hw2_backprop

October 10, 2017

In [1]:
```
import numpy as np
import matplotlib.pyplot as plt
```

In [2]:
```
# parameters
m = 2
l = 100
alpha = 0.1
MaxIter = 1000
tanh = lambda x: np.tanh(x)
# derivative for tanh
dtanh = lambda x: 4.0 / np.power(np.exp(x) + np.exp(-x), 2)
```

In [3]:
```
# dataset
x = np.array([-0.5, 0.5]).reshape(2,1)
y = np.array([-0.5, 0.5]).reshape(2,1)
```

In [4]:
```
def trainingAlgorithm(l, alpha, w_scaling, x, y, MaxIter):
    # weight initialization
    w = np.ones((l,))*w_scaling
    # array with activations
    m = x.shape[0]
    h = np.zeros((m,l))
    h[:,:1] = x
    cost = np.zeros(MaxIter)
    for epoch in range(MaxIter):
        # forward propagation to compute activations
        #and pre-nonlinear values
        u = np.zeros((m, l))
        for k in range(1,l):
            u[:, k:k+1] = h[:, k-1:k] * w[k]
            h[:, k:k+1] = tanh(u[:, k:k+1])

        # compute the cost for new iteration
        cost[epoch] = 1.0/(2*m) *np.sum(np.matmul(\
            np.transpose(y-h[:, l-1:l]),y-h[:, l-1:l]))

        # backward propagation
        g = np.zeros((m, l))
```

```
            g[:,l-1:l] = dtanh(u[:, l-1:l]) * (h[:, l-1:l] - y) / m
            for k in range(l-2,0,-1):
                g[:, k:k+1] = dtanh(u[:, k:k+1]) * g[:, k+1:k+2] * w[k+1]

            # gradient descent
            for k in range(1,l):
                w[k] = w[k] - alpha*np.sum(np.matmul\
                                (np.transpose(h[:,k-1:k]), g[:, k:k+1]))

            if epoch == 0:
                print("h[:,k]")
                print(h)
                print("g[:,k]")
                print(g)

        return w, cost
```

1 1 (a)

```
In [5]: w_scaling = 1.0
```

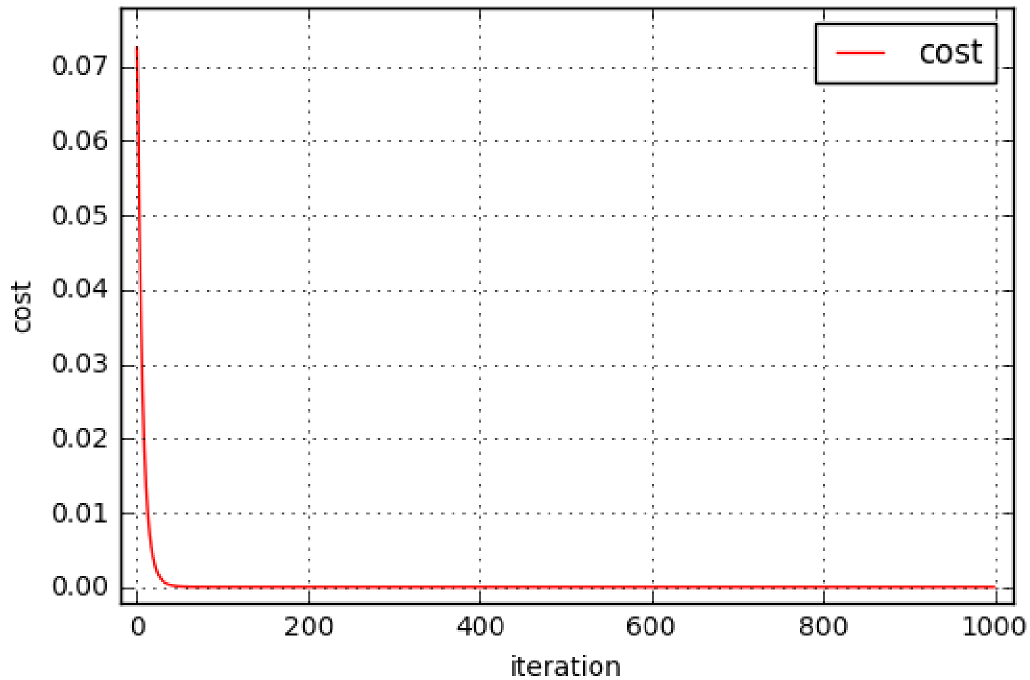
```
In [6]: aw, acost = trainingAlgorithm(l, alpha, w_scaling, x, y, MaxIter)
```

```
h[:,k]
[[-0.5        -0.46211716 -0.43180818 -0.40683132 -0.38577888 -0.36771555
  -0.35199191 -0.33814089 -0.32581665 -0.31475686 -0.30475838 -0.29566112
  -0.28733694 -0.2796819  -0.27261064 -0.26605237 -0.25994775 -0.25424666
  -0.24890641 -0.2438904  -0.23916701 -0.23470881 -0.23049181 -0.22649498
  -0.22269978 -0.21908976 -0.21565033 -0.21236842 -0.20923235 -0.20623162
  -0.20335674 -0.20059914 -0.19795105 -0.1954054  -0.19295574 -0.19059617
  -0.18832129 -0.18612616 -0.18400621 -0.18195725 -0.17997538 -0.17805703
  -0.17619887 -0.17439781 -0.17265097 -0.1709557  -0.1693095  -0.16771005
  -0.16615517 -0.16464282 -0.1631711  -0.16173823 -0.16034251 -0.15898238
  -0.15765634 -0.15636298 -0.155101   -0.15386913 -0.1526662  -0.15149109
  -0.15034275 -0.14922017 -0.1481224  -0.14704854 -0.14599774 -0.14496918
  -0.14396208 -0.14297572 -0.14200938 -0.1410624  -0.14013414 -0.13922399
  -0.13833137 -0.13745572 -0.13659651 -0.13575324 -0.13492541 -0.13411256
  -0.13331425 -0.13253004 -0.13175952 -0.13100231 -0.13025801 -0.12952628
  -0.12880675 -0.12809909 -0.12740299 -0.12671812 -0.12604419 -0.12538091
  -0.124728   -0.1240852  -0.12345225 -0.12282889 -0.12221489 -0.12161002
  -0.12101405 -0.12042676 -0.11984795 -0.11927742]
 [ 0.5         0.46211716  0.43180818  0.40683132  0.38577888  0.36771555
   0.35199191  0.33814089  0.32581665  0.31475686  0.30475838  0.29566112
   0.28733694  0.2796819   0.27261064  0.26605237  0.25994775  0.25424666
   0.24890641  0.2438904   0.23916701  0.23470881  0.23049181  0.22649498
   0.22269978  0.21908976  0.21565033  0.21236842  0.20923235  0.20623162
   0.20335674  0.20059914  0.19795105  0.1954054   0.19295574  0.19059617
   0.18832129  0.18612616  0.18400621  0.18195725  0.17997538  0.17805703
```

```
   0.17619887  0.17439781  0.17265097  0.1709557   0.1693095   0.16771005
   0.16615517  0.16464282  0.1631711   0.16173823  0.16034251  0.15898238
   0.15765634  0.15636298  0.155101    0.15386913  0.1526662   0.15149109
   0.15034275  0.14922017  0.1481224   0.14704854  0.14599774  0.14496918
   0.14396208  0.14297572  0.14200938  0.1410624   0.14013414  0.13922399
   0.13833137  0.13745572  0.13659651  0.13575324  0.13492541  0.13411256
   0.13331425  0.13253004  0.13175952  0.13100231  0.13025801  0.12952628
   0.12880675  0.12809909  0.12740299  0.12671812  0.12604419  0.12538091
   0.124728    0.1240852   0.12345225  0.12282889  0.12221489  0.12161002
   0.12101405  0.12042676  0.11984795  0.11927742]]
g[:,k]
[[ 0.          0.00252209  0.00320694  0.00394195  0.0047238   0.00554974
   0.00641747  0.00732503  0.0082707   0.00925296  0.01027048  0.01132204
   0.01240657  0.01352306  0.01467063  0.01584843  0.0170557   0.01829172
   0.01955584  0.02084743  0.02216591  0.02351075  0.02488142  0.02627745
   0.02769838  0.02914377  0.03061321  0.03210632  0.03362271  0.03516204
   0.03672397  0.03830816  0.03991431  0.04154213  0.04319131  0.0448616
   0.04655271  0.0482644   0.04999643  0.05174855  0.05352053  0.05531215
   0.0571232   0.05895347  0.06080276  0.06267088  0.06455763  0.06646284
   0.06838632  0.0703279   0.07228741  0.07426469  0.07625959  0.07827194
   0.08030159  0.0823484   0.08441223  0.08649293  0.08859037  0.09070442
   0.09283494  0.0949818   0.0971449   0.09932409  0.10151927  0.10373032
   0.10595712  0.10819956  0.11045754  0.11273095  0.11501968  0.11732363
   0.11964271  0.1219768   0.12432582  0.12668968  0.12906827  0.13146151
   0.1338693   0.13629157  0.13872822  0.14117917  0.14364434  0.14612364
   0.148617    0.15112433  0.15364556  0.15618061  0.1587294   0.16129187
   0.16386793  0.16645752  0.16906057  0.17167701  0.17430676  0.17694977
   0.17960596  0.18227528  0.18495764  0.187653  ]
 [ 0.         -0.00252209 -0.00320694 -0.00394195 -0.0047238  -0.00554974
  -0.00641747 -0.00732503 -0.0082707  -0.00925296 -0.01027048 -0.01132204
  -0.01240657 -0.01352306 -0.01467063 -0.01584843 -0.0170557  -0.01829172
  -0.01955584 -0.02084743 -0.02216591 -0.02351075 -0.02488142 -0.02627745
  -0.02769838 -0.02914377 -0.03061321 -0.03210632 -0.03362271 -0.03516204
  -0.03672397 -0.03830816 -0.03991431 -0.04154213 -0.04319131 -0.0448616
  -0.04655271 -0.0482644  -0.04999643 -0.05174855 -0.05352053 -0.05531215
  -0.0571232  -0.05895347 -0.06080276 -0.06267088 -0.06455763 -0.06646284
  -0.06838632 -0.0703279  -0.07228741 -0.07426469 -0.07625959 -0.07827194
  -0.08030159 -0.0823484  -0.08441223 -0.08649293 -0.08859037 -0.09070442
  -0.09283494 -0.0949818  -0.0971449  -0.09932409 -0.10151927 -0.10373032
  -0.10595712 -0.10819956 -0.11045754 -0.11273095 -0.11501968 -0.11732363
  -0.11964271 -0.1219768  -0.12432582 -0.12668968 -0.12906827 -0.13146151
  -0.1338693  -0.13629157 -0.13872822 -0.14117917 -0.14364434 -0.14612364
  -0.148617   -0.15112433 -0.15364556 -0.15618061 -0.1587294  -0.16129187
  -0.16386793 -0.16645752 -0.16906057 -0.17167701 -0.17430676 -0.17694977
  -0.17960596 -0.18227528 -0.18495764 -0.187653  ]]
```

```
In [7]: plt.plot(acost, 'r-', label='cost')
```

```
#plt.title("Cost for 1 (a) case")
plt.xlabel('iteration')
plt.axis([-20,1020,-0.002,0.078])
plt.ylabel('cost')
plt.legend(loc='best')
plt.grid(True)
plt.show()
```



2 1 (b)

```
In [8]: w_scaling = 5.0
```

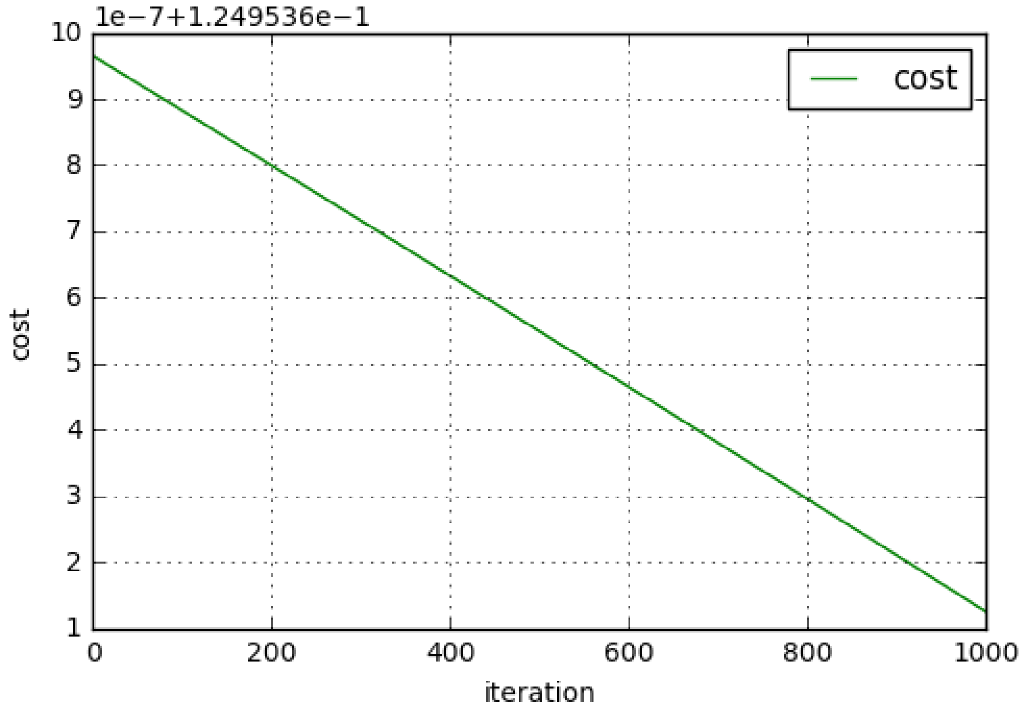
```
In [9]: bw, bcost = trainingAlgorithm(l, alpha, w_scaling, x, y, MaxIter)
```

```
h[:,k]
[[-0.5         -0.9866143   -0.9998962   -0.99990911 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
```

```
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912 -0.99990912
  -0.99990912 -0.99990912 -0.99990912 -0.99990912]
 [ 0.5         0.9866143   0.9998962   0.99990911  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912  0.99990912  0.99990912
   0.99990912  0.99990912  0.99990912  0.99990912]]
g[:,k]
[[  0.00000000e+000  -6.42041275e-301  -4.82878913e-300  -4.65226859e-297
   -5.11880141e-294  -5.63284554e-291  -6.19851213e-288  -6.82098459e-285
   -7.50596753e-282  -8.25973843e-279  -9.08920517e-276  -1.00019694e-272
   -1.10063960e-269  -1.21116900e-266  -1.33279808e-263  -1.46664150e-260
   -1.61392587e-257  -1.77600096e-254  -1.95435210e-251  -2.15061377e-248
   -2.36658461e-245  -2.60424385e-242  -2.86576953e-239  -3.15355837e-236
   -3.47024779e-233  -3.81874008e-230  -4.20222897e-227  -4.62422891e-224
   -5.08860731e-221  -5.59961993e-218  -6.16194991e-215  -6.78075068e-212
   -7.46169321e-209  -8.21101796e-206  -9.03559205e-203  -9.94297225e-200
   -1.09414742e-196  -1.20402486e-193  -1.32493650e-190  -1.45799045e-187
   -1.60440605e-184  -1.76552513e-181  -1.94282426e-178  -2.13792827e-175
   -2.35262519e-172  -2.58888260e-169  -2.84886565e-166  -3.13495695e-163
   -3.44977837e-160  -3.79621505e-157  -4.17744191e-154  -4.59695267e-151
   -5.05859192e-148  -5.56659030e-145  -6.12560335e-142  -6.74075409e-139
   -7.41768005e-136  -8.16258487e-133  -8.98229517e-130  -9.88432315e-127
   -1.08769354e-123  -1.19692286e-120  -1.31712130e-117  -1.44939042e-114
   -1.59494239e-111  -1.75511110e-108  -1.93136441e-105  -2.12531760e-102
   -2.33874812e-099  -2.57361195e-096  -2.83206148e-093  -3.11646525e-090
   -3.42942968e-087  -3.77382290e-084  -4.15280107e-081  -4.56983733e-078
   -5.02875356e-075  -5.53375550e-072  -6.08947118e-069  -6.70099343e-066
   -7.37392651e-063  -8.11443746e-060  -8.92931267e-057  -9.82601999e-054
   -1.08127772e-050  -1.18986275e-047  -1.30935220e-044  -1.44084112e-041
```

```
-1.58553454e-038  -1.74475849e-035  -1.91997217e-032  -2.11278131e-029
-2.32495290e-026  -2.55843137e-023  -2.81535642e-020  -3.09808263e-017
-3.40920102e-014  -3.75156282e-011  -4.12830557e-008  -4.54288192e-005]
[  0.00000000e+000   6.42041275e-301   4.82878913e-300   4.65226859e-297
   5.11880141e-294   5.63284554e-291   6.19851213e-288   6.82098459e-285
   7.50596753e-282   8.25973843e-279   9.08920517e-276   1.00019694e-272
   1.10063960e-269   1.21116900e-266   1.33279808e-263   1.46664150e-260
   1.61392587e-257   1.77600096e-254   1.95435210e-251   2.15061377e-248
   2.36658461e-245   2.60424385e-242   2.86576953e-239   3.15355837e-236
   3.47024779e-233   3.81874008e-230   4.20222897e-227   4.62422891e-224
   5.08860731e-221   5.59961993e-218   6.16194991e-215   6.78075068e-212
   7.46169321e-209   8.21101796e-206   9.03559205e-203   9.94297225e-200
   1.09414742e-196   1.20402486e-193   1.32493650e-190   1.45799045e-187
   1.60440605e-184   1.76552513e-181   1.94282426e-178   2.13792827e-175
   2.35262519e-172   2.58888260e-169   2.84886565e-166   3.13495695e-163
   3.44977837e-160   3.79621505e-157   4.17744191e-154   4.59695267e-151
   5.05859192e-148   5.56659030e-145   6.12560335e-142   6.74075409e-139
   7.41768005e-136   8.16258487e-133   8.98229517e-130   9.88432315e-127
   1.08769354e-123   1.19692286e-120   1.31712130e-117   1.44939042e-114
   1.59494239e-111   1.75511110e-108   1.93136441e-105   2.12531760e-102
   2.33874812e-099   2.57361195e-096   2.83206148e-093   3.11646525e-090
   3.42942968e-087   3.77382290e-084   4.15280107e-081   4.56983733e-078
   5.02875356e-075   5.53375550e-072   6.08947118e-069   6.70099343e-066
   7.37392651e-063   8.11443746e-060   8.92931267e-057   9.82601999e-054
   1.08127772e-050   1.18986275e-047   1.30935220e-044   1.44084112e-041
   1.58553454e-038   1.74475849e-035   1.91997217e-032   2.11278131e-029
   2.32495290e-026   2.55843137e-023   2.81535642e-020   3.09808263e-017
   3.40920102e-014   3.75156282e-011   4.12830557e-008   4.54288192e-005]]
```

In [16]:
```
plt.plot(bcost, 'g-', label='cost')
#plt.title("Cost for 1 (a) case")
plt.xlabel('iteration')
plt.ylabel('cost')
plt.legend(loc='best')
plt.grid(True)
plt.show()
```



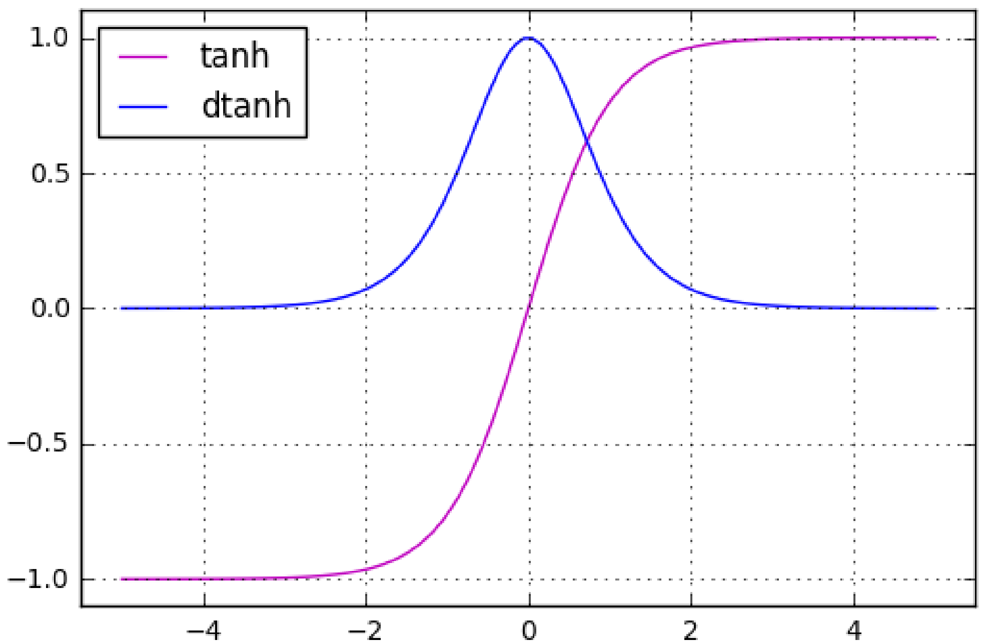
There is a difference between h[:,k] and g[:,k] in 1(b) and 1(a). It is observed in the fact that activation gradients g[:,k] values are very close to zero in 1(b) while in 1(a) they are different and not that close to 0, which results in very tiny change in the values of activations and eventually causes fail in training and non-convergance; also h[:,k] in 1 (a) change from -0.5..-0.1 for x[0]=-0.5 and 0.5..0.1 for x[1]=0.5, while h[:,k] in 1(b) are change from -0.5..-0.999 and 0.5..0.999 respectively, so the values of activations increase till [-0.999,0.999] in the direction of forward propagation in 1(b), while in 1(a) h[:,k] become closer to 0

3 1 (c)

We can see that as k increases h[:,k] gets closer to [-1, 1], that can be explained by the way we compute, h[:,k] = tanh(h[:,k-1]*w[k]) =approx [tanh(-15*), tanh(15*)], and because of tanh function that sends big positive values to +1 and big negative values to -1, the value of h[:,k] turns to [-1, 1] As from 1 (b) it is visible that as k decreases the values of g[:,k] changes significantly and eventually turns to 0. This can be explained as following, g[:,k] = dtanh(h[:,k-1]*w[k])*g[k+1]*w[k+1] =approx [dtanh(-15*)*g[k+1]*w[k+1], dtanh(*15)*g[k+1]*w[k+1]] and since dtanh(-5) and dtanh(5) approaches 0, g[:,k] gets multiplied by previous value of g[:,k-1] which is close to 0 as well as dtanh(h[:,k-1]*w[k]) which is always close to 0. That's why activation gradients vanish And lastly, dw[k] = h[:,k-1]^T* g[:,k] which is approaching 0 due to multiplication by g[:,k] Hence weights don't get updated and training fails

```
In [11]: p_x = np.linspace(-5,5,num=100)
         p_y = tanh(p_x)
         plt.plot(p_x, p_y, 'm-',  label='tanh')
         dy = dtanh(p_x)
```

```
plt.plot(p_x, dy, 'b-',  label='dtanh')
plt.axis([-5.5,5.5,-1.1,1.1])
plt.legend(loc='best')
plt.grid(True)
plt.show()
```



4 1 (d)

```
In [12]: w_scaling = 0.9

In [13]: dw, dcost = trainingAlgorithm(l, alpha, w_scaling, x, y, MaxIter)
```

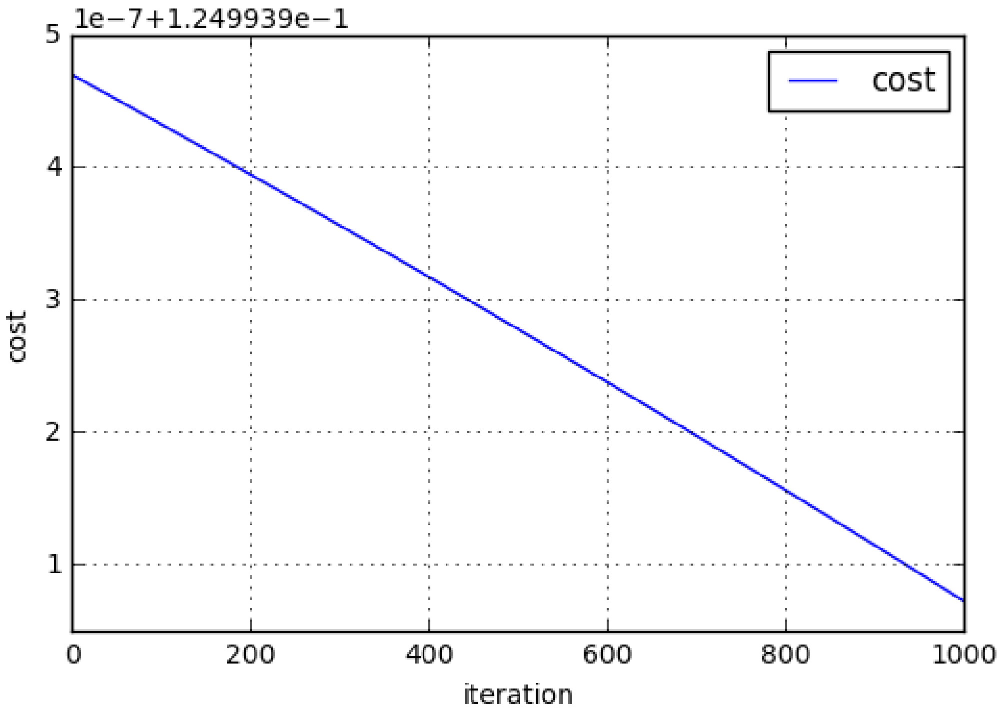
```
h[:,k]
[[ -5.00000000e-01  -4.21899005e-01  -3.62454815e-01  -3.15110593e-01
   -2.76233247e-01  -2.43611525e-01  -2.15803461e-01  -1.91817209e-01
   -1.70940671e-01  -1.52644199e-01  -1.36521989e-01  -1.22255180e-01
   -1.09587776e-01  -9.83104276e-02  -8.82492156e-02  -7.92577058e-02
   -7.12111962e-02  -6.40024697e-02  -5.75385988e-02  -5.17384988e-02
   -4.65310233e-02  -4.18534568e-02  -3.76503057e-02  -3.38723119e-02
   -3.04756405e-02  -2.74212005e-02  -2.46740714e-02  -2.22030146e-02
   -1.99800538e-02  -1.79801105e-02  -1.61806871e-02  -1.45615891e-02
   -1.31046799e-02  -1.17936651e-02  -1.06139000e-02  -9.55221944e-03
   -8.59678570e-03  -7.73695275e-03  -6.96314493e-03  -6.26674840e-03
   -5.64001376e-03  -5.07596879e-03  -4.56834013e-03  -4.11148295e-03
   -3.70031776e-03  -3.33027368e-03  -2.99723733e-03  -2.69750706e-03
```

```
  -2.42775158e-03  -2.18497295e-03  -1.96647312e-03  -1.76982396e-03
  -1.59284021e-03  -1.43355521e-03  -1.29019897e-03  -1.16117855e-03
  -1.04506032e-03  -9.40554010e-04  -8.46498406e-04  -7.61848418e-04
  -6.85663469e-04  -6.17097044e-04  -5.55387282e-04  -4.99848513e-04
  -4.49863631e-04  -4.04877246e-04  -3.64389505e-04  -3.27950543e-04
  -2.95155480e-04  -2.65639926e-04  -2.39075929e-04  -2.15168332e-04
  -1.93651497e-04  -1.74286345e-04  -1.56857709e-04  -1.41171938e-04
  -1.27054743e-04  -1.14349268e-04  -1.02914341e-04  -9.26229068e-05
  -8.33606159e-05  -7.50245542e-05  -6.75220986e-05  -6.07698887e-05
  -5.46928998e-05  -4.92236098e-05  -4.43012488e-05  -3.98711239e-05
  -3.58840115e-05  -3.22956103e-05  -2.90660493e-05  -2.61594443e-05
  -2.35434999e-05  -2.11891499e-05  -1.90702349e-05  -1.71632114e-05
  -1.54468903e-05  -1.39022012e-05  -1.25119811e-05  -1.12607830e-05]
 [  5.00000000e-01   4.21899005e-01   3.62454815e-01   3.15110593e-01
    2.76233247e-01   2.43611525e-01   2.15803461e-01   1.91817209e-01
    1.70940671e-01   1.52644199e-01   1.36521989e-01   1.22255180e-01
    1.09587776e-01   9.83104276e-02   8.82492156e-02   7.92577058e-02
    7.12111962e-02   6.40024697e-02   5.75385988e-02   5.17384988e-02
    4.65310233e-02   4.18534568e-02   3.76503057e-02   3.38723119e-02
    3.04756405e-02   2.74212005e-02   2.46740714e-02   2.22030146e-02
    1.99800538e-02   1.79801105e-02   1.61806871e-02   1.45615891e-02
    1.31046799e-02   1.17936651e-02   1.06139000e-02   9.55221944e-03
    8.59678570e-03   7.73695275e-03   6.96314493e-03   6.26674840e-03
    5.64001376e-03   5.07596879e-03   4.56834013e-03   4.11148295e-03
    3.70031776e-03   3.33027368e-03   2.99723733e-03   2.69750706e-03
    2.42775158e-03   2.18497295e-03   1.96647312e-03   1.76982396e-03
    1.59284021e-03   1.43355521e-03   1.29019897e-03   1.16117855e-03
    1.04506032e-03   9.40554010e-04   8.46498406e-04   7.61848418e-04
    6.85663469e-04   6.17097044e-04   5.55387282e-04   4.99848513e-04
    4.49863631e-04   4.04877246e-04   3.64389505e-04   3.27950543e-04
    2.95155480e-04   2.65639926e-04   2.39075929e-04   2.15168332e-04
    1.93651497e-04   1.74286345e-04   1.56857709e-04   1.41171938e-04
    1.27054743e-04   1.14349268e-04   1.02914341e-04   9.26229068e-05
    8.33606159e-05   7.50245542e-05   6.75220986e-05   6.07698887e-05
    5.46928998e-05   4.92236098e-05   4.43012488e-05   3.98711239e-05
    3.58840115e-05   3.22956103e-05   2.90660493e-05   2.61594443e-05
    2.35434999e-05   2.11891499e-05   1.90702349e-05   1.71632114e-05
    1.54468903e-05   1.39022012e-05   1.25119811e-05   1.12607830e-05]]
g[:,k]
[[  0.00000000e+00   3.62234928e-06   4.89638262e-06   6.26325019e-06
    7.72635264e-06   9.29401422e-06   1.09782011e-05   1.27938241e-05
    1.47583775e-05   1.68917873e-05   1.92163995e-05   2.17570688e-05
    2.45413236e-05   2.75995941e-05   3.09654954e-05   3.46761613e-05
    3.87726294e-05   4.33002765e-05   4.83093085e-05   5.38553077e-05
    5.99998427e-05   6.68111471e-05   7.43648737e-05   8.27449322e-05
    9.20444192e-05   1.02366651e-04   1.13826312e-04   1.26550726e-04
    1.40681270e-04   1.56374947e-04   1.73806130e-04   1.93168497e-04
    2.14677183e-04   2.38571174e-04   2.65115957e-04   2.94606474e-04
```

```
   3.27370398e-04   3.63771771e-04   4.04215053e-04   4.49149614e-04
   4.99074727e-04   5.54545114e-04   6.16177114e-04   6.84655526e-04
   7.60741222e-04   8.45279599e-04   9.39209970e-04   1.04357601e-03
   1.15953734e-03   1.28838241e-03   1.43154285e-03   1.59060931e-03
   1.76734922e-03   1.96372634e-03   2.18192264e-03   2.42436252e-03
   2.69373976e-03   2.99304745e-03   3.32561122e-03   3.69512623e-03
   4.10569819e-03   4.56188902e-03   5.06876751e-03   5.63196564e-03
   6.25774116e-03   6.95304714e-03   7.72560920e-03   8.58401137e-03
   9.53779143e-03   1.05975470e-02   1.17750530e-02   1.30833930e-02
   1.45371040e-02   1.61523384e-02   1.79470432e-02   1.99411596e-02
   2.21568444e-02   2.46187164e-02   2.73541297e-02   3.03934778e-02
   3.37705311e-02   3.75228126e-02   4.16920143e-02   4.63244605e-02
   5.14716230e-02   5.71906924e-02   6.35452139e-02   7.06057934e-02
   7.84508817e-02   8.71676464e-02   9.68529405e-02   1.07614378e-01
   1.19571532e-01   1.32857258e-01   1.47619175e-01   1.64021306e-01
   1.82245895e-01   2.02495439e-01   2.24994933e-01   2.49994370e-01]
 [  0.00000000e+00  -3.62234928e-06  -4.89638262e-06  -6.26325019e-06
  -7.72635264e-06  -9.29401422e-06  -1.09782011e-05  -1.27938241e-05
  -1.47583775e-05  -1.68917873e-05  -1.92163995e-05  -2.17570688e-05
  -2.45413236e-05  -2.75995941e-05  -3.09654954e-05  -3.46761613e-05
  -3.87726294e-05  -4.33002765e-05  -4.83093085e-05  -5.38553077e-05
  -5.99998427e-05  -6.68111471e-05  -7.43648737e-05  -8.27449322e-05
  -9.20444192e-05  -1.02366651e-04  -1.13826312e-04  -1.26550726e-04
  -1.40681270e-04  -1.56374947e-04  -1.73806130e-04  -1.93168497e-04
  -2.14677183e-04  -2.38571174e-04  -2.65115957e-04  -2.94606474e-04
  -3.27370398e-04  -3.63771771e-04  -4.04215053e-04  -4.49149614e-04
  -4.99074727e-04  -5.54545114e-04  -6.16177114e-04  -6.84655526e-04
  -7.60741222e-04  -8.45279599e-04  -9.39209970e-04  -1.04357601e-03
  -1.15953734e-03  -1.28838241e-03  -1.43154285e-03  -1.59060931e-03
  -1.76734922e-03  -1.96372634e-03  -2.18192264e-03  -2.42436252e-03
  -2.69373976e-03  -2.99304745e-03  -3.32561122e-03  -3.69512623e-03
  -4.10569819e-03  -4.56188902e-03  -5.06876751e-03  -5.63196564e-03
  -6.25774116e-03  -6.95304714e-03  -7.72560920e-03  -8.58401137e-03
  -9.53779143e-03  -1.05975470e-02  -1.17750530e-02  -1.30833930e-02
  -1.45371040e-02  -1.61523384e-02  -1.79470432e-02  -1.99411596e-02
  -2.21568444e-02  -2.46187164e-02  -2.73541297e-02  -3.03934778e-02
  -3.37705311e-02  -3.75228126e-02  -4.16920143e-02  -4.63244605e-02
  -5.14716230e-02  -5.71906924e-02  -6.35452139e-02  -7.06057934e-02
  -7.84508817e-02  -8.71676464e-02  -9.68529405e-02  -1.07614378e-01
  -1.19571532e-01  -1.32857258e-01  -1.47619175e-01  -1.64021306e-01
  -1.82245895e-01  -2.02495439e-01  -2.24994933e-01  -2.49994370e-01]]
```

```
In [18]: plt.plot(dcost, 'b-', label='cost')
         #plt.title("Cost for 1 (a) case")
         plt.xlabel('iteration')
         plt.ylabel('cost')
         plt.legend(loc='best')
```

```
plt.grid(True)
plt.show()
```



There is a difference between h[:,k] and g[:,k] in 1(c), 1(b) and 1(a). It is observed in the fact that activation gradients abs(g[:,k]) values are very close to zero (~ absolute values 2e-200..2e-005) in 1(b) while in 1(a) they are different around 3e-03..1e-01 and in 1 (c) around 6e-06..2e-01, so gradients in 1(c) are closer to 0 than in 1(a) but more distant form 0 than in 1(b) h[:,k] in 1 (a) change from -0.5..-0.1 for x[0]=-0.5 and 0.5..0.1 for x[1]=0.5, while h[:,k] in 1(b) are change from -0.5..-0.999 and 0.5..0.999 respectively, and in 1(c) abs(h[:,k]) change in range -5e-01..-1e-05 and 5e-01..1e-05 for x=[-0.5, 0.5], so the values of activations in 1(c) approach 0 much faster than in 1(a) in the direction of forward propagation

5 1 (e)

We can see that as k increases h[:,k] from [-5e-01, 5e-01] gets closer to [-1e-05, 1e-05], that can be explained by the way we compute, h[:,k] = tanh(h[:,k-1]*w[k]) =approx [tanh(h[0,k-1]*0.9), tanh(h[1,k-1]*0.9)], and because of tanh function is monotonically increasing abs(h[:,k-1]) > abs(h[:,k]) (since x = [-0.5, 0.5]) As from 1 (d) it is visible that as k decreases the values of abs(g[:,k]) change from 2e-01 to 6e-06. This can be explained as following

So, since dw[k] = h[:,k-1]^T * g[:,k] and on first half of the network abs(h[:,k]) are around 5e-01..4e-03 and abs(g[:,k]) are around 6e-06..1e-03, their product results in similar small values around 1e-06 and similar for second half of the network where activation values are small and activation gradients are bigger, but still their total product is a small value of order 1e-06, which give us small weight gradients for the whole network

Hence weights get very small gradients and training fails