729319928

Iteration Number: 9431
Loss: 27.13797112231318
l2 norm of gradients: 0.1684830143296825
l2 norm of weights: 5.594635986896075

Iteration Number: 9432
Loss: 27.13654734252709
l2 norm of gradients: 0.16847080564877082
l2 norm of weights: 5.594574684845687

Iteration Number: 9433
Loss: 27.13512376845648
l2 norm of gradients: 0.16845859888292217
l2 norm of weights: 5.594513387047586

Iteration Number: 9434
Loss: 27.133700400052472
l2 norm of gradients: 0.16844639403158715
l2 norm of weights: 5.5944520935012445

Iteration Number: 9435
Loss: 27.132277237268898
l2 norm of gradients: 0.16843419109421665
l2 norm of weights: 5.5943908042061326

Iteration Number: 9436
Loss: 27.130854280058394
l2 norm of gradients: 0.16842199007026176
l2 norm of weights: 5.594329519161723

Iteration Number: 9437
Loss: 27.129431528371487
l2 norm of gradients: 0.16840979095917383
l2 norm of weights: 5.594268238367485

Iteration Number: 9438
Loss: 27.12800898216305
l2 norm of gradients: 0.16839759376040428

```
12 norm of weights: 5.594206961822894
-----
Iteration Number: 9439
Loss: 27.126586641384385
12 norm of gradients: 0.16838539847340495
12 norm of weights: 5.594145689527418
-----
Iteration Number: 9440
Loss: 27.125164505988312
12 norm of gradients: 0.16837320509762788
12 norm of weights: 5.594084421480533
-----
Iteration Number: 9441
Loss: 27.123742575928397
12 norm of gradients: 0.16836101363252526
12 norm of weights: 5.594023157681708
-----
Iteration Number: 9442
Loss: 27.122320851156026
12 norm of gradients: 0.16834882407754953
12 norm of weights: 5.5939618981304156
-----
Iteration Number: 9443
Loss: 27.120899331624333
12 norm of gradients: 0.16833663643215332
12 norm of weights: 5.59390064282613
-----
Iteration Number: 9444
Loss: 27.11947801728732
12 norm of gradients: 0.16832445069578963
12 norm of weights: 5.593839391768321
-----
Iteration Number: 9445
Loss: 27.11805690809446
12 norm of gradients: 0.16831226686791156
12 norm of weights: 5.593778144956463
-----
Iteration Number: 9446
Loss: 27.116636004002114
12 norm of gradients: 0.16830008494797244
12 norm of weights: 5.593716902390027
-----
Iteration Number: 9447
Loss: 27.115215304960056
12 norm of gradients: 0.16828790493542584
12 norm of weights: 5.593655664068487
-----
Iteration Number: 9448
```

```

Loss: 27.11379481092399
12 norm of gradients: 0.16827572682972555
12 norm of weights: 5.593594429991317
-----
Iteration Number: 9449
Loss: 27.112374521844483
12 norm of gradients: 0.16826355063032564
12 norm of weights: 5.593533200157987
-----
Iteration Number: 9450
Loss: 27.110954437673755
12 norm of gradients: 0.16825137633668036
12 norm of weights: 5.593471974567972
-----
Iteration Number: 9451
Loss: 27.109534558365162
12 norm of gradients: 0.16823920394824418
12 norm of weights: 5.593410753220744
-----
Iteration Number: 9452
Loss: 27.10811488387252
12 norm of gradients: 0.1682270334644718
12 norm of weights: 5.593349536115778
-----
Iteration Number: 9453
Loss: 27.106695414149094
12 norm of gradients: 0.16821486488481815
12 norm of weights: 5.593288323252546
-----
Iteration Number: 9454
Loss: 27.10527614914699
12 norm of gradients: 0.16820269820873834
12 norm of weights: 5.593227114630522
-----
Iteration Number: 9455
Loss: 27.103857088818167
12 norm of gradients: 0.16819053343568782
12 norm of weights: 5.59316591024918
-----
Iteration Number: 9456
Loss: 27.102438233117393
12 norm of gradients: 0.16817837056512214
12 norm of weights: 5.593104710107992
-----
Iteration Number: 9457
Loss: 27.10101958199599
12 norm of gradients: 0.1681662095964971
12 norm of weights: 5.593043514206434

```



```

-----
Iteration Number: 9458
Loss: 27.09960113540689
l2 norm of gradients: 0.1681540505292688
l2 norm of weights: 5.592982322543979
-----
Iteration Number: 9459
Loss: 27.098182893303157
l2 norm of gradients: 0.16814189336289342
l2 norm of weights: 5.592921135120101
-----
Iteration Number: 9460
Loss: 27.096764855638803
l2 norm of gradients: 0.16812973809682752
l2 norm of weights: 5.592859951934274
-----
Iteration Number: 9461
Loss: 27.0953470223649
l2 norm of gradients: 0.16811758473052774
l2 norm of weights: 5.592798772985972
-----
Iteration Number: 9462
Loss: 27.093929393435584
l2 norm of gradients: 0.1681054332634511
l2 norm of weights: 5.592737598274671
-----
Iteration Number: 9463
Loss: 27.092511968804992
l2 norm of gradients: 0.16809328369505466
l2 norm of weights: 5.592676427799844
-----
Iteration Number: 9464
Loss: 27.091094748424155
l2 norm of gradients: 0.16808113602479582
l2 norm of weights: 5.592615261560966
-----
Iteration Number: 9465
Loss: 27.089677732246653
l2 norm of gradients: 0.16806899025213223
l2 norm of weights: 5.592554099557511
-----
Iteration Number: 9466
Loss: 27.0882609202264
l2 norm of gradients: 0.1680568463765216
l2 norm of weights: 5.592492941788955
-----
Iteration Number: 9467
Loss: 27.086844312317105

```

```

12 norm of gradients: 0.16804470439742206
12 norm of weights: 5.592431788254773
-----
Iteration Number: 9468
Loss: 27.085427908468063
12 norm of gradients: 0.16803256431429175
12 norm of weights: 5.5923706389544385
-----
Iteration Number: 9469
Loss: 27.084011708635828
12 norm of gradients: 0.16802042612658927
12 norm of weights: 5.592309493887428
-----
Iteration Number: 9470
Loss: 27.082595712771926
12 norm of gradients: 0.16800828983377322
12 norm of weights: 5.592248353053216
-----
Iteration Number: 9471
Loss: 27.081179920831353
12 norm of gradients: 0.1679961554353026
12 norm of weights: 5.592187216451279
-----
Iteration Number: 9472
Loss: 27.079764332765606
12 norm of gradients: 0.16798402293063647
12 norm of weights: 5.592126084081091
-----
Iteration Number: 9473
Loss: 27.078348948528813
12 norm of gradients: 0.16797189231923412
12 norm of weights: 5.592064955942128
-----
Iteration Number: 9474
Loss: 27.076933768071733
12 norm of gradients: 0.1679597636005553
12 norm of weights: 5.592003832033867
-----
Iteration Number: 9475
Loss: 27.075518791350966
12 norm of gradients: 0.16794763677405966
12 norm of weights: 5.591942712355783
-----
Iteration Number: 9476
Loss: 27.074104018316614
12 norm of gradients: 0.16793551183920724
12 norm of weights: 5.59188159690735
-----

```

Iteration Number: 9477
Loss: 27.072689448925892
l2 norm of gradients: 0.1679233887954583
l2 norm of weights: 5.5918204856880465

Iteration Number: 9478
Loss: 27.071275083129066
l2 norm of gradients: 0.16791126764227315
l2 norm of weights: 5.591759378697348

Iteration Number: 9479
Loss: 27.069860920877424
l2 norm of gradients: 0.16789914837911263
l2 norm of weights: 5.591698275934731

Iteration Number: 9480
Loss: 27.068446962128498
l2 norm of gradients: 0.1678870310054375
l2 norm of weights: 5.5916371773996705

Iteration Number: 9481
Loss: 27.067033206832622
l2 norm of gradients: 0.16787491552070885
l2 norm of weights: 5.5915760830916446

Iteration Number: 9482
Loss: 27.0656196549446
l2 norm of gradients: 0.1678628019243881
l2 norm of weights: 5.591514993010128

Iteration Number: 9483
Loss: 27.064206306419077
l2 norm of gradients: 0.16785069021593665
l2 norm of weights: 5.5914539071546

Iteration Number: 9484
Loss: 27.062793161206844
l2 norm of gradients: 0.1678385803948163
l2 norm of weights: 5.591392825524536

Iteration Number: 9485
Loss: 27.061380219261768
l2 norm of gradients: 0.167826472460489
l2 norm of weights: 5.591331748119411

Iteration Number: 9486
Loss: 27.059967480538354
l2 norm of gradients: 0.1678143664124169

```
12 norm of weights: 5.591270674938705
-----
Iteration Number: 9487
Loss: 27.058554944987822
12 norm of gradients: 0.16780226225006245
12 norm of weights: 5.591209605981893
-----
Iteration Number: 9488
Loss: 27.057142612565237
12 norm of gradients: 0.16779015997288824
12 norm of weights: 5.591148541248455
-----
Iteration Number: 9489
Loss: 27.055730483224046
12 norm of gradients: 0.16777805958035702
12 norm of weights: 5.591087480737865
-----
Iteration Number: 9490
Loss: 27.054318556918254
12 norm of gradients: 0.1677659610719319
12 norm of weights: 5.591026424449603
-----
Iteration Number: 9491
Loss: 27.052906833599337
12 norm of gradients: 0.1677538644470761
12 norm of weights: 5.5909653723831445
-----
Iteration Number: 9492
Loss: 27.05149531322221
12 norm of gradients: 0.1677417697052531
12 norm of weights: 5.590904324537968
-----
Iteration Number: 9493
Loss: 27.050083995739463
12 norm of gradients: 0.16772967684592657
12 norm of weights: 5.5908432809135515
-----
Iteration Number: 9494
Loss: 27.04867288110656
12 norm of gradients: 0.1677175858685604
12 norm of weights: 5.590782241509372
-----
Iteration Number: 9495
Loss: 27.047261969274892
12 norm of gradients: 0.16770549677261867
12 norm of weights: 5.590721206324909
-----
Iteration Number: 9496
```

```

Loss: 27.045851260197665
12 norm of gradients: 0.16769340955756573
12 norm of weights: 5.59066017535964
-----
Iteration Number: 9497
Loss: 27.044440753828656
12 norm of gradients: 0.16768132422286613
12 norm of weights: 5.590599148613043
-----
Iteration Number: 9498
Loss: 27.043030450125173
12 norm of gradients: 0.1676692407679846
12 norm of weights: 5.5905381260845965
-----
Iteration Number: 9499
Loss: 27.04162034903519
12 norm of gradients: 0.16765715919238608
12 norm of weights: 5.590477107773777
-----
Iteration Number: 9500
Loss: 27.040210450515666
12 norm of gradients: 0.16764507949553575
12 norm of weights: 5.590416093680066
-----
Iteration Number: 9501
Loss: 27.03880075451996
12 norm of gradients: 0.16763300167689899
12 norm of weights: 5.59035508380294
-----
Iteration Number: 9502
Loss: 27.037391261000195
12 norm of gradients: 0.1676209257359414
12 norm of weights: 5.59029407814188
-----
Iteration Number: 9503
Loss: 27.035981969912488
12 norm of gradients: 0.16760885167212877
12 norm of weights: 5.590233076696363
-----
Iteration Number: 9504
Loss: 27.034572881207364
12 norm of gradients: 0.16759677948492718
12 norm of weights: 5.590172079465868
-----
Iteration Number: 9505
Loss: 27.033163994842045
12 norm of gradients: 0.1675847091738028
12 norm of weights: 5.590111086449874

```

```

-----
Iteration Number: 9506
Loss: 27.03175531076565
l2 norm of gradients: 0.1675726407382221
l2 norm of weights: 5.590050097647861
-----
Iteration Number: 9507
Loss: 27.03034682893733
l2 norm of gradients: 0.1675605741776517
l2 norm of weights: 5.589989113059309
-----
Iteration Number: 9508
Loss: 27.02893854930611
l2 norm of gradients: 0.16754850949155842
l2 norm of weights: 5.589928132683695
-----
Iteration Number: 9509
Loss: 27.02753047182773
l2 norm of gradients: 0.16753644667940948
l2 norm of weights: 5.5898671565205
-----
Iteration Number: 9510
Loss: 27.02612259645733
l2 norm of gradients: 0.16752438574067205
l2 norm of weights: 5.589806184569204
-----
Iteration Number: 9511
Loss: 27.024714923144767
l2 norm of gradients: 0.16751232667481364
l2 norm of weights: 5.589745216829285
-----
Iteration Number: 9512
Loss: 27.023307451847327
l2 norm of gradients: 0.16750026948130195
l2 norm of weights: 5.5896842533002244
-----
Iteration Number: 9513
Loss: 27.02190018251863
l2 norm of gradients: 0.1674882141596049
l2 norm of weights: 5.589623293981502
-----
Iteration Number: 9514
Loss: 27.020493115110938
l2 norm of gradients: 0.1674761607091906
l2 norm of weights: 5.589562338872597
-----
Iteration Number: 9515
Loss: 27.019086249579626

```

```

12 norm of gradients: 0.16746410912952742
12 norm of weights: 5.58950138797299
-----
Iteration Number: 9516
Loss: 27.017679585876657
12 norm of gradients: 0.16745205942008382
12 norm of weights: 5.589440441282161
-----
Iteration Number: 9517
Loss: 27.016273123955905
12 norm of gradients: 0.16744001158032865
12 norm of weights: 5.589379498799592
-----
Iteration Number: 9518
Loss: 27.01486686377298
12 norm of gradients: 0.16742796560973083
12 norm of weights: 5.589318560524761
-----
Iteration Number: 9519
Loss: 27.01346080528202
12 norm of gradients: 0.1674159215077594
12 norm of weights: 5.58925762645715
-----
Iteration Number: 9520
Loss: 27.012054948434656
12 norm of gradients: 0.16740387927388395
12 norm of weights: 5.589196696596239
-----
Iteration Number: 9521
Loss: 27.010649293186724
12 norm of gradients: 0.1673918389075739
12 norm of weights: 5.5891357709415095
-----
Iteration Number: 9522
Loss: 27.00924383949067
12 norm of gradients: 0.16737980040829906
12 norm of weights: 5.589074849492442
-----
Iteration Number: 9523
Loss: 27.007838587302135
12 norm of gradients: 0.16736776377552945
12 norm of weights: 5.589013932248518
-----
Iteration Number: 9524
Loss: 27.00643353657496
12 norm of gradients: 0.1673557290087353
12 norm of weights: 5.588953019209217
-----

```

Iteration Number: 9525
Loss: 27.00502868726028
l2 norm of gradients: 0.16734369610738695
l2 norm of weights: 5.588892110374022

Iteration Number: 9526
Loss: 27.00362403931567
l2 norm of gradients: 0.16733166507095504
l2 norm of weights: 5.588831205742414

Iteration Number: 9527
Loss: 27.002219592692704
l2 norm of gradients: 0.16731963589891038
l2 norm of weights: 5.588770305313874

Iteration Number: 9528
Loss: 27.000815347348382
l2 norm of gradients: 0.16730760859072402
l2 norm of weights: 5.588709409087884

Iteration Number: 9529
Loss: 26.999411303233767
l2 norm of gradients: 0.1672955831458672
l2 norm of weights: 5.588648517063925

Iteration Number: 9530
Loss: 26.99800746030223
l2 norm of gradients: 0.1672835595638113
l2 norm of weights: 5.58858762924148

Iteration Number: 9531
Loss: 26.996603818512042
l2 norm of gradients: 0.16727153784402804
l2 norm of weights: 5.588526745620029

Iteration Number: 9532
Loss: 26.995200377813624
l2 norm of gradients: 0.16725951798598923
l2 norm of weights: 5.588465866199055

Iteration Number: 9533
Loss: 26.993797138163174
l2 norm of gradients: 0.16724749998916688
l2 norm of weights: 5.58840499097804

Iteration Number: 9534
Loss: 26.992394099514218
l2 norm of gradients: 0.1672354838530333


```

12 norm of weights: 5.588344119956467
-----
Iteration Number: 9535
Loss: 26.990991261819406
12 norm of gradients: 0.16722346957706094
12 norm of weights: 5.588283253133817
-----
Iteration Number: 9536
Loss: 26.989588625034795
12 norm of gradients: 0.16721145716072242
12 norm of weights: 5.588222390509573
-----
Iteration Number: 9537
Loss: 26.988186189113282
12 norm of gradients: 0.16719944660349068
12 norm of weights: 5.588161532083217
-----
Iteration Number: 9538
Loss: 26.986783954010924
12 norm of gradients: 0.16718743790483873
12 norm of weights: 5.588100677854232
-----
Iteration Number: 9539
Loss: 26.985381919679195
12 norm of gradients: 0.16717543106423993
12 norm of weights: 5.588039827822102
-----
Iteration Number: 9540
Loss: 26.98398008607438
12 norm of gradients: 0.16716342608116766
12 norm of weights: 5.587978981986308
-----
Iteration Number: 9541
Loss: 26.982578453150754
12 norm of gradients: 0.16715142295509566
12 norm of weights: 5.587918140346333
-----
Iteration Number: 9542
Loss: 26.98117702086018
12 norm of gradients: 0.16713942168549784
12 norm of weights: 5.587857302901662
-----
Iteration Number: 9543
Loss: 26.979775789160392
12 norm of gradients: 0.16712742227184818
12 norm of weights: 5.587796469651775
-----
Iteration Number: 9544

```

Loss: 26.97837475800484
l2 norm of gradients: 0.16711542471362112
l2 norm of weights: 5.587735640596158

Iteration Number: 9545
Loss: 26.976973927345703
l2 norm of gradients: 0.167103429010291
l2 norm of weights: 5.587674815734294

Iteration Number: 9546
Loss: 26.975573297138027
l2 norm of gradients: 0.16709143516133265
l2 norm of weights: 5.587613995065665

Iteration Number: 9547
Loss: 26.974172867337465
l2 norm of gradients: 0.16707944316622086
l2 norm of weights: 5.587553178589756

Iteration Number: 9548
Loss: 26.972772637896306
l2 norm of gradients: 0.1670674530244308
l2 norm of weights: 5.58749236630605

Iteration Number: 9549
Loss: 26.971372608772555
l2 norm of gradients: 0.1670554647354377
l2 norm of weights: 5.58743155821403

Iteration Number: 9550
Loss: 26.969972779917466
l2 norm of gradients: 0.1670434782987171
l2 norm of weights: 5.5873707543131825

Iteration Number: 9551
Loss: 26.968573151285888
l2 norm of gradients: 0.1670314937137447
l2 norm of weights: 5.587309954602989

Iteration Number: 9552
Loss: 26.9671737228315
l2 norm of gradients: 0.16701951097999643
l2 norm of weights: 5.587249159082934

Iteration Number: 9553
Loss: 26.965774494511674
l2 norm of gradients: 0.16700753009694835
l2 norm of weights: 5.5871883677525025

```

-----
Iteration Number: 9554
Loss: 26.964375466277655
l2 norm of gradients: 0.16699555106407674
l2 norm of weights: 5.587127580611179
-----
Iteration Number: 9555
Loss: 26.962976638085333
l2 norm of gradients: 0.16698357388085816
l2 norm of weights: 5.587066797658447
-----
Iteration Number: 9556
Loss: 26.96157800988989
l2 norm of gradients: 0.16697159854676927
l2 norm of weights: 5.587006018893792
-----
Iteration Number: 9557
Loss: 26.96017958164478
l2 norm of gradients: 0.166959625061287
l2 norm of weights: 5.5869452443166985
-----
Iteration Number: 9558
Loss: 26.95878135330371
l2 norm of gradients: 0.16694765342388843
l2 norm of weights: 5.58688447392665
-----
Iteration Number: 9559
Loss: 26.95738332482357
l2 norm of gradients: 0.1669356836340508
l2 norm of weights: 5.586823707723133
-----
Iteration Number: 9560
Loss: 26.955985496156856
l2 norm of gradients: 0.16692371569125175
l2 norm of weights: 5.586762945705632
-----
Iteration Number: 9561
Loss: 26.954587867257956
l2 norm of gradients: 0.16691174959496885
l2 norm of weights: 5.586702187873631
-----
Iteration Number: 9562
Loss: 26.953190438084043
l2 norm of gradients: 0.16689978534468008
l2 norm of weights: 5.586641434226617
-----
Iteration Number: 9563
Loss: 26.951793208587198

```

```

12 norm of gradients: 0.16688782293986346
12 norm of weights: 5.5865806847640735
-----
Iteration Number: 9564
Loss: 26.950396178722343
12 norm of gradients: 0.16687586237999727
12 norm of weights: 5.586519939485487
-----
Iteration Number: 9565
Loss: 26.94899934844477
12 norm of gradients: 0.16686390366456008
12 norm of weights: 5.586459198390343
-----
Iteration Number: 9566
Loss: 26.9476027177095
12 norm of gradients: 0.16685194679303053
12 norm of weights: 5.586398461478126
-----
Iteration Number: 9567
Loss: 26.946206286471607
12 norm of gradients: 0.1668399917648875
12 norm of weights: 5.586337728748324
-----
Iteration Number: 9568
Loss: 26.944810054680488
12 norm of gradients: 0.16682803857961004
12 norm of weights: 5.58627700020042
-----
Iteration Number: 9569
Loss: 26.943414022298498
12 norm of gradients: 0.16681608723667748
12 norm of weights: 5.586216275833902
-----
Iteration Number: 9570
Loss: 26.942018189275306
12 norm of gradients: 0.16680413773556926
12 norm of weights: 5.586155555648255
-----
Iteration Number: 9571
Loss: 26.94062255556874
12 norm of gradients: 0.1667921900757651
12 norm of weights: 5.586094839642967
-----
Iteration Number: 9572
Loss: 26.93922712112992
12 norm of gradients: 0.16678024425674481
12 norm of weights: 5.586034127817522
-----

```

Iteration Number: 9573
Loss: 26.937831885917426
l2 norm of gradients: 0.16676830027798845
l2 norm of weights: 5.585973420171406

Iteration Number: 9574
Loss: 26.936436849882387
l2 norm of gradients: 0.16675635813897632
l2 norm of weights: 5.585912716704107

Iteration Number: 9575
Loss: 26.935042012983875
l2 norm of gradients: 0.16674441783918884
l2 norm of weights: 5.585852017415111

Iteration Number: 9576
Loss: 26.93364737517188
l2 norm of gradients: 0.16673247937810665
l2 norm of weights: 5.585791322303906

Iteration Number: 9577
Loss: 26.932252936404833
l2 norm of gradients: 0.1667205427552106
l2 norm of weights: 5.585730631369977

Iteration Number: 9578
Loss: 26.930858696634935
l2 norm of gradients: 0.16670860796998174
l2 norm of weights: 5.58566994461281

Iteration Number: 9579
Loss: 26.929464655819704
l2 norm of gradients: 0.16669667502190133
l2 norm of weights: 5.585609262031896

Iteration Number: 9580
Loss: 26.92807081391238
l2 norm of gradients: 0.1666847439104507
l2 norm of weights: 5.585548583626718

Iteration Number: 9581
Loss: 26.926677170866313
l2 norm of gradients: 0.16667281463511155
l2 norm of weights: 5.585487909396765

Iteration Number: 9582
Loss: 26.925283726639726
l2 norm of gradients: 0.1666608871953657

```
l2 norm of weights: 5.585427239341524
-----
Iteration Number: 9583
Loss: 26.923890481185154
l2 norm of gradients: 0.16664896159069514
l2 norm of weights: 5.585366573460482
-----
Iteration Number: 9584
Loss: 26.922497434460013
l2 norm of gradients: 0.16663703782058206
l2 norm of weights: 5.585305911753128
-----
Iteration Number: 9585
Loss: 26.921104586414852
l2 norm of gradients: 0.16662511588450887
l2 norm of weights: 5.585245254218949
-----
Iteration Number: 9586
Loss: 26.91971193700891
l2 norm of gradients: 0.16661319578195813
l2 norm of weights: 5.585184600857431
-----
Iteration Number: 9587
Loss: 26.918319486193074
l2 norm of gradients: 0.1666012775124127
l2 norm of weights: 5.585123951668064
-----
Iteration Number: 9588
Loss: 26.916927233927314
l2 norm of gradients: 0.1665893610753554
l2 norm of weights: 5.585063306650335
-----
Iteration Number: 9589
Loss: 26.915535180163907
l2 norm of gradients: 0.1665774464702695
l2 norm of weights: 5.585002665803732
-----
Iteration Number: 9590
Loss: 26.914143324856624
l2 norm of gradients: 0.1665655336966384
l2 norm of weights: 5.5849420291277445
-----
Iteration Number: 9591
Loss: 26.912751667961363
l2 norm of gradients: 0.16655362275394564
l2 norm of weights: 5.584881396621858
-----
Iteration Number: 9592
```

Loss: 26.91136020943496
l2 norm of gradients: 0.1665417136416749
l2 norm of weights: 5.584820768285564

Iteration Number: 9593
Loss: 26.909968949229093
l2 norm of gradients: 0.1665298063593101
l2 norm of weights: 5.5847601441183485

Iteration Number: 9594
Loss: 26.90857788730312
l2 norm of gradients: 0.16651790090633542
l2 norm of weights: 5.584699524119702

Iteration Number: 9595
Loss: 26.90718702360894
l2 norm of gradients: 0.1665059972822351
l2 norm of weights: 5.5846389082891115

Iteration Number: 9596
Loss: 26.905796358103366
l2 norm of gradients: 0.16649409548649383
l2 norm of weights: 5.584578296626066

Iteration Number: 9597
Loss: 26.904405890738737
l2 norm of gradients: 0.16648219551859614
l2 norm of weights: 5.584517689130056

Iteration Number: 9598
Loss: 26.903015621472655
l2 norm of gradients: 0.16647029737802693
l2 norm of weights: 5.58445708580057

Iteration Number: 9599
Loss: 26.901625550259958
l2 norm of gradients: 0.16645840106427134
l2 norm of weights: 5.584396486637095

Iteration Number: 9600
Loss: 26.900235677054614
l2 norm of gradients: 0.16644650657681465
l2 norm of weights: 5.584335891639122

Iteration Number: 9601
Loss: 26.898846001814544
l2 norm of gradients: 0.16643461391514225
l2 norm of weights: 5.584275300806141

```

-----
Iteration Number: 9602
Loss: 26.897456524491357
l2 norm of gradients: 0.1664227230787398
l2 norm of weights: 5.58421471413764
-----
Iteration Number: 9603
Loss: 26.896067245043728
l2 norm of gradients: 0.16641083406709317
l2 norm of weights: 5.584154131633109
-----
Iteration Number: 9604
Loss: 26.894678163423613
l2 norm of gradients: 0.1663989468796884
l2 norm of weights: 5.584093553292036
-----
Iteration Number: 9605
Loss: 26.893289279587954
l2 norm of gradients: 0.16638706151601168
l2 norm of weights: 5.584032979113915
-----
Iteration Number: 9606
Loss: 26.89190059349193
l2 norm of gradients: 0.1663751779755494
l2 norm of weights: 5.583972409098232
-----
Iteration Number: 9607
Loss: 26.890512105091716
l2 norm of gradients: 0.16636329625778823
l2 norm of weights: 5.583911843244479
-----
Iteration Number: 9608
Loss: 26.8891238143414
l2 norm of gradients: 0.16635141636221484
l2 norm of weights: 5.583851281552144
-----
Iteration Number: 9609
Loss: 26.88773572119492
l2 norm of gradients: 0.1663395382883163
l2 norm of weights: 5.583790724020719
-----
Iteration Number: 9610
Loss: 26.886347825609622
l2 norm of gradients: 0.16632766203557967
l2 norm of weights: 5.5837301706496945
-----
Iteration Number: 9611
Loss: 26.88496012754071

```


12 norm of gradients: 0.16631578760349244
12 norm of weights: 5.58366962143856

Iteration Number: 9612
Loss: 26.883572626941955
12 norm of gradients: 0.166303914991542
12 norm of weights: 5.5836090763868045

Iteration Number: 9613
Loss: 26.882185323769704
12 norm of gradients: 0.16629204419921614
12 norm of weights: 5.583548535493922

Iteration Number: 9614
Loss: 26.88079821798134
12 norm of gradients: 0.16628017522600283
12 norm of weights: 5.5834879987593995

Iteration Number: 9615
Loss: 26.87941130952977
12 norm of gradients: 0.16626830807138995
12 norm of weights: 5.58342746618273

Iteration Number: 9616
Loss: 26.87802459837083
12 norm of gradients: 0.16625644273486603
12 norm of weights: 5.5833669377634045

Iteration Number: 9617
Loss: 26.87663808446008
12 norm of gradients: 0.16624457921591942
12 norm of weights: 5.583306413500914

Iteration Number: 9618
Loss: 26.875251767751628
12 norm of gradients: 0.1662327175140388
12 norm of weights: 5.583245893394748

Iteration Number: 9619
Loss: 26.87386564820338
12 norm of gradients: 0.16622085762871303
12 norm of weights: 5.5831853774444

Iteration Number: 9620
Loss: 26.872479725770102
12 norm of gradients: 0.16620899955943108
12 norm of weights: 5.583124865649358

Iteration Number: 9621
Loss: 26.8710940004067
l2 norm of gradients: 0.16619714330568222
l2 norm of weights: 5.583064358009117

Iteration Number: 9622
Loss: 26.869708472067387
l2 norm of gradients: 0.16618528886695583
l2 norm of weights: 5.583003854523167

Iteration Number: 9623
Loss: 26.868323140710192
l2 norm of gradients: 0.16617343624274145
l2 norm of weights: 5.5829433551909995

Iteration Number: 9624
Loss: 26.86693800628875
l2 norm of gradients: 0.16616158543252893
l2 norm of weights: 5.5828828600121065

Iteration Number: 9625
Loss: 26.86555306875953
l2 norm of gradients: 0.1661497364358082
l2 norm of weights: 5.582822368985979

Iteration Number: 9626
Loss: 26.86416832807952
l2 norm of gradients: 0.16613788925206938
l2 norm of weights: 5.582761882112111

Iteration Number: 9627
Loss: 26.862783784200097
l2 norm of gradients: 0.16612604388080277
l2 norm of weights: 5.582701399389992

Iteration Number: 9628
Loss: 26.8613994370819
l2 norm of gradients: 0.1661142003214989
l2 norm of weights: 5.582640920819117

Iteration Number: 9629
Loss: 26.860015286675363
l2 norm of gradients: 0.16610235857364847
l2 norm of weights: 5.582580446398975

Iteration Number: 9630
Loss: 26.85863133294099
l2 norm of gradients: 0.16609051863674243

```

12 norm of weights: 5.582519976129062
-----
Iteration Number: 9631
Loss: 26.857247575830527
12 norm of gradients: 0.16607868051027166
12 norm of weights: 5.582459510008866
-----
Iteration Number: 9632
Loss: 26.855864015302657
12 norm of gradients: 0.16606684419372758
12 norm of weights: 5.582399048037884
-----
Iteration Number: 9633
Loss: 26.854480651309864
12 norm of gradients: 0.1660550096866015
12 norm of weights: 5.582338590215605
-----
Iteration Number: 9634
Loss: 26.853097483812487
12 norm of gradients: 0.16604317698838508
12 norm of weights: 5.582278136541525
-----
Iteration Number: 9635
Loss: 26.851714512760033
12 norm of gradients: 0.16603134609857018
12 norm of weights: 5.5822176870151345
-----
Iteration Number: 9636
Loss: 26.850331738112995
12 norm of gradients: 0.1660195170166486
12 norm of weights: 5.582157241635928
-----
Iteration Number: 9637
Loss: 26.848949159826393
12 norm of gradients: 0.16600768974211264
12 norm of weights: 5.582096800403398
-----
Iteration Number: 9638
Loss: 26.84756677785417
12 norm of gradients: 0.16599586427445462
12 norm of weights: 5.582036363317037
-----
Iteration Number: 9639
Loss: 26.846184592152337
12 norm of gradients: 0.16598404061316702
12 norm of weights: 5.581975930376339
-----
Iteration Number: 9640

```

Loss: 26.844802602677966
l2 norm of gradients: 0.16597221875774257
l2 norm of weights: 5.581915501580799

Iteration Number: 9641
Loss: 26.843420809386934
l2 norm of gradients: 0.16596039870767412
l2 norm of weights: 5.581855076929908

Iteration Number: 9642
Loss: 26.842039212233736
l2 norm of gradients: 0.1659485804624548
l2 norm of weights: 5.581794656423159

Iteration Number: 9643
Loss: 26.840657811173408
l2 norm of gradients: 0.16593676402157786
l2 norm of weights: 5.581734240060049

Iteration Number: 9644
Loss: 26.83927660616431
l2 norm of gradients: 0.16592494938453667
l2 norm of weights: 5.58167382784007

Iteration Number: 9645
Loss: 26.837895597160912
l2 norm of gradients: 0.16591313655082485
l2 norm of weights: 5.581613419762715

Iteration Number: 9646
Loss: 26.836514784119814
l2 norm of gradients: 0.16590132551993622
l2 norm of weights: 5.581553015827479

Iteration Number: 9647
Loss: 26.835134166996994
l2 norm of gradients: 0.16588951629136472
l2 norm of weights: 5.581492616033857

Iteration Number: 9648
Loss: 26.8337537457459
l2 norm of gradients: 0.16587770886460454
l2 norm of weights: 5.581432220381342

Iteration Number: 9649
Loss: 26.83237352032535
l2 norm of gradients: 0.16586590323915
l2 norm of weights: 5.581371828869429

```
-----  
Iteration Number: 9650  
Loss: 26.830993490687128  
l2 norm of gradients: 0.16585409941449558  
l2 norm of weights: 5.581311441497612  
-----  
Iteration Number: 9651  
Loss: 26.829613656793413  
l2 norm of gradients: 0.16584229739013603  
l2 norm of weights: 5.581251058265385  
-----  
Iteration Number: 9652  
Loss: 26.828234018596078  
l2 norm of gradients: 0.16583049716556617  
l2 norm of weights: 5.581190679172244  
-----  
Iteration Number: 9653  
Loss: 26.826854576051005  
l2 norm of gradients: 0.16581869874028105  
l2 norm of weights: 5.581130304217684  
-----  
Iteration Number: 9654  
Loss: 26.825475329115058  
l2 norm of gradients: 0.16580690211377594  
l2 norm of weights: 5.5810699334011975  
-----  
Iteration Number: 9655  
Loss: 26.82409627774536  
l2 norm of gradients: 0.1657951072855462  
l2 norm of weights: 5.581009566722282  
-----  
Iteration Number: 9656  
Loss: 26.822717421895508  
l2 norm of gradients: 0.16578331425508747  
l2 norm of weights: 5.58094920418043  
-----  
Iteration Number: 9657  
Loss: 26.821338761521783  
l2 norm of gradients: 0.16577152302189543  
l2 norm of weights: 5.58088884577514  
-----  
Iteration Number: 9658  
Loss: 26.819960296583286  
l2 norm of gradients: 0.16575973358546614  
l2 norm of weights: 5.580828491505906  
-----  
Iteration Number: 9659  
Loss: 26.81858202703247
```

```

12 norm of gradients: 0.16574794594529563
12 norm of weights: 5.580768141372221
-----
Iteration Number: 9660
Loss: 26.81720395282628
12 norm of gradients: 0.16573616010088021
12 norm of weights: 5.580707795373583
-----
Iteration Number: 9661
Loss: 26.815826073923088
12 norm of gradients: 0.16572437605171644
12 norm of weights: 5.580647453509488
-----
Iteration Number: 9662
Loss: 26.81444839027574
12 norm of gradients: 0.16571259379730088
12 norm of weights: 5.580587115779429
-----
Iteration Number: 9663
Loss: 26.813070901842565
12 norm of gradients: 0.16570081333713038
12 norm of weights: 5.580526782182904
-----
Iteration Number: 9664
Loss: 26.81169360857859
12 norm of gradients: 0.165689034670702
12 norm of weights: 5.580466452719409
-----
Iteration Number: 9665
Loss: 26.81031651044042
12 norm of gradients: 0.16567725779751286
12 norm of weights: 5.580406127388438
-----
Iteration Number: 9666
Loss: 26.8089396073851
12 norm of gradients: 0.16566548271706036
12 norm of weights: 5.58034580618949
-----
Iteration Number: 9667
Loss: 26.807562899365195
12 norm of gradients: 0.165653709428842
12 norm of weights: 5.580285489122058
-----
Iteration Number: 9668
Loss: 26.806186386340983
12 norm of gradients: 0.1656419379323556
12 norm of weights: 5.580225176185642
-----

```

Iteration Number: 9669
Loss: 26.80481006826732
l2 norm of gradients: 0.16563016822709895
l2 norm of weights: 5.580164867379735

Iteration Number: 9670
Loss: 26.80343394510008
l2 norm of gradients: 0.16561840031257014
l2 norm of weights: 5.580104562703835

Iteration Number: 9671
Loss: 26.802058016795755
l2 norm of gradients: 0.1656066341882674
l2 norm of weights: 5.580044262157439

Iteration Number: 9672
Loss: 26.800682283309325
l2 norm of gradients: 0.16559486985368918
l2 norm of weights: 5.579983965740043

Iteration Number: 9673
Loss: 26.799306744599228
l2 norm of gradients: 0.1655831073083341
l2 norm of weights: 5.579923673451144

Iteration Number: 9674
Loss: 26.79793140062054
l2 norm of gradients: 0.16557134655170083
l2 norm of weights: 5.579863385290239

Iteration Number: 9675
Loss: 26.79655625133058
l2 norm of gradients: 0.16555958758328845
l2 norm of weights: 5.579803101256825

Iteration Number: 9676
Loss: 26.795181296682895
l2 norm of gradients: 0.16554783040259602
l2 norm of weights: 5.5797428213504

Iteration Number: 9677
Loss: 26.793806536636037
l2 norm of gradients: 0.1655360750091228
l2 norm of weights: 5.579682545570459

Iteration Number: 9678
Loss: 26.792431971145884
l2 norm of gradients: 0.16552432140236828

```
l2 norm of weights: 5.579622273916502
-----
Iteration Number: 9679
Loss: 26.791057600169257
l2 norm of gradients: 0.16551256958183208
l2 norm of weights: 5.5795620063880245
-----
Iteration Number: 9680
Loss: 26.78968342366266
l2 norm of gradients: 0.16550081954701404
l2 norm of weights: 5.579501742984526
-----
Iteration Number: 9681
Loss: 26.78830944157998
l2 norm of gradients: 0.16548907129741416
l2 norm of weights: 5.5794414837055015
-----
Iteration Number: 9682
Loss: 26.78693565388123
l2 norm of gradients: 0.16547732483253264
l2 norm of weights: 5.579381228550451
-----
Iteration Number: 9683
Loss: 26.785562060519503
l2 norm of gradients: 0.1654655801518697
l2 norm of weights: 5.579320977518871
-----
Iteration Number: 9684
Loss: 26.784188661455982
l2 norm of gradients: 0.165453837254926
l2 norm of weights: 5.57926073061026
-----
Iteration Number: 9685
Loss: 26.78281545664012
l2 norm of gradients: 0.16544209614120214
l2 norm of weights: 5.579200487824117
-----
Iteration Number: 9686
Loss: 26.78144244603344
l2 norm of gradients: 0.165430356810199
l2 norm of weights: 5.579140249159939
-----
Iteration Number: 9687
Loss: 26.780069629591086
l2 norm of gradients: 0.16541861926141752
l2 norm of weights: 5.579080014617224
-----
Iteration Number: 9688
```


Loss: 26.778697007270573
l2 norm of gradients: 0.16540688349435903
l2 norm of weights: 5.579019784195471

Iteration Number: 9689
Loss: 26.777324579026402
l2 norm of gradients: 0.16539514950852488
l2 norm of weights: 5.578959557894179

Iteration Number: 9690
Loss: 26.775952344815863
l2 norm of gradients: 0.16538341730341655
l2 norm of weights: 5.578899335712846

Iteration Number: 9691
Loss: 26.774580304596125
l2 norm of gradients: 0.16537168687853585
l2 norm of weights: 5.578839117650969

Iteration Number: 9692
Loss: 26.773208458322472
l2 norm of gradients: 0.16535995823338454
l2 norm of weights: 5.57877890370805

Iteration Number: 9693
Loss: 26.77183680595251
l2 norm of gradients: 0.16534823136746488
l2 norm of weights: 5.578718693883586

Iteration Number: 9694
Loss: 26.77046534744183
l2 norm of gradients: 0.1653365062802789
l2 norm of weights: 5.578658488177077

Iteration Number: 9695
Loss: 26.769094082749156
l2 norm of gradients: 0.16532478297132908
l2 norm of weights: 5.57859828658802

Iteration Number: 9696
Loss: 26.76772301182763
l2 norm of gradients: 0.16531306144011806
l2 norm of weights: 5.578538089115917

Iteration Number: 9697
Loss: 26.76635213463793
l2 norm of gradients: 0.16530134168614852
l2 norm of weights: 5.578477895760266

```

-----
Iteration Number: 9698
Loss: 26.764981451133213
l2 norm of gradients: 0.16528962370892336
l2 norm of weights: 5.578417706520566
-----
Iteration Number: 9699
Loss: 26.763610961271503
l2 norm of gradients: 0.16527790750794577
l2 norm of weights: 5.578357521396316
-----
Iteration Number: 9700
Loss: 26.76224066501065
l2 norm of gradients: 0.16526619308271887
l2 norm of weights: 5.578297340387017
-----
Iteration Number: 9701
Loss: 26.760870562303943
l2 norm of gradients: 0.1652544804327462
l2 norm of weights: 5.578237163492169
-----
Iteration Number: 9702
Loss: 26.759500653111893
l2 norm of gradients: 0.1652427695575313
l2 norm of weights: 5.5781769907112695
-----
Iteration Number: 9703
Loss: 26.758130937388707
l2 norm of gradients: 0.16523106045657793
l2 norm of weights: 5.578116822043821
-----
Iteration Number: 9704
Loss: 26.756761415092214
l2 norm of gradients: 0.1652193531293901
l2 norm of weights: 5.578056657489322
-----
Iteration Number: 9705
Loss: 26.755392086178162
l2 norm of gradients: 0.1652076475754718
l2 norm of weights: 5.577996497047273
-----
Iteration Number: 9706
Loss: 26.754022950604423
l2 norm of gradients: 0.1651959437943274
l2 norm of weights: 5.577936340717175
-----
Iteration Number: 9707
Loss: 26.752654008327287

```

```

12 norm of gradients: 0.16518424178546132
12 norm of weights: 5.577876188498526
-----
Iteration Number: 9708
Loss: 26.75128525930335
12 norm of gradients: 0.16517254154837815
12 norm of weights: 5.577816040390829
-----
Iteration Number: 9709
Loss: 26.74991670348943
12 norm of gradients: 0.1651608430825827
12 norm of weights: 5.577755896393583
-----
Iteration Number: 9710
Loss: 26.748548340843165
12 norm of gradients: 0.1651491463875799
12 norm of weights: 5.577695756506291
-----
Iteration Number: 9711
Loss: 26.74718017131905
12 norm of gradients: 0.1651374514628749
12 norm of weights: 5.57763562072845
-----
Iteration Number: 9712
Loss: 26.74581219487518
12 norm of gradients: 0.16512575830797296
12 norm of weights: 5.577575489059563
-----
Iteration Number: 9713
Loss: 26.74444441147028
12 norm of gradients: 0.16511406692237957
12 norm of weights: 5.577515361499131
-----
Iteration Number: 9714
Loss: 26.743076821060004
12 norm of gradients: 0.1651023773056003
12 norm of weights: 5.577455238046655
-----
Iteration Number: 9715
Loss: 26.741709423599463
12 norm of gradients: 0.16509068945714095
12 norm of weights: 5.577395118701635
-----
Iteration Number: 9716
Loss: 26.74034221904627
12 norm of gradients: 0.16507900337650752
12 norm of weights: 5.577335003463574
-----

```

Iteration Number: 9717
Loss: 26.73897520735916
l2 norm of gradients: 0.16506731906320607
l2 norm of weights: 5.577274892331973

Iteration Number: 9718
Loss: 26.737608388493577
l2 norm of gradients: 0.16505563651674296
l2 norm of weights: 5.577214785306332

Iteration Number: 9719
Loss: 26.7362417624064
l2 norm of gradients: 0.1650439557366246
l2 norm of weights: 5.577154682386153

Iteration Number: 9720
Loss: 26.73487532905622
l2 norm of gradients: 0.16503227672235762
l2 norm of weights: 5.577094583570939

Iteration Number: 9721
Loss: 26.733509088397295
l2 norm of gradients: 0.16502059947344885
l2 norm of weights: 5.57703448886019

Iteration Number: 9722
Loss: 26.7321430403888
l2 norm of gradients: 0.16500892398940523
l2 norm of weights: 5.576974398253409

Iteration Number: 9723
Loss: 26.730777184985087
l2 norm of gradients: 0.16499725026973383
l2 norm of weights: 5.576914311750098

Iteration Number: 9724
Loss: 26.729411522146115
l2 norm of gradients: 0.164985578313942
l2 norm of weights: 5.576854229349759

Iteration Number: 9725
Loss: 26.728046051827683
l2 norm of gradients: 0.1649739081215372
l2 norm of weights: 5.576794151051893

Iteration Number: 9726
Loss: 26.726680773987304
l2 norm of gradients: 0.164962239692027

```
12 norm of weights: 5.576734076856003
-----
Iteration Number: 9727
Loss: 26.725315688580558
12 norm of gradients: 0.16495057302491928
12 norm of weights: 5.576674006761594
-----
Iteration Number: 9728
Loss: 26.72395079556582
12 norm of gradients: 0.16493890811972187
12 norm of weights: 5.576613940768164
-----
Iteration Number: 9729
Loss: 26.722586094900215
12 norm of gradients: 0.16492724497594294
12 norm of weights: 5.576553878875218
-----
Iteration Number: 9730
Loss: 26.721221586539546
12 norm of gradients: 0.16491558359309078
12 norm of weights: 5.576493821082257
-----
Iteration Number: 9731
Loss: 26.71985727044106
12 norm of gradients: 0.16490392397067383
12 norm of weights: 5.576433767388786
-----
Iteration Number: 9732
Loss: 26.718493146564306
12 norm of gradients: 0.16489226610820068
12 norm of weights: 5.576373717794308
-----
Iteration Number: 9733
Loss: 26.717129214864336
12 norm of gradients: 0.16488061000518017
12 norm of weights: 5.5763136722983235
-----
Iteration Number: 9734
Loss: 26.715765475297562
12 norm of gradients: 0.16486895566112114
12 norm of weights: 5.576253630900337
-----
Iteration Number: 9735
Loss: 26.714401927822877
12 norm of gradients: 0.16485730307553276
12 norm of weights: 5.576193593599851
-----
Iteration Number: 9736
```

Loss: 26.713038572397345
l2 norm of gradients: 0.1648456522479243
l2 norm of weights: 5.576133560396371

Iteration Number: 9737
Loss: 26.71167540897686
l2 norm of gradients: 0.16483400317780517
l2 norm of weights: 5.576073531289398

Iteration Number: 9738
Loss: 26.71031243752138
l2 norm of gradients: 0.16482235586468497
l2 norm of weights: 5.576013506278437

Iteration Number: 9739
Loss: 26.708949657983826
l2 norm of gradients: 0.16481071030807343
l2 norm of weights: 5.57595348536299

Iteration Number: 9740
Loss: 26.707587070325836
l2 norm of gradients: 0.1647990665074805
l2 norm of weights: 5.575893468542561

Iteration Number: 9741
Loss: 26.706224674500266
l2 norm of gradients: 0.1647874244624163
l2 norm of weights: 5.575833455816655

Iteration Number: 9742
Loss: 26.70486247046773
l2 norm of gradients: 0.16477578417239097
l2 norm of weights: 5.575773447184775

Iteration Number: 9743
Loss: 26.70350045818458
l2 norm of gradients: 0.16476414563691497
l2 norm of weights: 5.575713442646426

Iteration Number: 9744
Loss: 26.702138637608876
l2 norm of gradients: 0.16475250885549894
l2 norm of weights: 5.575653442201111

Iteration Number: 9745
Loss: 26.700777008694047
l2 norm of gradients: 0.16474087382765348
l2 norm of weights: 5.575593445848335

```

-----
Iteration Number: 9746
Loss: 26.69941557140274
l2 norm of gradients: 0.1647292405528896
l2 norm of weights: 5.5755334535876
-----
Iteration Number: 9747
Loss: 26.698054325689128
l2 norm of gradients: 0.1647176090307183
l2 norm of weights: 5.575473465418413
-----
Iteration Number: 9748
Loss: 26.69669327151077
l2 norm of gradients: 0.1647059792606508
l2 norm of weights: 5.575413481340278
-----
Iteration Number: 9749
Loss: 26.69533240882646
l2 norm of gradients: 0.16469435124219853
l2 norm of weights: 5.575353501352699
-----
Iteration Number: 9750
Loss: 26.693971737593383
l2 norm of gradients: 0.1646827249748729
l2 norm of weights: 5.575293525455181
-----
Iteration Number: 9751
Loss: 26.692611257766494
l2 norm of gradients: 0.16467110045818575
l2 norm of weights: 5.575233553647228
-----
Iteration Number: 9752
Loss: 26.691250969304967
l2 norm of gradients: 0.16465947769164893
l2 norm of weights: 5.575173585928347
-----
Iteration Number: 9753
Loss: 26.689890872167528
l2 norm of gradients: 0.16464785667477436
l2 norm of weights: 5.5751136222980415
-----
Iteration Number: 9754
Loss: 26.688530966309685
l2 norm of gradients: 0.1646362374070743
l2 norm of weights: 5.575053662755817
-----
Iteration Number: 9755
Loss: 26.687171251689197

```

```
12 norm of gradients: 0.1646246198880611
12 norm of weights: 5.574993707301178
-----
Iteration Number: 9756
Loss: 26.685811728264245
12 norm of gradients: 0.16461300411724727
12 norm of weights: 5.574933755933629
-----
Iteration Number: 9757
Loss: 26.68445239599241
12 norm of gradients: 0.16460139009414546
12 norm of weights: 5.574873808652678
-----
Iteration Number: 9758
Loss: 26.68309325482932
12 norm of gradients: 0.16458977781826845
12 norm of weights: 5.574813865457829
-----
Iteration Number: 9759
Loss: 26.681734304733393
12 norm of gradients: 0.16457816728912933
12 norm of weights: 5.5747539263485875
-----
Iteration Number: 9760
Loss: 26.680375545664276
12 norm of gradients: 0.16456655850624108
12 norm of weights: 5.574693991324461
-----
Iteration Number: 9761
Loss: 26.67901697757591
12 norm of gradients: 0.16455495146911717
12 norm of weights: 5.574634060384952
-----
Iteration Number: 9762
Loss: 26.677658600428884
12 norm of gradients: 0.164543346177271
12 norm of weights: 5.574574133529568
-----
Iteration Number: 9763
Loss: 26.67630041417878
12 norm of gradients: 0.16453174263021614
12 norm of weights: 5.574514210757816
-----
Iteration Number: 9764
Loss: 26.674942418783743
12 norm of gradients: 0.16452014082746647
12 norm of weights: 5.574454292069202
-----
```


Iteration Number: 9765
Loss: 26.673584614203747
l2 norm of gradients: 0.16450854076853585
l2 norm of weights: 5.57439437746323

Iteration Number: 9766
Loss: 26.672227000391022
l2 norm of gradients: 0.16449694245293844
l2 norm of weights: 5.574334466939409

Iteration Number: 9767
Loss: 26.670869577307226
l2 norm of gradients: 0.16448534588018845
l2 norm of weights: 5.574274560497244

Iteration Number: 9768
Loss: 26.669512344910466
l2 norm of gradients: 0.1644737510498003
l2 norm of weights: 5.574214658136242

Iteration Number: 9769
Loss: 26.668155303155448
l2 norm of gradients: 0.16446215796128855
l2 norm of weights: 5.574154759855909

Iteration Number: 9770
Loss: 26.666798452002453
l2 norm of gradients: 0.164450566614168
l2 norm of weights: 5.574094865655752

Iteration Number: 9771
Loss: 26.665441791408227
l2 norm of gradients: 0.16443897700795349
l2 norm of weights: 5.5740349755352785

Iteration Number: 9772
Loss: 26.664085321329903
l2 norm of gradients: 0.16442738914216004
l2 norm of weights: 5.573975089493994

Iteration Number: 9773
Loss: 26.662729041725505
l2 norm of gradients: 0.16441580301630287
l2 norm of weights: 5.573915207531407

Iteration Number: 9774
Loss: 26.661372952552355
l2 norm of gradients: 0.1644042186298974

```
l2 norm of weights: 5.5738553296470235
-----
Iteration Number: 9775
Loss: 26.660017053770282
l2 norm of gradients: 0.16439263598245912
l2 norm of weights: 5.573795455840352
-----
Iteration Number: 9776
Loss: 26.658661345333435
l2 norm of gradients: 0.16438105507350365
l2 norm of weights: 5.573735586110898
-----
Iteration Number: 9777
Loss: 26.657305827202812
l2 norm of gradients: 0.16436947590254686
l2 norm of weights: 5.57367572045817
-----
Iteration Number: 9778
Loss: 26.65595049933481
l2 norm of gradients: 0.16435789846910476
l2 norm of weights: 5.573615858881675
-----
Iteration Number: 9779
Loss: 26.654595361686773
l2 norm of gradients: 0.16434632277269345
l2 norm of weights: 5.573556001380922
-----
Iteration Number: 9780
Loss: 26.65324041421719
l2 norm of gradients: 0.16433474881282925
l2 norm of weights: 5.573496147955418
-----
Iteration Number: 9781
Loss: 26.651885656884357
l2 norm of gradients: 0.16432317658902867
l2 norm of weights: 5.57343629860467
-----
Iteration Number: 9782
Loss: 26.650531089645302
l2 norm of gradients: 0.16431160610080825
l2 norm of weights: 5.573376453328186
-----
Iteration Number: 9783
Loss: 26.649176712457763
l2 norm of gradients: 0.1643000373476848
l2 norm of weights: 5.573316612125475
-----
Iteration Number: 9784
```

Loss: 26.647822525280585
l2 norm of gradients: 0.1642884703291752
l2 norm of weights: 5.573256774996045

Iteration Number: 9785
Loss: 26.64646852806871
l2 norm of gradients: 0.1642769050447966
l2 norm of weights: 5.573196941939403

Iteration Number: 9786
Loss: 26.645114720785426
l2 norm of gradients: 0.16426534149406616
l2 norm of weights: 5.573137112955059

Iteration Number: 9787
Loss: 26.643761103383643
l2 norm of gradients: 0.16425377967650132
l2 norm of weights: 5.57307728804252

Iteration Number: 9788
Loss: 26.6424076758226
l2 norm of gradients: 0.1642422195916196
l2 norm of weights: 5.573017467201295

Iteration Number: 9789
Loss: 26.641054438061072
l2 norm of gradients: 0.1642306612389387
l2 norm of weights: 5.572957650430893

Iteration Number: 9790
Loss: 26.639701390056917
l2 norm of gradients: 0.16421910461797656
l2 norm of weights: 5.572897837730823

Iteration Number: 9791
Loss: 26.638348531767463
l2 norm of gradients: 0.16420754972825102
l2 norm of weights: 5.572838029100592

Iteration Number: 9792
Loss: 26.63699586315292
l2 norm of gradients: 0.16419599656928038
l2 norm of weights: 5.572778224539711

Iteration Number: 9793
Loss: 26.635643384166166
l2 norm of gradients: 0.16418444514058297
l2 norm of weights: 5.572718424047688

```

-----
Iteration Number: 9794
Loss: 26.634291094770248
l2 norm of gradients: 0.16417289544167712
l2 norm of weights: 5.5726586276240315
-----
Iteration Number: 9795
Loss: 26.63293899492277
l2 norm of gradients: 0.1641613474720816
l2 norm of weights: 5.572598835268253
-----
Iteration Number: 9796
Loss: 26.631587084577845
l2 norm of gradients: 0.16414980123131515
l2 norm of weights: 5.572539046979859
-----
Iteration Number: 9797
Loss: 26.6302353636965
l2 norm of gradients: 0.16413825671889665
l2 norm of weights: 5.572479262758361
-----
Iteration Number: 9798
Loss: 26.62888383223755
l2 norm of gradients: 0.16412671393434522
l2 norm of weights: 5.572419482603268
-----
Iteration Number: 9799
Loss: 26.627532490157307
l2 norm of gradients: 0.16411517287718014
l2 norm of weights: 5.572359706514089
-----
Iteration Number: 9800
Loss: 26.626181337413158
l2 norm of gradients: 0.16410363354692076
l2 norm of weights: 5.572299934490334
-----
Iteration Number: 9801
Loss: 26.62483037396595
l2 norm of gradients: 0.16409209594308657
l2 norm of weights: 5.572240166531513
-----
Iteration Number: 9802
Loss: 26.623479599771148
l2 norm of gradients: 0.16408056006519736
l2 norm of weights: 5.572180402637136
-----
Iteration Number: 9803
Loss: 26.62212901478844

```

```
l2 norm of gradients: 0.16406902591277298
l2 norm of weights: 5.572120642806713
-----
Iteration Number: 9804
Loss: 26.620778618974676
l2 norm of gradients: 0.16405749348533336
l2 norm of weights: 5.572060887039753
-----
Iteration Number: 9805
Loss: 26.619428412289015
l2 norm of gradients: 0.16404596278239872
l2 norm of weights: 5.572001135335768
-----
Iteration Number: 9806
Loss: 26.618078394691008
l2 norm of gradients: 0.16403443380348934
l2 norm of weights: 5.571941387694267
-----
Iteration Number: 9807
Loss: 26.616728566135627
l2 norm of gradients: 0.16402290654812568
l2 norm of weights: 5.571881644114761
-----
Iteration Number: 9808
Loss: 26.61537892658249
l2 norm of gradients: 0.16401138101582835
l2 norm of weights: 5.57182190459676
-----
Iteration Number: 9809
Loss: 26.6140294759901
l2 norm of gradients: 0.16399985720611812
l2 norm of weights: 5.571762169139776
-----
Iteration Number: 9810
Loss: 26.612680214317066
l2 norm of gradients: 0.16398833511851588
l2 norm of weights: 5.571702437743317
-----
Iteration Number: 9811
Loss: 26.611331141522726
l2 norm of gradients: 0.16397681475254272
l2 norm of weights: 5.5716427104068975
-----
Iteration Number: 9812
Loss: 26.60998225756108
l2 norm of gradients: 0.16396529610771987
l2 norm of weights: 5.571582987130025
-----
```

Iteration Number: 9813
Loss: 26.608633562394697
l2 norm of gradients: 0.16395377918356868
l2 norm of weights: 5.571523267912212

Iteration Number: 9814
Loss: 26.60728505597845
l2 norm of gradients: 0.16394226397961062
l2 norm of weights: 5.57146355275297

Iteration Number: 9815
Loss: 26.605936738274234
l2 norm of gradients: 0.16393075049536748
l2 norm of weights: 5.571403841651811

Iteration Number: 9816
Loss: 26.604588609237307
l2 norm of gradients: 0.16391923873036096
l2 norm of weights: 5.5713441346082435

Iteration Number: 9817
Loss: 26.60324066882687
l2 norm of gradients: 0.16390772868411307
l2 norm of weights: 5.57128443162178

Iteration Number: 9818
Loss: 26.60189291700293
l2 norm of gradients: 0.16389622035614593
l2 norm of weights: 5.571224732691934

Iteration Number: 9819
Loss: 26.6005453537224
l2 norm of gradients: 0.16388471374598182
l2 norm of weights: 5.571165037818214

Iteration Number: 9820
Loss: 26.59919797894385
l2 norm of gradients: 0.1638732088531432
l2 norm of weights: 5.571105347000135

Iteration Number: 9821
Loss: 26.597850792622882
l2 norm of gradients: 0.1638617056771525
l2 norm of weights: 5.571045660237205

Iteration Number: 9822
Loss: 26.596503794722388
l2 norm of gradients: 0.16385020421753257

```

12 norm of weights: 5.570985977528939
-----
Iteration Number: 9823
Loss: 26.59515698519963
12 norm of gradients: 0.16383870447380625
12 norm of weights: 5.5709262988748485
-----
Iteration Number: 9824
Loss: 26.593810364011226
12 norm of gradients: 0.16382720644549653
12 norm of weights: 5.570866624274445
-----
Iteration Number: 9825
Loss: 26.592463931117557
12 norm of gradients: 0.1638157101321266
12 norm of weights: 5.57080695372724
-----
Iteration Number: 9826
Loss: 26.59111768647566
12 norm of gradients: 0.16380421553321978
12 norm of weights: 5.570747287232747
-----
Iteration Number: 9827
Loss: 26.589771630044403
12 norm of gradients: 0.1637927226482995
12 norm of weights: 5.5706876247904775
-----
Iteration Number: 9828
Loss: 26.588425761781927
12 norm of gradients: 0.1637812314768894
12 norm of weights: 5.570627966399945
-----
Iteration Number: 9829
Loss: 26.587080081648104
12 norm of gradients: 0.16376974201851324
12 norm of weights: 5.570568312060661
-----
Iteration Number: 9830
Loss: 26.58573458959954
12 norm of gradients: 0.16375825427269494
12 norm of weights: 5.5705086617721395
-----
Iteration Number: 9831
Loss: 26.584389285596753
12 norm of gradients: 0.16374676823895853
12 norm of weights: 5.5704490155338915
-----
Iteration Number: 9832

```

Loss: 26.58304416959506
l2 norm of gradients: 0.16373528391682826
l2 norm of weights: 5.570389373345432

Iteration Number: 9833
Loss: 26.581699241557487
l2 norm of gradients: 0.16372380130582842
l2 norm of weights: 5.570329735206272

Iteration Number: 9834
Loss: 26.580354501440144
l2 norm of gradients: 0.16371232040548356
l2 norm of weights: 5.570270101115925

Iteration Number: 9835
Loss: 26.57900994919922
l2 norm of gradients: 0.16370084121531833
l2 norm of weights: 5.570210471073906

Iteration Number: 9836
Loss: 26.577665584798897
l2 norm of gradients: 0.16368936373485748
l2 norm of weights: 5.570150845079726

Iteration Number: 9837
Loss: 26.57632140819364
l2 norm of gradients: 0.163677887963626
l2 norm of weights: 5.5700912231329

Iteration Number: 9838
Loss: 26.574977419342016
l2 norm of gradients: 0.163666413901149
l2 norm of weights: 5.57003160523294

Iteration Number: 9839
Loss: 26.57363361820533
l2 norm of gradients: 0.16365494154695168
l2 norm of weights: 5.5699719913793615

Iteration Number: 9840
Loss: 26.572290004738644
l2 norm of gradients: 0.16364347090055942
l2 norm of weights: 5.569912381571677

Iteration Number: 9841
Loss: 26.570946578904508
l2 norm of gradients: 0.16363200196149774
l2 norm of weights: 5.569852775809399


```
-----  
Iteration Number: 9842  
Loss: 26.56960334065862  
l2 norm of gradients: 0.16362053472929244  
l2 norm of weights: 5.569793174092044  
-----  
Iteration Number: 9843  
Loss: 26.56826028996091  
l2 norm of gradients: 0.16360906920346913  
l2 norm of weights: 5.569733576419124  
-----  
Iteration Number: 9844  
Loss: 26.566917426769006  
l2 norm of gradients: 0.16359760538355395  
l2 norm of weights: 5.569673982790154  
-----  
Iteration Number: 9845  
Loss: 26.565574751044107  
l2 norm of gradients: 0.163586143269073  
l2 norm of weights: 5.569614393204648  
-----  
Iteration Number: 9846  
Loss: 26.56423226274266  
l2 norm of gradients: 0.16357468285955248  
l2 norm of weights: 5.5695548076621195  
-----  
Iteration Number: 9847  
Loss: 26.56288996182316  
l2 norm of gradients: 0.16356322415451882  
l2 norm of weights: 5.569495226162084  
-----  
Iteration Number: 9848  
Loss: 26.56154784824705  
l2 norm of gradients: 0.16355176715349865  
l2 norm of weights: 5.569435648704056  
-----  
Iteration Number: 9849  
Loss: 26.560205921969235  
l2 norm of gradients: 0.1635403118560185  
l2 norm of weights: 5.569376075287549  
-----  
Iteration Number: 9850  
Loss: 26.558864182951613  
l2 norm of gradients: 0.16352885826160538  
l2 norm of weights: 5.569316505912078  
-----  
Iteration Number: 9851  
Loss: 26.557522631151613
```

12 norm of gradients: 0.1635174063697862
12 norm of weights: 5.569256940577159

Iteration Number: 9852
Loss: 26.55618126652888
12 norm of gradients: 0.16350595618008815
12 norm of weights: 5.569197379282305

Iteration Number: 9853
Loss: 26.55484008904075
12 norm of gradients: 0.16349450769203844
12 norm of weights: 5.569137822027032

Iteration Number: 9854
Loss: 26.55349909864757
12 norm of gradients: 0.16348306090516457
12 norm of weights: 5.569078268810855

Iteration Number: 9855
Loss: 26.5521582953065
12 norm of gradients: 0.16347161581899403
12 norm of weights: 5.569018719633289

Iteration Number: 9856
Loss: 26.55081767897949
12 norm of gradients: 0.16346017243305463
12 norm of weights: 5.568959174493849

Iteration Number: 9857
Loss: 26.549477249621642
12 norm of gradients: 0.16344873074687416
12 norm of weights: 5.568899633392051

Iteration Number: 9858
Loss: 26.548137007194974
12 norm of gradients: 0.1634372907599806
12 norm of weights: 5.56884009632741

Iteration Number: 9859
Loss: 26.546796951654688
12 norm of gradients: 0.16342585247190217
12 norm of weights: 5.568780563299441

Iteration Number: 9860
Loss: 26.54545708296388
12 norm of gradients: 0.16341441588216712
12 norm of weights: 5.568721034307661

Iteration Number: 9861
Loss: 26.544117401078605
l2 norm of gradients: 0.16340298099030393
l2 norm of weights: 5.5686615093515845

Iteration Number: 9862
Loss: 26.542777905960076
l2 norm of gradients: 0.1633915477958411
l2 norm of weights: 5.5686019884307285

Iteration Number: 9863
Loss: 26.541438597566188
l2 norm of gradients: 0.16338011629830745
l2 norm of weights: 5.568542471544608

Iteration Number: 9864
Loss: 26.540099475853722
l2 norm of gradients: 0.16336868649723174
l2 norm of weights: 5.568482958692739

Iteration Number: 9865
Loss: 26.538760540785503
l2 norm of gradients: 0.16335725839214302
l2 norm of weights: 5.568423449874637

Iteration Number: 9866
Loss: 26.53742179231734
l2 norm of gradients: 0.1633458319825705
l2 norm of weights: 5.56836394508982

Iteration Number: 9867
Loss: 26.53608323040946
l2 norm of gradients: 0.16333440726804338
l2 norm of weights: 5.568304444337804

Iteration Number: 9868
Loss: 26.53474485502169
l2 norm of gradients: 0.1633229842480912
l2 norm of weights: 5.568244947618104

Iteration Number: 9869
Loss: 26.53340666611227
l2 norm of gradients: 0.16331156292224344
l2 norm of weights: 5.568185454930239

Iteration Number: 9870
Loss: 26.53206866364095
l2 norm of gradients: 0.1633001432900299

```

12 norm of weights: 5.568125966273723
-----
Iteration Number: 9871
Loss: 26.530730847566645
12 norm of gradients: 0.16328872535098035
12 norm of weights: 5.5680664816480725
-----
Iteration Number: 9872
Loss: 26.52939321784587
12 norm of gradients: 0.16327730910462493
12 norm of weights: 5.568007001052807
-----
Iteration Number: 9873
Loss: 26.528055774441427
12 norm of gradients: 0.16326589455049367
12 norm of weights: 5.567947524487442
-----
Iteration Number: 9874
Loss: 26.526718517310115
12 norm of gradients: 0.1632544816881169
12 norm of weights: 5.567888051951494
-----
Iteration Number: 9875
Loss: 26.525381446411345
12 norm of gradients: 0.1632430705170251
12 norm of weights: 5.567828583444481
-----
Iteration Number: 9876
Loss: 26.524044561705907
12 norm of gradients: 0.1632316610367488
12 norm of weights: 5.567769118965919
-----
Iteration Number: 9877
Loss: 26.522707863150767
12 norm of gradients: 0.16322025324681871
12 norm of weights: 5.5677096585153265
-----
Iteration Number: 9878
Loss: 26.52137135070623
12 norm of gradients: 0.1632088471467657
12 norm of weights: 5.567650202092222
-----
Iteration Number: 9879
Loss: 26.520035024332373
12 norm of gradients: 0.1631974427361208
12 norm of weights: 5.5675907496961194
-----
Iteration Number: 9880

```

```

Loss: 26.51869888398529
l2 norm of gradients: 0.1631860400144151
l2 norm of weights: 5.5675313013265395
-----
Iteration Number: 9881
Loss: 26.5173629296267
l2 norm of gradients: 0.16317463898117993
l2 norm of weights: 5.567471856982999
-----
Iteration Number: 9882
Loss: 26.51602716121585
l2 norm of gradients: 0.16316323963594664
l2 norm of weights: 5.5674124166650145
-----
Iteration Number: 9883
Loss: 26.514691578711638
l2 norm of gradients: 0.1631518419782469
l2 norm of weights: 5.567352980372105
-----
Iteration Number: 9884
Loss: 26.513356182071174
l2 norm of gradients: 0.16314044600761227
l2 norm of weights: 5.567293548103788
-----
Iteration Number: 9885
Loss: 26.512020971258206
l2 norm of gradients: 0.1631290517235747
l2 norm of weights: 5.5672341198595845
-----
Iteration Number: 9886
Loss: 26.5106859462277
l2 norm of gradients: 0.16311765912566614
l2 norm of weights: 5.567174695639008
-----
Iteration Number: 9887
Loss: 26.509351106940876
l2 norm of gradients: 0.16310626821341873
l2 norm of weights: 5.56711527544158
-----
Iteration Number: 9888
Loss: 26.508016453356856
l2 norm of gradients: 0.16309487898636468
l2 norm of weights: 5.567055859266818
-----
Iteration Number: 9889
Loss: 26.506681985434284
l2 norm of gradients: 0.16308349144403647
l2 norm of weights: 5.566996447114239

```

```

-----
Iteration Number: 9890
Loss: 26.505347703133374
l2 norm of gradients: 0.1630721055859666
l2 norm of weights: 5.5669370389833635
-----
Iteration Number: 9891
Loss: 26.504013606414453
l2 norm of gradients: 0.1630607214116878
l2 norm of weights: 5.566877634873709
-----
Iteration Number: 9892
Loss: 26.502679695233727
l2 norm of gradients: 0.16304933892073278
l2 norm of weights: 5.566818234784796
-----
Iteration Number: 9893
Loss: 26.501345969553974
l2 norm of gradients: 0.16303795811263463
l2 norm of weights: 5.566758838716142
-----
Iteration Number: 9894
Loss: 26.500012429332404
l2 norm of gradients: 0.1630265789869264
l2 norm of weights: 5.566699446667265
-----
Iteration Number: 9895
Loss: 26.49867907452911
l2 norm of gradients: 0.16301520154314128
l2 norm of weights: 5.566640058637686
-----
Iteration Number: 9896
Loss: 26.497345905102463
l2 norm of gradients: 0.16300382578081277
l2 norm of weights: 5.566580674626923
-----
Iteration Number: 9897
Loss: 26.496012921013545
l2 norm of gradients: 0.16299245169947424
l2 norm of weights: 5.566521294634496
-----
Iteration Number: 9898
Loss: 26.494680122220785
l2 norm of gradients: 0.16298107929865943
l2 norm of weights: 5.566461918659924
-----
Iteration Number: 9899
Loss: 26.493347508683616

```

```

12 norm of gradients: 0.16296970857790216
12 norm of weights: 5.566402546702726
-----
Iteration Number: 9900
Loss: 26.49201508036133
12 norm of gradients: 0.1629583395367363
12 norm of weights: 5.566343178762422
-----
Iteration Number: 9901
Loss: 26.490682837214827
12 norm of gradients: 0.16294697217469595
12 norm of weights: 5.56628381483853
-----
Iteration Number: 9902
Loss: 26.48935077920249
12 norm of gradients: 0.1629356064913153
12 norm of weights: 5.566224454930572
-----
Iteration Number: 9903
Loss: 26.488018906283845
12 norm of gradients: 0.16292424248612872
12 norm of weights: 5.566165099038067
-----
Iteration Number: 9904
Loss: 26.486687218418613
12 norm of gradients: 0.1629128801586707
12 norm of weights: 5.566105747160536
-----
Iteration Number: 9905
Loss: 26.48535571556578
12 norm of gradients: 0.16290151950847584
12 norm of weights: 5.566046399297497
-----
Iteration Number: 9906
Loss: 26.484024397685367
12 norm of gradients: 0.16289016053507896
12 norm of weights: 5.565987055448472
-----
Iteration Number: 9907
Loss: 26.482693264737158
12 norm of gradients: 0.16287880323801485
12 norm of weights: 5.56592771561298
-----
Iteration Number: 9908
Loss: 26.481362316679174
12 norm of gradients: 0.16286744761681862
12 norm of weights: 5.565868379790541
-----

```

Iteration Number: 9909
Loss: 26.48003155347351
l2 norm of gradients: 0.16285609367102544
l2 norm of weights: 5.565809047980677

Iteration Number: 9910
Loss: 26.478700975078052
l2 norm of gradients: 0.16284474140017058
l2 norm of weights: 5.5657497201829065

Iteration Number: 9911
Loss: 26.477370581451904
l2 norm of gradients: 0.16283339080378953
l2 norm of weights: 5.565690396396752

Iteration Number: 9912
Loss: 26.476040372557268
l2 norm of gradients: 0.16282204188141788
l2 norm of weights: 5.565631076621733

Iteration Number: 9913
Loss: 26.474710348351735
l2 norm of gradients: 0.1628106946325913
l2 norm of weights: 5.56557176085737

Iteration Number: 9914
Loss: 26.473380508795003
l2 norm of gradients: 0.16279934905684568
l2 norm of weights: 5.565512449103185

Iteration Number: 9915
Loss: 26.47205085384708
l2 norm of gradients: 0.162788005153717
l2 norm of weights: 5.565453141358699

Iteration Number: 9916
Loss: 26.470721383467573
l2 norm of gradients: 0.1627766629227414
l2 norm of weights: 5.565393837623431

Iteration Number: 9917
Loss: 26.46939209761468
l2 norm of gradients: 0.16276532236345514
l2 norm of weights: 5.565334537896906

Iteration Number: 9918
Loss: 26.468062996252694
l2 norm of gradients: 0.16275398347539458


```

12 norm of weights: 5.565275242178641
-----
Iteration Number: 9919
Loss: 26.466734079334636
12 norm of gradients: 0.16274264625809637
12 norm of weights: 5.56521595046816
-----
Iteration Number: 9920
Loss: 26.4654053468264
12 norm of gradients: 0.1627313107110971
12 norm of weights: 5.565156662764984
-----
Iteration Number: 9921
Loss: 26.464076798683713
12 norm of gradients: 0.16271997683393352
12 norm of weights: 5.565097379068635
-----
Iteration Number: 9922
Loss: 26.462748434868793
12 norm of gradients: 0.16270864462614268
12 norm of weights: 5.565038099378633
-----
Iteration Number: 9923
Loss: 26.461420255339263
12 norm of gradients: 0.16269731408726165
12 norm of weights: 5.564978823694501
-----
Iteration Number: 9924
Loss: 26.460092260057536
12 norm of gradients: 0.16268598521682756
12 norm of weights: 5.564919552015762
-----
Iteration Number: 9925
Loss: 26.458764448981988
12 norm of gradients: 0.16267465801437786
12 norm of weights: 5.564860284341934
-----
Iteration Number: 9926
Loss: 26.457436822071088
12 norm of gradients: 0.16266333247945
12 norm of weights: 5.564801020672544
-----
Iteration Number: 9927
Loss: 26.45610937928711
12 norm of gradients: 0.16265200861158155
12 norm of weights: 5.56474176100711
-----
Iteration Number: 9928

```

Loss: 26.454782120587673
l2 norm of gradients: 0.1626406864103103
l2 norm of weights: 5.564682505345157

Iteration Number: 9929
Loss: 26.453455045934543
l2 norm of gradients: 0.1626293658751742
l2 norm of weights: 5.564623253686207

Iteration Number: 9930
Loss: 26.452128155284854
l2 norm of gradients: 0.16261804700571114
l2 norm of weights: 5.56456400602978

Iteration Number: 9931
Loss: 26.450801448602167
l2 norm of gradients: 0.16260672980145943
l2 norm of weights: 5.564504762375402

Iteration Number: 9932
Loss: 26.449474925843216
l2 norm of gradients: 0.16259541426195723
l2 norm of weights: 5.564445522722593

Iteration Number: 9933
Loss: 26.4481485869698
l2 norm of gradients: 0.16258410038674304
l2 norm of weights: 5.564386287070877

Iteration Number: 9934
Loss: 26.44682243194045
l2 norm of gradients: 0.1625727881753554
l2 norm of weights: 5.564327055419777

Iteration Number: 9935
Loss: 26.44549646071764
l2 norm of gradients: 0.162561477627333
l2 norm of weights: 5.564267827768814

Iteration Number: 9936
Loss: 26.44417067325845
l2 norm of gradients: 0.16255016874221467
l2 norm of weights: 5.564208604117513

Iteration Number: 9937
Loss: 26.44284506952298
l2 norm of gradients: 0.16253886151953933
l2 norm of weights: 5.564149384465396

```
-----  
Iteration Number: 9938  
Loss: 26.441519649473463  
l2 norm of gradients: 0.16252755595884621  
l2 norm of weights: 5.564090168811987  
-----  
Iteration Number: 9939  
Loss: 26.44019441306598  
l2 norm of gradients: 0.16251625205967438  
l2 norm of weights: 5.56403095715681  
-----  
Iteration Number: 9940  
Loss: 26.438869360265752  
l2 norm of gradients: 0.16250494982156327  
l2 norm of weights: 5.563971749499385  
-----  
Iteration Number: 9941  
Loss: 26.437544491029957  
l2 norm of gradients: 0.16249364924405238  
l2 norm of weights: 5.563912545839239  
-----  
Iteration Number: 9942  
Loss: 26.436219805318185  
l2 norm of gradients: 0.16248235032668135  
l2 norm of weights: 5.563853346175894  
-----  
Iteration Number: 9943  
Loss: 26.4348953030899  
l2 norm of gradients: 0.16247105306898987  
l2 norm of weights: 5.563794150508874  
-----  
Iteration Number: 9944  
Loss: 26.43357098430776  
l2 norm of gradients: 0.16245975747051788  
l2 norm of weights: 5.563734958837704  
-----  
Iteration Number: 9945  
Loss: 26.43224684892925  
l2 norm of gradients: 0.16244846353080544  
l2 norm of weights: 5.563675771161905  
-----  
Iteration Number: 9946  
Loss: 26.4309228969163  
l2 norm of gradients: 0.16243717124939266  
l2 norm of weights: 5.563616587481004  
-----  
Iteration Number: 9947  
Loss: 26.429599128228187
```

```

12 norm of gradients: 0.16242588062581984
12 norm of weights: 5.563557407794522
-----
Iteration Number: 9948
Loss: 26.428275542825244
12 norm of gradients: 0.16241459165962746
12 norm of weights: 5.563498232101987
-----
Iteration Number: 9949
Loss: 26.42695214066804
12 norm of gradients: 0.16240330435035594
12 norm of weights: 5.563439060402919
-----
Iteration Number: 9950
Loss: 26.42562892171435
12 norm of gradients: 0.1623920186975461
12 norm of weights: 5.563379892696846
-----
Iteration Number: 9951
Loss: 26.424305885927165
12 norm of gradients: 0.16238073470073866
12 norm of weights: 5.56332072898329
-----
Iteration Number: 9952
Loss: 26.422983033263712
12 norm of gradients: 0.16236945235947464
12 norm of weights: 5.563261569261775
-----
Iteration Number: 9953
Loss: 26.42166036368906
12 norm of gradients: 0.1623581716732951
12 norm of weights: 5.563202413531829
-----
Iteration Number: 9954
Loss: 26.42033787715809
12 norm of gradients: 0.16234689264174126
12 norm of weights: 5.563143261792974
-----
Iteration Number: 9955
Loss: 26.41901557363279
12 norm of gradients: 0.16233561526435444
12 norm of weights: 5.5630841140447345
-----
Iteration Number: 9956
Loss: 26.41769345307398
12 norm of gradients: 0.16232433954067613
12 norm of weights: 5.563024970286638
-----

```

Iteration Number: 9957
Loss: 26.416371515441814
l2 norm of gradients: 0.16231306547024793
l2 norm of weights: 5.562965830518207

Iteration Number: 9958
Loss: 26.41504976069664
l2 norm of gradients: 0.1623017930526116
l2 norm of weights: 5.5629066947389685

Iteration Number: 9959
Loss: 26.413728188798338
l2 norm of gradients: 0.162290522287309
l2 norm of weights: 5.562847562948446

Iteration Number: 9960
Loss: 26.41240679970766
l2 norm of gradients: 0.1622792531738821
l2 norm of weights: 5.562788435146165

Iteration Number: 9961
Loss: 26.411085593383486
l2 norm of gradients: 0.1622679857118731
l2 norm of weights: 5.562729311331651

Iteration Number: 9962
Loss: 26.409764569788166
l2 norm of gradients: 0.1622567199008242
l2 norm of weights: 5.562670191504431

Iteration Number: 9963
Loss: 26.40844372887991
l2 norm of gradients: 0.1622454557402778
l2 norm of weights: 5.562611075664029

Iteration Number: 9964
Loss: 26.407123070620244
l2 norm of gradients: 0.16223419322977642
l2 norm of weights: 5.562551963809971

Iteration Number: 9965
Loss: 26.40580259496848
l2 norm of gradients: 0.1622229323688628
l2 norm of weights: 5.562492855941784

Iteration Number: 9966
Loss: 26.40448230188659
l2 norm of gradients: 0.16221167315707954

```
12 norm of weights: 5.5624337520589915
-----
Iteration Number: 9967
Loss: 26.403162191333895
12 norm of gradients: 0.1622004155939697
12 norm of weights: 5.562374652161121
-----
Iteration Number: 9968
Loss: 26.40184226327069
12 norm of gradients: 0.1621891596790763
12 norm of weights: 5.562315556247698
-----
Iteration Number: 9969
Loss: 26.40052251765798
12 norm of gradients: 0.1621779054119425
12 norm of weights: 5.562256464318248
-----
Iteration Number: 9970
Loss: 26.399202954455692
12 norm of gradients: 0.16216665279211154
12 norm of weights: 5.562197376372301
-----
Iteration Number: 9971
Loss: 26.397883573623282
12 norm of gradients: 0.16215540181912697
12 norm of weights: 5.5621382924093785
-----
Iteration Number: 9972
Loss: 26.396564375123894
12 norm of gradients: 0.16214415249253228
12 norm of weights: 5.562079212429008
-----
Iteration Number: 9973
Loss: 26.39524535891391
12 norm of gradients: 0.16213290481187115
12 norm of weights: 5.562020136430719
-----
Iteration Number: 9974
Loss: 26.393926524958385
12 norm of gradients: 0.1621216587766874
12 norm of weights: 5.561961064414034
-----
Iteration Number: 9975
Loss: 26.39260787321266
12 norm of gradients: 0.16211041438652504
12 norm of weights: 5.5619019963784835
-----
Iteration Number: 9976
```

```

Loss: 26.39128940364231
l2 norm of gradients: 0.16209917164092807
l2 norm of weights: 5.561842932323592
-----
Iteration Number: 9977
Loss: 26.389971116205118
l2 norm of gradients: 0.1620879305394407
l2 norm of weights: 5.561783872248887
-----
Iteration Number: 9978
Loss: 26.38865301086146
l2 norm of gradients: 0.16207669108160733
l2 norm of weights: 5.561724816153895
-----
Iteration Number: 9979
Loss: 26.3873350875727
l2 norm of gradients: 0.1620654532669724
l2 norm of weights: 5.561665764038143
-----
Iteration Number: 9980
Loss: 26.38601734629882
l2 norm of gradients: 0.1620542170950804
l2 norm of weights: 5.56160671590116
-----
Iteration Number: 9981
Loss: 26.384699787000976
l2 norm of gradients: 0.16204298256547622
l2 norm of weights: 5.5615476717424706
-----
Iteration Number: 9982
Loss: 26.383382409639058
l2 norm of gradients: 0.16203174967770456
l2 norm of weights: 5.561488631561605
-----
Iteration Number: 9983
Loss: 26.38206521417287
l2 norm of gradients: 0.1620205184313105
l2 norm of weights: 5.561429595358089
-----
Iteration Number: 9984
Loss: 26.380748200565467
l2 norm of gradients: 0.1620092888258391
l2 norm of weights: 5.561370563131449
-----
Iteration Number: 9985
Loss: 26.37943136877489
l2 norm of gradients: 0.16199806086083562
l2 norm of weights: 5.561311534881216

```

```
-----  
Iteration Number: 9986  
Loss: 26.37811471876197  
l2 norm of gradients: 0.1619868345358454  
l2 norm of weights: 5.561252510606915  
-----  
Iteration Number: 9987  
Loss: 26.376798250488843  
l2 norm of gradients: 0.16197560985041384  
l2 norm of weights: 5.5611934903080735  
-----  
Iteration Number: 9988  
Loss: 26.37548196391586  
l2 norm of gradients: 0.16196438680408673  
l2 norm of weights: 5.561134473984222  
-----  
Iteration Number: 9989  
Loss: 26.374165859002968  
l2 norm of gradients: 0.1619531653964097  
l2 norm of weights: 5.561075461634887  
-----  
Iteration Number: 9990  
Loss: 26.372849935711354  
l2 norm of gradients: 0.1619419456269286  
l2 norm of weights: 5.5610164532595965  
-----  
Iteration Number: 9991  
Loss: 26.371534194000585  
l2 norm of gradients: 0.1619307274951895  
l2 norm of weights: 5.560957448857879  
-----  
Iteration Number: 9992  
Loss: 26.37021863383227  
l2 norm of gradients: 0.16191951100073854  
l2 norm of weights: 5.560898448429263  
-----  
Iteration Number: 9993  
Loss: 26.368903255167176  
l2 norm of gradients: 0.16190829614312185  
l2 norm of weights: 5.560839451973277  
-----  
Iteration Number: 9994  
Loss: 26.36758805796518  
l2 norm of gradients: 0.16189708292188593  
l2 norm of weights: 5.5607804594894485  
-----  
Iteration Number: 9995  
Loss: 26.366273042188276
```



```

l2 norm of gradients: 0.16188587133657725
l2 norm of weights: 5.560721470977307
-----
Iteration Number: 9996
Loss: 26.364958207796427
l2 norm of gradients: 0.1618746613867424
l2 norm of weights: 5.5606624864363825
-----
Iteration Number: 9997
Loss: 26.363643554750873
l2 norm of gradients: 0.1618634530719282
l2 norm of weights: 5.560603505866201
-----
Iteration Number: 9998
Loss: 26.362329083010465
l2 norm of gradients: 0.16185224639168153
l2 norm of weights: 5.560544529266292
-----
Iteration Number: 9999
Loss: 26.36101479253773
l2 norm of gradients: 0.16184104134554933
l2 norm of weights: 5.560485556636186
-----

```

```
[34]: print("Train accuracy: ", accuracy(xtrain_normal, ytrain, model))
```

```
Train accuracy: 0.8307692307692308
```

```
[35]: #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 0.9280719280719281
0.01 2 0.922077922077922
0.01 1 0.9250749250749251
0.01 0.1 0.9270729270729271
0.01 0.01 0.9180819180819181
0.001 5 0.8111888111888111
0.001 2 0.8231768231768232

```

```

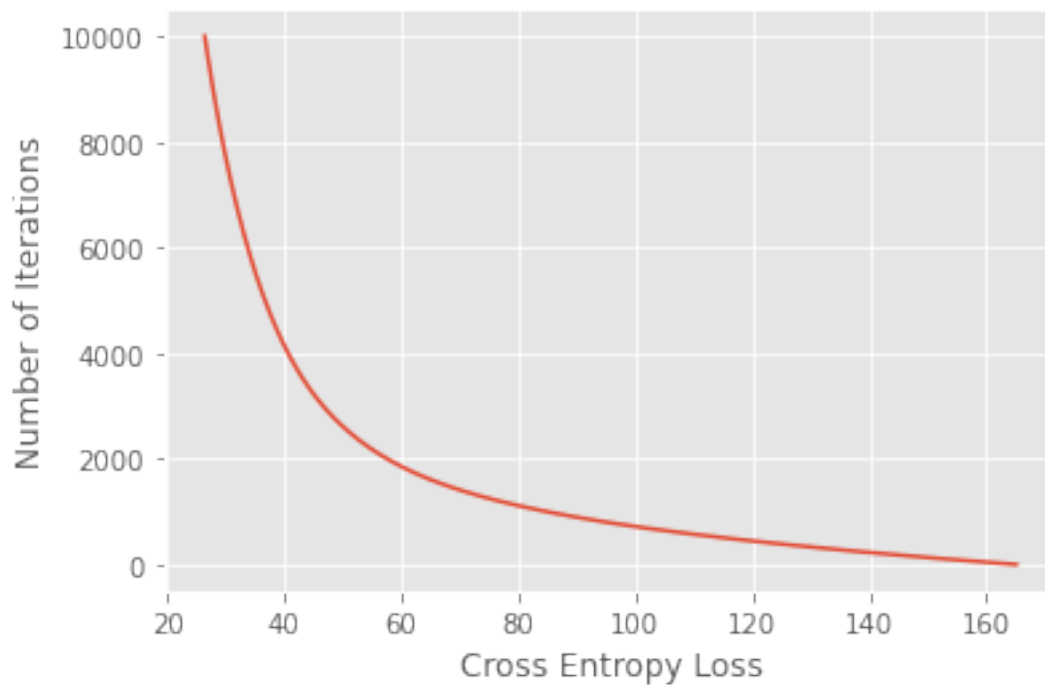
0.001 1 0.8221778221778222
0.001 0.1 0.8251748251748252
0.001 0.01 0.8331668331668332
0.0001 5 0.6823176823176823
0.0001 2 0.7612387612387612
0.0001 1 0.46053946053946054
0.0001 0.1 0.7462537462537463
0.0001 0.01 0.7412587412587412
1e-05 5 0.5894105894105894
1e-05 2 0.5724275724275725
1e-05 1 0.4965034965034965
1e-05 0.1 0.5634365634365635
1e-05 0.01 0.4565434565434565

```

```
[36]: print("Test accuracy: ", accuracy(xtest_normal, ytest, best_model))
```

Test accuracy: 0.941

```
[37]: 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()
```



[]: