

exercise

October 19, 2020

```
[6]: import numpy as np
from sklearn.datasets import load_breast_cancer

data = load_breast_cancer()
y = np.matrix(data.target).T
X = np.matrix(data.data)
M = X.shape[0]
N = X.shape[1]

# Normalize each input feature

def normalize(X):
    M = X.shape[0]
    XX = X - np.tile(np.mean(X,0), [M,1])
    XX = np.divide(XX, np.tile(np.std(XX,0), [M,1]))
    return XX

XX = normalize(X)
```

```
[7]: # Let's start with a 3-layer network with sigmoid activation functions,
# 6 units in layer 1, and 5 units in layer 2.
h2 = 5
h1 = 6

W = [[], np.random.normal(0,0.1, [N,h1]),
      np.random.normal(0,0.1, [h1,1])]
b = [[], np.random.normal(0,0.1, [h1,1]),
      np.random.normal(0,0.1, [1,1])]
L = len(W)-1
print(len(W))
```

3

```
[8]: def act(z):
    return 1/(1+np.exp(-z))

def actder(z):
    az = act(z)
```

```

    prod = np.multiply(az,1-az)
    return prod

def actder_relu(z):
    der = np.ones((z.shape[0],1))
    der[z < 0] = 0
    return der

def ff(x,W,b):
    L = len(W)-1
    a = x
    for l in range(1,L+1):
        z = W[l].T*a+b[l]
        a = act(z)
    return a

def loss(y,yhat):
    return -((1-y) * np.log(1-yhat) + y * np.log(yhat))

```

```

[9]: asd = np.array([[1],[-2],[3]])
    actder_relu(asd)

```

```

[9]: array([[1.],
           [0.],
           [1.]])

```

```

[10]: # Use mini-batch size 1

alpha = 0.01
max_iter = 1000
for iter in range(0, max_iter):
    loss_this_iter = 0
    order = np.random.permutation(M)
    for i in range(0,M):

        # Grab the pattern order[i]

        x_this = XX[order[i],:].T
        y_this = y[order[i],0]

        # Feed forward step

        a = [x_this]
        z = [[]]
        delta = [[]]

```

```

dW = [[]]
db = [[]]
for l in range(1,L+1):
    z.append(W[l].T*a[l-1]+b[l])
    a.append(act(z[l]))
    # Just to give arrays the right shape for the backprop step
    delta.append([]); dW.append([]); db.append([])

loss_this_pattern = loss(y_this, a[L][0,0])
loss_this_iter = loss_this_iter + loss_this_pattern

# Backprop step

delta[L] = a[L] - y_this
for l in range(L,0,-1):
    db[l] = delta[l].copy()
    dW[l] = a[l-1] * delta[l].T
    if l == L:
        delta[l-1] = np.multiply(actder(z[l-1]), W[l] * delta[l])
    elif l > 1:
        delta[l-1] = np.multiply(actder_relu(z[l-1]), W[l] * delta[l])

# Check delta calculation

if False:
    print('Target: %f' % y_this)
    print('y_hat: %f' % a[L][0,0])
    print(db)
    y_pred = ff(x_this,W,b)
    diff = 1e-3
    W[1][10,0] = W[1][10,0] + diff
    y_pred_db = ff(x_this,W,b)
    L1 = loss(y_this,y_pred)
    L2 = loss(y_this,y_pred_db)
    db_finite_difference = (L2-L1)/diff
    print('Original out %f, perturbed out %f' %
          (y_pred[0,0], y_pred_db[0,0]))
    print('Theoretical dW %f, calculated db %f' %
          (dW[1][10,0], db_finite_difference[0,0]))

for l in range(1,L+1):
    W[l] = W[l] - alpha * dW[l]
    b[l] = b[l] - alpha * db[l]

print('Iteration %d loss %f' % (iter, loss_this_iter))

```

Iteration 0 loss 357.589422

Iteration 1 loss 251.613570
Iteration 2 loss 160.577862
Iteration 3 loss 114.689459
Iteration 4 loss 90.423708
Iteration 5 loss 76.067088
Iteration 6 loss 66.849790
Iteration 7 loss 60.367138
Iteration 8 loss 55.638421
Iteration 9 loss 52.112172
Iteration 10 loss 49.361210
Iteration 11 loss 47.077578
Iteration 12 loss 45.235756
Iteration 13 loss 43.649921
Iteration 14 loss 42.317802
Iteration 15 loss 41.160336
Iteration 16 loss 40.149695
Iteration 17 loss 39.223714
Iteration 18 loss 38.483771
Iteration 19 loss 37.693011
Iteration 20 loss 37.068589
Iteration 21 loss 36.503450
Iteration 22 loss 35.932096
Iteration 23 loss 35.431291
Iteration 24 loss 34.966018
Iteration 25 loss 34.543172
Iteration 26 loss 34.181654
Iteration 27 loss 33.811210
Iteration 28 loss 33.450442
Iteration 29 loss 33.143131
Iteration 30 loss 32.837665
Iteration 31 loss 32.556737
Iteration 32 loss 32.306591
Iteration 33 loss 32.032675
Iteration 34 loss 31.711356
Iteration 35 loss 31.567284
Iteration 36 loss 31.374420
Iteration 37 loss 31.136843
Iteration 38 loss 30.956596
Iteration 39 loss 30.730318
Iteration 40 loss 30.581424
Iteration 41 loss 30.413134
Iteration 42 loss 30.270662
Iteration 43 loss 30.074773
Iteration 44 loss 29.939028
Iteration 45 loss 29.716713
Iteration 46 loss 29.656442
Iteration 47 loss 29.528918
Iteration 48 loss 29.394977

Iteration 49 loss 29.301523
Iteration 50 loss 29.128447
Iteration 51 loss 29.049073
Iteration 52 loss 28.913848
Iteration 53 loss 28.831160
Iteration 54 loss 28.729609
Iteration 55 loss 28.584862
Iteration 56 loss 28.483993
Iteration 57 loss 28.413723
Iteration 58 loss 28.334458
Iteration 59 loss 28.219552
Iteration 60 loss 28.139634
Iteration 61 loss 28.056188
Iteration 62 loss 27.973022
Iteration 63 loss 27.894645
Iteration 64 loss 27.760312
Iteration 65 loss 27.702875
Iteration 66 loss 27.690334
Iteration 67 loss 27.536163
Iteration 68 loss 27.481785
Iteration 69 loss 27.435447
Iteration 70 loss 27.381592
Iteration 71 loss 27.249124
Iteration 72 loss 27.119201
Iteration 73 loss 27.147534
Iteration 74 loss 26.854483
Iteration 75 loss 27.036668
Iteration 76 loss 26.937934
Iteration 77 loss 26.857793
Iteration 78 loss 26.712116
Iteration 79 loss 26.742991
Iteration 80 loss 26.697962
Iteration 81 loss 26.638544
Iteration 82 loss 26.629605
Iteration 83 loss 26.500019
Iteration 84 loss 26.381903
Iteration 85 loss 26.363486
Iteration 86 loss 26.303300
Iteration 87 loss 26.234330
Iteration 88 loss 26.203544
Iteration 89 loss 26.049858
Iteration 90 loss 26.096364
Iteration 91 loss 26.019254
Iteration 92 loss 25.925220
Iteration 93 loss 25.953019
Iteration 94 loss 25.834108
Iteration 95 loss 25.795679
Iteration 96 loss 25.643351

Iteration 97 loss 25.704615
Iteration 98 loss 25.458688
Iteration 99 loss 25.663750
Iteration 100 loss 25.481128
Iteration 101 loss 25.427551
Iteration 102 loss 25.389798
Iteration 103 loss 25.340792
Iteration 104 loss 25.140500
Iteration 105 loss 25.187375
Iteration 106 loss 25.105012
Iteration 107 loss 24.982339
Iteration 108 loss 25.134499
Iteration 109 loss 24.940789
Iteration 110 loss 24.944372
Iteration 111 loss 24.914682
Iteration 112 loss 24.773893
Iteration 113 loss 24.794191
Iteration 114 loss 24.667396
Iteration 115 loss 24.642217
Iteration 116 loss 24.631808
Iteration 117 loss 24.540476
Iteration 118 loss 24.488338
Iteration 119 loss 24.477856
Iteration 120 loss 24.345612
Iteration 121 loss 24.319775
Iteration 122 loss 24.236840
Iteration 123 loss 24.228469
Iteration 124 loss 24.168148
Iteration 125 loss 24.145690
Iteration 126 loss 24.092512
Iteration 127 loss 23.962636
Iteration 128 loss 23.972078
Iteration 129 loss 23.937603
Iteration 130 loss 23.915367
Iteration 131 loss 23.796804
Iteration 132 loss 23.742622
Iteration 133 loss 23.739897
Iteration 134 loss 23.686056
Iteration 135 loss 23.650928
Iteration 136 loss 23.619325
Iteration 137 loss 23.552967
Iteration 138 loss 23.522431
Iteration 139 loss 23.458279
Iteration 140 loss 23.410587
Iteration 141 loss 23.404861
Iteration 142 loss 23.334604
Iteration 143 loss 23.253537
Iteration 144 loss 23.282210

Iteration 145 loss 23.004951
Iteration 146 loss 23.271777
Iteration 147 loss 23.074161
Iteration 148 loss 23.048105
Iteration 149 loss 23.006262
Iteration 150 loss 23.005883
Iteration 151 loss 22.927480
Iteration 152 loss 22.904680
Iteration 153 loss 22.876070
Iteration 154 loss 22.824149
Iteration 155 loss 22.764083
Iteration 156 loss 22.650193
Iteration 157 loss 22.707256
Iteration 158 loss 22.651042
Iteration 159 loss 22.624524
Iteration 160 loss 22.573404
Iteration 161 loss 22.468289
Iteration 162 loss 22.519759
Iteration 163 loss 22.500244
Iteration 164 loss 22.411416
Iteration 165 loss 22.366984
Iteration 166 loss 22.326893
Iteration 167 loss 22.233600
Iteration 168 loss 22.248969
Iteration 169 loss 22.231117
Iteration 170 loss 22.160574
Iteration 171 loss 22.151871
Iteration 172 loss 22.081901
Iteration 173 loss 22.029322
Iteration 174 loss 22.056497
Iteration 175 loss 21.948761
Iteration 176 loss 21.939897
Iteration 177 loss 21.904314
Iteration 178 loss 21.796231
Iteration 179 loss 21.847699
Iteration 180 loss 21.791956
Iteration 181 loss 21.666987
Iteration 182 loss 21.706875
Iteration 183 loss 21.687334
Iteration 184 loss 21.634467
Iteration 185 loss 21.610604
Iteration 186 loss 21.529894
Iteration 187 loss 21.524587
Iteration 188 loss 21.436326
Iteration 189 loss 21.454309
Iteration 190 loss 21.387915
Iteration 191 loss 21.361292
Iteration 192 loss 21.328619

Iteration 193 loss 21.293006
Iteration 194 loss 21.218268
Iteration 195 loss 21.137968
Iteration 196 loss 21.250182
Iteration 197 loss 21.140777
Iteration 198 loss 21.013419
Iteration 199 loss 21.080172
Iteration 200 loss 21.022840
Iteration 201 loss 20.952847
Iteration 202 loss 20.942989
Iteration 203 loss 20.903565
Iteration 204 loss 20.885069
Iteration 205 loss 20.854738
Iteration 206 loss 20.753434
Iteration 207 loss 20.771520
Iteration 208 loss 20.735776
Iteration 209 loss 20.680114
Iteration 210 loss 20.550839
Iteration 211 loss 20.636556
Iteration 212 loss 20.524181
Iteration 213 loss 20.518625
Iteration 214 loss 20.489297
Iteration 215 loss 20.415283
Iteration 216 loss 20.390730
Iteration 217 loss 20.369348
Iteration 218 loss 20.329953
Iteration 219 loss 20.287236
Iteration 220 loss 20.220635
Iteration 221 loss 20.185107
Iteration 222 loss 20.130109
Iteration 223 loss 20.114417
Iteration 224 loss 20.054571
Iteration 225 loss 20.027928
Iteration 226 loss 19.975795
Iteration 227 loss 19.912647
Iteration 228 loss 19.847263
Iteration 229 loss 19.826227
Iteration 230 loss 19.766675
Iteration 231 loss 19.693110
Iteration 232 loss 19.641108
Iteration 233 loss 19.624625
Iteration 234 loss 19.549509
Iteration 235 loss 19.470797
Iteration 236 loss 19.409827
Iteration 237 loss 19.385430
Iteration 238 loss 19.314047
Iteration 239 loss 19.254073
Iteration 240 loss 19.199734

Iteration 241 loss 19.132697
Iteration 242 loss 19.058565
Iteration 243 loss 19.019035
Iteration 244 loss 18.949936
Iteration 245 loss 18.880153
Iteration 246 loss 18.824071
Iteration 247 loss 18.748290
Iteration 248 loss 18.663270
Iteration 249 loss 18.638748
Iteration 250 loss 18.549766
Iteration 251 loss 18.529859
Iteration 252 loss 18.412552
Iteration 253 loss 18.399905
Iteration 254 loss 18.328221
Iteration 255 loss 18.288440
Iteration 256 loss 18.217732
Iteration 257 loss 18.162450
Iteration 258 loss 18.118553
Iteration 259 loss 18.034958
Iteration 260 loss 18.008641
Iteration 261 loss 17.895437
Iteration 262 loss 17.912604
Iteration 263 loss 17.828802
Iteration 264 loss 17.728277
Iteration 265 loss 17.756118
Iteration 266 loss 17.700455
Iteration 267 loss 17.640390
Iteration 268 loss 17.583681
Iteration 269 loss 17.551821
Iteration 270 loss 17.500037
Iteration 271 loss 17.446179
Iteration 272 loss 17.407380
Iteration 273 loss 17.345532
Iteration 274 loss 17.293134
Iteration 275 loss 17.267308
Iteration 276 loss 17.220579
Iteration 277 loss 17.174352
Iteration 278 loss 17.123008
Iteration 279 loss 17.055855
Iteration 280 loss 17.031000
Iteration 281 loss 16.995542
Iteration 282 loss 16.949585
Iteration 283 loss 16.882100
Iteration 284 loss 16.857898
Iteration 285 loss 16.802533
Iteration 286 loss 16.772854
Iteration 287 loss 16.739889
Iteration 288 loss 16.662486

Iteration 289 loss 16.658703
Iteration 290 loss 16.604208
Iteration 291 loss 16.544853
Iteration 292 loss 16.531940
Iteration 293 loss 16.477049
Iteration 294 loss 16.445012
Iteration 295 loss 16.405196
Iteration 296 loss 16.339591
Iteration 297 loss 16.325115
Iteration 298 loss 16.270055
Iteration 299 loss 16.244799
Iteration 300 loss 16.196377
Iteration 301 loss 16.156263
Iteration 302 loss 16.124266
Iteration 303 loss 16.051768
Iteration 304 loss 16.039851
Iteration 305 loss 15.992157
Iteration 306 loss 15.950537
Iteration 307 loss 15.915393
Iteration 308 loss 15.867877
Iteration 309 loss 15.821899
Iteration 310 loss 15.790354
Iteration 311 loss 15.759927
Iteration 312 loss 15.710811
Iteration 313 loss 15.676887
Iteration 314 loss 15.645131
Iteration 315 loss 15.603630
Iteration 316 loss 15.561568
Iteration 317 loss 15.526981
Iteration 318 loss 15.482098
Iteration 319 loss 15.424708
Iteration 320 loss 15.411278
Iteration 321 loss 15.325564
Iteration 322 loss 15.320609
Iteration 323 loss 15.285253
Iteration 324 loss 15.258080
Iteration 325 loss 15.211579
Iteration 326 loss 15.167244
Iteration 327 loss 15.127652
Iteration 328 loss 15.090070
Iteration 329 loss 15.065783
Iteration 330 loss 14.983460
Iteration 331 loss 14.967501
Iteration 332 loss 14.931019
Iteration 333 loss 14.902321
Iteration 334 loss 14.886153
Iteration 335 loss 14.830620
Iteration 336 loss 14.797859

Iteration 337 loss 14.755034
Iteration 338 loss 14.726610
Iteration 339 loss 14.648632
Iteration 340 loss 14.654029
Iteration 341 loss 14.614755
Iteration 342 loss 14.572982
Iteration 343 loss 14.538936
Iteration 344 loss 14.512223
Iteration 345 loss 14.469219
Iteration 346 loss 14.428186
Iteration 347 loss 14.396888
Iteration 348 loss 14.359285
Iteration 349 loss 14.328757
Iteration 350 loss 14.283302
Iteration 351 loss 14.246772
Iteration 352 loss 14.222071
Iteration 353 loss 14.184194
Iteration 354 loss 14.149372
Iteration 355 loss 14.110171
Iteration 356 loss 14.062876
Iteration 357 loss 14.015557
Iteration 358 loss 14.008809
Iteration 359 loss 13.965183
Iteration 360 loss 13.920762
Iteration 361 loss 13.898832
Iteration 362 loss 13.850458
Iteration 363 loss 13.792208
Iteration 364 loss 13.787141
Iteration 365 loss 13.750042
Iteration 366 loss 13.727160
Iteration 367 loss 13.666723
Iteration 368 loss 13.633254
Iteration 369 loss 13.624408
Iteration 370 loss 13.575717
Iteration 371 loss 13.541196
Iteration 372 loss 13.515242
Iteration 373 loss 13.481162
Iteration 374 loss 13.442549
Iteration 375 loss 13.399797
Iteration 376 loss 13.374896
Iteration 377 loss 13.345679
Iteration 378 loss 13.291266
Iteration 379 loss 13.282334
Iteration 380 loss 13.243023
Iteration 381 loss 13.213884
Iteration 382 loss 13.173732
Iteration 383 loss 13.140105
Iteration 384 loss 13.110433

Iteration 385 loss 13.076724
Iteration 386 loss 13.026361
Iteration 387 loss 13.011828
Iteration 388 loss 12.966525
Iteration 389 loss 12.939680
Iteration 390 loss 12.910647
Iteration 391 loss 12.883398
Iteration 392 loss 12.851706
Iteration 393 loss 12.807150
Iteration 394 loss 12.774635
Iteration 395 loss 12.750677
Iteration 396 loss 12.718362
Iteration 397 loss 12.678143
Iteration 398 loss 12.648215
Iteration 399 loss 12.618552
Iteration 400 loss 12.598356
Iteration 401 loss 12.563725
Iteration 402 loss 12.531512
Iteration 403 loss 12.498857
Iteration 404 loss 12.474562
Iteration 405 loss 12.435517
Iteration 406 loss 12.398755
Iteration 407 loss 12.380838
Iteration 408 loss 12.346586
Iteration 409 loss 12.312604
Iteration 410 loss 12.287648
Iteration 411 loss 12.254603
Iteration 412 loss 12.215050
Iteration 413 loss 12.195634
Iteration 414 loss 12.156011
Iteration 415 loss 12.130730
Iteration 416 loss 12.091684
Iteration 417 loss 12.064981
Iteration 418 loss 12.038372
Iteration 419 loss 12.011505
Iteration 420 loss 11.979448
Iteration 421 loss 11.952970
Iteration 422 loss 11.907712
Iteration 423 loss 11.896405
Iteration 424 loss 11.870540
Iteration 425 loss 11.833689
Iteration 426 loss 11.808513
Iteration 427 loss 11.765823
Iteration 428 loss 11.739036
Iteration 429 loss 11.727063
Iteration 430 loss 11.689357
Iteration 431 loss 11.650584
Iteration 432 loss 11.629070

Iteration 433 loss 11.607867
Iteration 434 loss 11.583692
Iteration 435 loss 11.550836
Iteration 436 loss 11.524895
Iteration 437 loss 11.499950
Iteration 438 loss 11.452571
Iteration 439 loss 11.443683
Iteration 440 loss 11.407100
Iteration 441 loss 11.383162
Iteration 442 loss 11.358387
Iteration 443 loss 11.329874
Iteration 444 loss 11.305287
Iteration 445 loss 11.271284
Iteration 446 loss 11.244124
Iteration 447 loss 11.221700
Iteration 448 loss 11.195614
Iteration 449 loss 11.162596
Iteration 450 loss 11.142093
Iteration 451 loss 11.094433
Iteration 452 loss 11.089176
Iteration 453 loss 11.056944
Iteration 454 loss 11.029198
Iteration 455 loss 10.998730
Iteration 456 loss 10.978595
Iteration 457 loss 10.950297
Iteration 458 loss 10.919131
Iteration 459 loss 10.881526
Iteration 460 loss 10.868494
Iteration 461 loss 10.843999
Iteration 462 loss 10.818234
Iteration 463 loss 10.795781
Iteration 464 loss 10.741833
Iteration 465 loss 10.734210
Iteration 466 loss 10.714077
Iteration 467 loss 10.692470
Iteration 468 loss 10.666476
Iteration 469 loss 10.634298
Iteration 470 loss 10.601661
Iteration 471 loss 10.586846
Iteration 472 loss 10.561497
Iteration 473 loss 10.539829
Iteration 474 loss 10.498835
Iteration 475 loss 10.478928
Iteration 476 loss 10.453404
Iteration 477 loss 10.437709
Iteration 478 loss 10.404670
Iteration 479 loss 10.390202
Iteration 480 loss 10.358291

Iteration 481 loss 10.331905
Iteration 482 loss 10.311355
Iteration 483 loss 10.287086
Iteration 484 loss 10.269450
Iteration 485 loss 10.242075
Iteration 486 loss 10.215961
Iteration 487 loss 10.194865
Iteration 488 loss 10.129906
Iteration 489 loss 10.142453
Iteration 490 loss 10.122849
Iteration 491 loss 10.057169
Iteration 492 loss 10.075342
Iteration 493 loss 10.050142
Iteration 494 loss 10.011874
Iteration 495 loss 9.984705
Iteration 496 loss 9.958153
Iteration 497 loss 9.931546
Iteration 498 loss 9.905567
Iteration 499 loss 9.906963
Iteration 500 loss 9.875368
Iteration 501 loss 9.853420
Iteration 502 loss 9.833714
Iteration 503 loss 9.810278
Iteration 504 loss 9.786087
Iteration 505 loss 9.759547
Iteration 506 loss 9.739146
Iteration 507 loss 9.713921
Iteration 508 loss 9.682397
Iteration 509 loss 9.673382
Iteration 510 loss 9.626281
Iteration 511 loss 9.631595
Iteration 512 loss 9.604699
Iteration 513 loss 9.574378
Iteration 514 loss 9.547557
Iteration 515 loss 9.524316
Iteration 516 loss 9.512358
Iteration 517 loss 9.483104
Iteration 518 loss 9.461259
Iteration 519 loss 9.445187
Iteration 520 loss 9.411163
Iteration 521 loss 9.397917
Iteration 522 loss 9.381069
Iteration 523 loss 9.351842
Iteration 524 loss 9.331292
Iteration 525 loss 9.309335
Iteration 526 loss 9.274840
Iteration 527 loss 9.260363
Iteration 528 loss 9.241335

Iteration 529 loss 9.223437
Iteration 530 loss 9.205346
Iteration 531 loss 9.179874
Iteration 532 loss 9.157341
Iteration 533 loss 9.134358
Iteration 534 loss 9.082479
Iteration 535 loss 9.102621
Iteration 536 loss 9.069247
Iteration 537 loss 9.048728
Iteration 538 loss 9.028186
Iteration 539 loss 9.006438
Iteration 540 loss 8.983125
Iteration 541 loss 8.954668
Iteration 542 loss 8.943305
Iteration 543 loss 8.904739
Iteration 544 loss 8.894440
Iteration 545 loss 8.863938
Iteration 546 loss 8.862150
Iteration 547 loss 8.807559
Iteration 548 loss 8.810612
Iteration 549 loss 8.794977
Iteration 550 loss 8.766818
Iteration 551 loss 8.754372
Iteration 552 loss 8.727150
Iteration 553 loss 8.711093
Iteration 554 loss 8.692852
Iteration 555 loss 8.655635
Iteration 556 loss 8.622357
Iteration 557 loss 8.627929
Iteration 558 loss 8.606422
Iteration 559 loss 8.586961
Iteration 560 loss 8.568857
Iteration 561 loss 8.546127
Iteration 562 loss 8.530762
Iteration 563 loss 8.503567
Iteration 564 loss 8.487928
Iteration 565 loss 8.463666
Iteration 566 loss 8.438650
Iteration 567 loss 8.419607
Iteration 568 loss 8.406970
Iteration 569 loss 8.364015
Iteration 570 loss 8.369172
Iteration 571 loss 8.337125
Iteration 572 loss 8.328673
Iteration 573 loss 8.309285
Iteration 574 loss 8.285798
Iteration 575 loss 8.261012
Iteration 576 loss 8.242696

Iteration 577 loss 8.228700
Iteration 578 loss 8.210309
Iteration 579 loss 8.187554
Iteration 580 loss 8.148266
Iteration 581 loss 8.146667
Iteration 582 loss 8.127951
Iteration 583 loss 8.097941
Iteration 584 loss 8.089779
Iteration 585 loss 8.071334
Iteration 586 loss 8.057064
Iteration 587 loss 8.029453
Iteration 588 loss 8.015590
Iteration 589 loss 7.993529
Iteration 590 loss 7.975731
Iteration 591 loss 7.955727
Iteration 592 loss 7.942386
Iteration 593 loss 7.919265
Iteration 594 loss 7.893938
Iteration 595 loss 7.883429
Iteration 596 loss 7.863059
Iteration 597 loss 7.839762
Iteration 598 loss 7.828854
Iteration 599 loss 7.809025
Iteration 600 loss 7.789274
Iteration 601 loss 7.762144
Iteration 602 loss 7.748797
Iteration 603 loss 7.736343
Iteration 604 loss 7.718317
Iteration 605 loss 7.676970
Iteration 606 loss 7.683893
Iteration 607 loss 7.662022
Iteration 608 loss 7.639205
Iteration 609 loss 7.619818
Iteration 610 loss 7.603522
Iteration 611 loss 7.583819
Iteration 612 loss 7.568465
Iteration 613 loss 7.550091
Iteration 614 loss 7.532645
Iteration 615 loss 7.507992
Iteration 616 loss 7.493834
Iteration 617 loss 7.477996
Iteration 618 loss 7.457972
Iteration 619 loss 7.446636
Iteration 620 loss 7.425860
Iteration 621 loss 7.403201
Iteration 622 loss 7.393185
Iteration 623 loss 7.370615
Iteration 624 loss 7.349453

Iteration 625 loss 7.340328
Iteration 626 loss 7.322386
Iteration 627 loss 7.304162
Iteration 628 loss 7.286052
Iteration 629 loss 7.262689
Iteration 630 loss 7.250777
Iteration 631 loss 7.220384
Iteration 632 loss 7.212816
Iteration 633 loss 7.193594
Iteration 634 loss 7.172994
Iteration 635 loss 7.170710
Iteration 636 loss 7.150273
Iteration 637 loss 7.132898
Iteration 638 loss 7.111587
Iteration 639 loss 7.096455
Iteration 640 loss 7.073664
Iteration 641 loss 7.065002
Iteration 642 loss 7.029438
Iteration 643 loss 7.028638
Iteration 644 loss 7.007292
Iteration 645 loss 6.987824
Iteration 646 loss 6.970582
Iteration 647 loss 6.955405
Iteration 648 loss 6.941311
Iteration 649 loss 6.925001
Iteration 650 loss 6.911487
Iteration 651 loss 6.899293
Iteration 652 loss 6.878267
Iteration 653 loss 6.862813
Iteration 654 loss 6.842171
Iteration 655 loss 6.823965
Iteration 656 loss 6.812086
Iteration 657 loss 6.795217
Iteration 658 loss 6.784509
Iteration 659 loss 6.766063
Iteration 660 loss 6.749791
Iteration 661 loss 6.733421
Iteration 662 loss 6.716061
Iteration 663 loss 6.702272
Iteration 664 loss 6.687701
Iteration 665 loss 6.670439
Iteration 666 loss 6.655188
Iteration 667 loss 6.631436
Iteration 668 loss 6.611598
Iteration 669 loss 6.603007
Iteration 670 loss 6.565556
Iteration 671 loss 6.566549
Iteration 672 loss 6.563841

Iteration 673 loss 6.542034
Iteration 674 loss 6.528755
Iteration 675 loss 6.484577
Iteration 676 loss 6.500719
Iteration 677 loss 6.484821
Iteration 678 loss 6.465616
Iteration 679 loss 6.437168
Iteration 680 loss 6.436064
Iteration 681 loss 6.420990
Iteration 682 loss 6.404142
Iteration 683 loss 6.390740
Iteration 684 loss 6.365110
Iteration 685 loss 6.349867
Iteration 686 loss 6.344507
Iteration 687 loss 6.324358
Iteration 688 loss 6.317384
Iteration 689 loss 6.301538
Iteration 690 loss 6.279269
Iteration 691 loss 6.270297
Iteration 692 loss 6.246600
Iteration 693 loss 6.233408
Iteration 694 loss 6.227855
Iteration 695 loss 6.207621
Iteration 696 loss 6.197001
Iteration 697 loss 6.183833
Iteration 698 loss 6.159299
Iteration 699 loss 6.141318
Iteration 700 loss 6.144012
Iteration 701 loss 6.124491
Iteration 702 loss 6.109051
Iteration 703 loss 6.097244
Iteration 704 loss 6.066103
Iteration 705 loss 6.064672
Iteration 706 loss 6.049934
Iteration 707 loss 6.039553
Iteration 708 loss 6.017158
Iteration 709 loss 6.008538
Iteration 710 loss 5.995267
Iteration 711 loss 5.967578
Iteration 712 loss 5.969489
Iteration 713 loss 5.953065
Iteration 714 loss 5.939837
Iteration 715 loss 5.925537
Iteration 716 loss 5.900453
Iteration 717 loss 5.895417
Iteration 718 loss 5.882372
Iteration 719 loss 5.871413
Iteration 720 loss 5.856240

Iteration 721 loss 5.842989
Iteration 722 loss 5.826803
Iteration 723 loss 5.803541
Iteration 724 loss 5.799796
Iteration 725 loss 5.790422
Iteration 726 loss 5.773634
Iteration 727 loss 5.749205
Iteration 728 loss 5.737444
Iteration 729 loss 5.733073
Iteration 730 loss 5.724628
Iteration 731 loss 5.708802
Iteration 732 loss 5.687894
Iteration 733 loss 5.678450
Iteration 734 loss 5.665801
Iteration 735 loss 5.655009
Iteration 736 loss 5.639995
Iteration 737 loss 5.625382
Iteration 738 loss 5.610714
Iteration 739 loss 5.599032
Iteration 740 loss 5.588841
Iteration 741 loss 5.562441
Iteration 742 loss 5.555930
Iteration 743 loss 5.543710
Iteration 744 loss 5.536797
Iteration 745 loss 5.525032
Iteration 746 loss 5.512195
Iteration 747 loss 5.498705
Iteration 748 loss 5.480040
Iteration 749 loss 5.474266
Iteration 750 loss 5.453232
Iteration 751 loss 5.445724
Iteration 752 loss 5.430512
Iteration 753 loss 5.419422
Iteration 754 loss 5.412156
Iteration 755 loss 5.399448
Iteration 756 loss 5.382079
Iteration 757 loss 5.373012
Iteration 758 loss 5.359009
Iteration 759 loss 5.342631
Iteration 760 loss 5.334983
Iteration 761 loss 5.327181
Iteration 762 loss 5.314218
Iteration 763 loss 5.300613
Iteration 764 loss 5.282381
Iteration 765 loss 5.275065
Iteration 766 loss 5.263539
Iteration 767 loss 5.249217
Iteration 768 loss 5.237069

Iteration 769 loss 5.228690
Iteration 770 loss 5.216860
Iteration 771 loss 5.199700
Iteration 772 loss 5.190319
Iteration 773 loss 5.179796
Iteration 774 loss 5.169225
Iteration 775 loss 5.156294
Iteration 776 loss 5.145013
Iteration 777 loss 5.133098
Iteration 778 loss 5.119535
Iteration 779 loss 5.110725
Iteration 780 loss 5.088281
Iteration 781 loss 5.087693
Iteration 782 loss 5.073932
Iteration 783 loss 5.064654
Iteration 784 loss 5.052728
Iteration 785 loss 5.040521
Iteration 786 loss 5.020774
Iteration 787 loss 5.018040
Iteration 788 loss 5.008001
Iteration 789 loss 4.995575
Iteration 790 loss 4.985226
Iteration 791 loss 4.973307
Iteration 792 loss 4.961448
Iteration 793 loss 4.950867
Iteration 794 loss 4.940119
Iteration 795 loss 4.921116
Iteration 796 loss 4.917860
Iteration 797 loss 4.903047
Iteration 798 loss 4.896117
Iteration 799 loss 4.881851
Iteration 800 loss 4.874871
Iteration 801 loss 4.861545
Iteration 802 loss 4.848362
Iteration 803 loss 4.840927
Iteration 804 loss 4.831883
Iteration 805 loss 4.821236
Iteration 806 loss 4.803160
Iteration 807 loss 4.798452
Iteration 808 loss 4.787957
Iteration 809 loss 4.773835
Iteration 810 loss 4.765292
Iteration 811 loss 4.755790
Iteration 812 loss 4.743357
Iteration 813 loss 4.734624
Iteration 814 loss 4.722648
Iteration 815 loss 4.700126
Iteration 816 loss 4.702543

Iteration 817 loss 4.695749
Iteration 818 loss 4.684012
Iteration 819 loss 4.672899
Iteration 820 loss 4.662938
Iteration 821 loss 4.650289
Iteration 822 loss 4.642826
Iteration 823 loss 4.632400
Iteration 824 loss 4.620177
Iteration 825 loss 4.612383
Iteration 826 loss 4.601426
Iteration 827 loss 4.589912
Iteration 828 loss 4.580599
Iteration 829 loss 4.571702
Iteration 830 loss 4.562244
Iteration 831 loss 4.551052
Iteration 832 loss 4.541653
Iteration 833 loss 4.532507
Iteration 834 loss 4.522079
Iteration 835 loss 4.512147
Iteration 836 loss 4.496819
Iteration 837 loss 4.492798
Iteration 838 loss 4.482205
Iteration 839 loss 4.469311
Iteration 840 loss 4.462484
Iteration 841 loss 4.454424
Iteration 842 loss 4.439450
Iteration 843 loss 4.435579
Iteration 844 loss 4.425584
Iteration 845 loss 4.415347
Iteration 846 loss 4.404245
Iteration 847 loss 4.397268
Iteration 848 loss 4.385198
Iteration 849 loss 4.377448
Iteration 850 loss 4.366359
Iteration 851 loss 4.360223
Iteration 852 loss 4.349507
Iteration 853 loss 4.341915
Iteration 854 loss 4.332218
Iteration 855 loss 4.315248
Iteration 856 loss 4.311946
Iteration 857 loss 4.304313
Iteration 858 loss 4.293414
Iteration 859 loss 4.286193
Iteration 860 loss 4.279153
Iteration 861 loss 4.267378
Iteration 862 loss 4.257663
Iteration 863 loss 4.246816
Iteration 864 loss 4.239892

Iteration 865 loss 4.231600
Iteration 866 loss 4.221704
Iteration 867 loss 4.214223
Iteration 868 loss 4.204091
Iteration 869 loss 4.197090
Iteration 870 loss 4.186763
Iteration 871 loss 4.178734
Iteration 872 loss 4.169923
Iteration 873 loss 4.158034
Iteration 874 loss 4.153470
Iteration 875 loss 4.141183
Iteration 876 loss 4.135914
Iteration 877 loss 4.126209
Iteration 878 loss 4.118149
Iteration 879 loss 4.109796
Iteration 880 loss 4.100238
Iteration 881 loss 4.092295
Iteration 882 loss 4.083085
Iteration 883 loss 4.075446
Iteration 884 loss 4.065731
Iteration 885 loss 4.059356
Iteration 886 loss 4.047515
Iteration 887 loss 4.041341
Iteration 888 loss 4.034187
Iteration 889 loss 4.025084
Iteration 890 loss 4.008743
Iteration 891 loss 4.009422
Iteration 892 loss 3.998892
Iteration 893 loss 3.991096
Iteration 894 loss 3.984898
Iteration 895 loss 3.974876
Iteration 896 loss 3.967790
Iteration 897 loss 3.959174
Iteration 898 loss 3.950298
Iteration 899 loss 3.943294
Iteration 900 loss 3.931211
Iteration 901 loss 3.924091
Iteration 902 loss 3.916300
Iteration 903 loss 3.911833
Iteration 904 loss 3.904029
Iteration 905 loss 3.895101
Iteration 906 loss 3.887555
Iteration 907 loss 3.880082
Iteration 908 loss 3.868232
Iteration 909 loss 3.865013
Iteration 910 loss 3.856748
Iteration 911 loss 3.848609
Iteration 912 loss 3.840056

Iteration 913 loss 3.833470
Iteration 914 loss 3.825935
Iteration 915 loss 3.815293
Iteration 916 loss 3.810092
Iteration 917 loss 3.802761
Iteration 918 loss 3.795207
Iteration 919 loss 3.788318
Iteration 920 loss 3.779683
Iteration 921 loss 3.772496
Iteration 922 loss 3.765340
Iteration 923 loss 3.756319
Iteration 924 loss 3.750153
Iteration 925 loss 3.740776
Iteration 926 loss 3.736385
Iteration 927 loss 3.728121
Iteration 928 loss 3.719758
Iteration 929 loss 3.708527
Iteration 930 loss 3.707706
Iteration 931 loss 3.699552
Iteration 932 loss 3.690553
Iteration 933 loss 3.684222
Iteration 934 loss 3.676489
Iteration 935 loss 3.669537
Iteration 936 loss 3.661327
Iteration 937 loss 3.655935
Iteration 938 loss 3.648288
Iteration 939 loss 3.641360
Iteration 940 loss 3.631817
Iteration 941 loss 3.624592
Iteration 942 loss 3.619779
Iteration 943 loss 3.606456
Iteration 944 loss 3.605987
Iteration 945 loss 3.598457
Iteration 946 loss 3.591634
Iteration 947 loss 3.585273
Iteration 948 loss 3.578307
Iteration 949 loss 3.570908
Iteration 950 loss 3.556848
Iteration 951 loss 3.557410
Iteration 952 loss 3.550948
Iteration 953 loss 3.543208
Iteration 954 loss 3.536028
Iteration 955 loss 3.529896
Iteration 956 loss 3.521313
Iteration 957 loss 3.517470
Iteration 958 loss 3.507251
Iteration 959 loss 3.501039
Iteration 960 loss 3.496280

```
Iteration 961 loss 3.489640
Iteration 962 loss 3.482220
Iteration 963 loss 3.476064
Iteration 964 loss 3.468522
Iteration 965 loss 3.462299
Iteration 966 loss 3.455085
Iteration 967 loss 3.449714
Iteration 968 loss 3.442050
Iteration 969 loss 3.436740
Iteration 970 loss 3.426219
Iteration 971 loss 3.420296
Iteration 972 loss 3.418482
Iteration 973 loss 3.408700
Iteration 974 loss 3.404890
Iteration 975 loss 3.397902
Iteration 976 loss 3.392709
Iteration 977 loss 3.385087
Iteration 978 loss 3.379405
Iteration 979 loss 3.373260
Iteration 980 loss 3.362862
Iteration 981 loss 3.358855
Iteration 982 loss 3.349339
Iteration 983 loss 3.346805
Iteration 984 loss 3.340693
Iteration 985 loss 3.336657
Iteration 986 loss 3.328797
Iteration 987 loss 3.323195
Iteration 988 loss 3.313572
Iteration 989 loss 3.309023
Iteration 990 loss 3.303983
Iteration 991 loss 3.296108
Iteration 992 loss 3.293139
Iteration 993 loss 3.284874
Iteration 994 loss 3.277524
Iteration 995 loss 3.273007
Iteration 996 loss 3.267989
Iteration 997 loss 3.261840
Iteration 998 loss 3.256057
Iteration 999 loss 3.249850
```

[]: