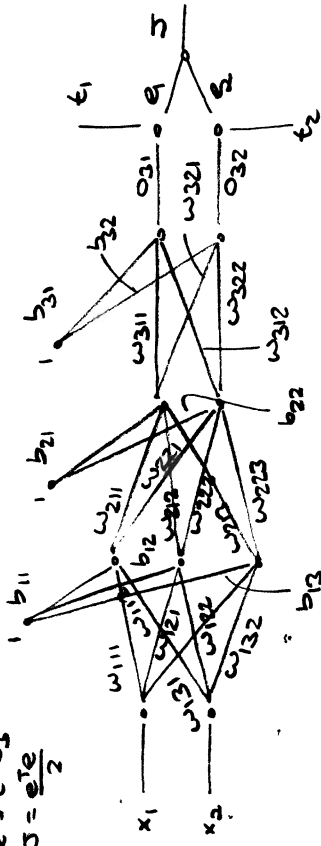


QUESTÃO #1:

- $u_1 = w_1 x_1 + b_1$
- $o_1 = f(u_1)$
- $u_2 = w_{21} o_1 + b_2$
- $o_2 = f(u_2)$
- $u_3 = w_{31} o_1 + w_{32} o_2 + b_3$
- $o_3 = f(u_3)$
- $e = t - o_3$
- $J = \frac{e^2}{2}$



$$u_1 = w_{11}x_1 + w_{12}x_2 + b_1 \quad o_{11} = \tanh(u_{11})$$

$$u_2 = w_{21}x_1 + w_{22}x_2 + b_2 \quad o_{12} = \tanh(u_{12})$$

$$u_3 = w_{31}x_1 + w_{32}x_2 + b_3 \quad o_{13} = \tanh(u_{13})$$

$$u_{21} = w_{21}o_{11} + w_{22}o_{12} + w_{23}o_{13} + b_{21} \quad o_{21} = \tanh(u_{21})$$

$$u_{22} = w_{22}o_{11} + w_{23}o_{12} + w_{24}o_{13} + b_{22} \quad o_{22} = \tanh(u_{22})$$

$$u_{31} = w_{31}o_{21} + w_{32}o_{22} + b_{31} \quad o_{31} = \tanh(u_{31})$$

$$u_{32} = w_{32}o_{21} + w_{33}o_{22} + b_{32} \quad o_{32} = \tanh(u_{32})$$

$$e_1 = t_1 - o_{31}$$

$$e_2 = t_2 - o_{32}$$

$$J = \frac{e_1^2 + e_2^2}{2}$$

$$\frac{\partial J}{\partial o_{31}} = \frac{\partial}{\partial o_{31}} \left( \frac{e_1^2 + e_2^2}{2} \right) = -e_1(1-o_{31}^2)$$

$$\frac{\partial J}{\partial o_{32}} = \frac{\partial}{\partial o_{32}} \left( \frac{e_1^2 + e_2^2}{2} \right) = -e_2(1-o_{32}^2)$$

$$w_{31} \rightarrow u_{31} \rightarrow o_{31} \rightarrow e_1 \rightarrow J$$

$$\frac{\partial J}{\partial w_{31}} = \frac{\partial}{\partial w_{31}} \left( \frac{e_1^2 + e_2^2}{2} \right) = o_{21}(1-o_{31}^2)(-e_1)$$

$$\frac{\partial J}{\partial w_{32}} = \frac{\partial}{\partial w_{32}} \left( \frac{e_1^2 + e_2^2}{2} \right) = o_{22}(1-o_{31}^2)(-e_1)$$

$$\frac{\partial J}{\partial w_{33}} = \frac{\partial}{\partial w_{33}} \left( \frac{e_1^2 + e_2^2}{2} \right) = o_{21}(1-o_{32}^2)(-e_2)$$

$$\frac{\partial J}{\partial w_{34}} = \frac{\partial}{\partial w_{34}} \left( \frac{e_1^2 + e_2^2}{2} \right) = o_{22}(1-o_{32}^2)(-e_2)$$

$$w_{21} \rightarrow u_{21} \rightarrow o_{21} \rightarrow e_1 \rightarrow J$$

$$w_{22} \rightarrow u_{22} \rightarrow o_{22} \rightarrow e_2 \rightarrow J$$

Handwritten mathematical derivations for the backpropagation algorithm, showing the flow of error gradients from the output layer back to the hidden and input layers. The equations are arranged in a columnar fashion, with arrows indicating the direction of gradient flow.

$$\frac{\partial J}{\partial w_{31}} = \frac{\partial}{\partial w_{31}} \left( \frac{e_1^2 + e_2^2}{2} \right) = -e_1(1-o_{31}^2)$$

$$\frac{\partial J}{\partial w_{32}} = \frac{\partial}{\partial w_{32}} \left( \frac{e_1^2 + e_2^2}{2} \right) = -e_2(1-o_{32}^2)$$

$$\frac{\partial J}{\partial w_{21}} = \frac{\partial}{\partial w_{21}} \left( \frac{e_1^2 + e_2^2}{2} \right) = o_{11}(1-o_{21}^2)(-e_1) + o_{12}(1-o_{21}^2)(-e_2)$$

$$\frac{\partial J}{\partial w_{22}} = \frac{\partial}{\partial w_{22}} \left( \frac{e_1^2 + e_2^2}{2} \right) = o_{11}(1-o_{22}^2)(-e_1) + o_{12}(1-o_{22}^2)(-e_2)$$

$$\frac{\partial J}{\partial w_{11}} = \frac{\partial}{\partial w_{11}} \left( \frac{e_1^2 + e_2^2}{2} \right) = o_{11}o_{21}(1-o_{31}^2)(-e_1) + o_{11}o_{22}(1-o_{31}^2)(-e_2) + o_{12}o_{21}(1-o_{32}^2)(-e_1) + o_{12}o_{22}(1-o_{32}^2)(-e_2)$$





### Questao 3, Item (a)

#### % Inicializacao We b

```
>> rand('state',0)
>> W1 = round(20*(rand(3,2)-0.5))/10
W1 =
    0.9000    0
   -0.5000    0.8000
    0.2000    0.5000

>> b1 = round(20*(rand(3,1)-0.5))/10
b1 =
   -0.1000
   -1.0000
    0.6000

>> W2 = round(20*(rand(2,3)-0.5))/10
W2 =
   -0.1000    0.6000    0.5000
   -0.2000    0.8000   -0.6000

>> b2 = round(20*(rand(2,1)-0.5))/10
b2 =
   -0.2000
    0.9000

>> W3 = round(20*(rand(2,2)-0.5))/10
W3 =
    0.8000    0.8000
   -0.2000   -0.9000

>> b3 = round(20*(rand(2,1)-0.5))/10
b3 =
   -0.3000
    0.6000

>> eta = 0.01;
```

#### % Iteracao #1 - Feed-Forward

```
>> x = [-0.5 ; -0.5]
x =
   -0.5000
   -0.5000

>> t = [1 ; -1]
t =
     1
    -1

>> u1 = W1*x+b1
u1 =
   -0.5500
   -1.1500
    0.2500
```

```
>> o1 = tanh(u1)
o1 =
   -0.5005
   -0.8178
    0.2449
```

```
>> u2 = W2*o1+b2
u2 =
   -0.5181
   -0.0013
```

```
>> o2 = tanh(u2)
o2 =
   -0.4763
   -0.0013
```

```
>> u3 = W3*o2+b3
u3 =
   -0.6820
    0.6964
```

```
>> o3 = tanh(u3)
o3 =
   -0.5928
    0.6021
```

```
>> e = t-o3
e =
    1.5928
   -1.6021
```

#### % Iteracao #1 - Backpropagation

```
>> db3 = (eye(2)-diag(o3)^2)*e
db3 =
    1.0330
   -1.0213

>> dW3 = kron(o2', db3)
dW3 =
   -0.4920   -0.0013
    0.4864    0.0013
```

```
>> db2 = (eye(2)-diag(o2)^2)*W3'*db3
db2 =
    0.7969
    1.7456

>> dW2 = kron(o1', db2)
dW2 =
   -0.3989   -0.6517    0.1952
   -0.8737   -1.4275    0.4275

>> db1 = (eye(3)-diag(o1)^2)*W2'*db2
db1 =
    0.2019
    0.6210
   -0.6100

>> dW1 = kron(x', db1)
dW1 =
   -0.1010   -0.1010
   -0.3105   -0.3105
    0.3050    0.3050
```

#### % Iteracao #1 - Atualizacoes

```
>> b1 = b1 + eta*db1
b1 =
   -0.0980
   -0.9938
    0.5939

>> W1 = W1 + eta*dW1
W1 =
    0.8990   -0.0010
   -0.5031    0.7969
    0.2030    0.5030

>> b2 = b2 + eta*db2
b2 =
   -0.1920
    0.9175

>> W2 = W2 + eta*dW2
W2 =
   -0.1040    0.5935    0.5020
    0.1913    0.7857   -0.5957

>> b3 = b3 + eta*db3
b3 =
   -0.2897
    0.5898

>> W3 = W3 + eta*dW3
W3 =
    0.7951    0.8000
   -0.1951   -0.9000
```

#### % Iteracao #2 - Feed-Forward

```
>> x = [-0.5 ; 0.5]
x =
   -0.5000
    0.5000

>> t = [-1 ; 1]
t =
    -1
     1

>> u1 = W1*x+b1
u1 =
   -0.5480
   -0.3438
    0.7439
```

```
>> o1 = tanh(u1)
o1 =
   -0.4990
   -0.3309
    0.6315
```

```
>> u2 = W2*o1+b2
u2 =
   -0.0195
    0.1859
```

```
>> o2 = tanh(u2)
o2 =
   -0.0195
    0.1837
```

```
>> u3 = W3*o2+b3
u3 =
   -0.1582
    0.4282
```

```
>> o3 = tanh(u3)
o3 =
   -0.1569
    0.4038
```

```
>> e = t-o3
e =
   -0.8431
    0.5962
```

#### % Iteracao #2 - Backpropagation

```
>> db3 = (eye(2)-diag(o3)^2)*e
db3 =
   -0.8224
    0.4989

>> dW3 = kron(o2', db3)
dW3 =
    0.0160   -0.1511
   -0.0097    0.0917

>> db2 = (eye(2)-diag(o2)^2)*W3'*db3
db2 =
   -0.7509
   -1.0695

>> dW2 = kron(o1', db2)
dW2 =
    0.3747    0.2484   -0.4742
    0.5337    0.3539   -0.6754

>> db1 = (eye(3)-diag(o1)^2)*W2'*db2
db1 =
   -0.0950
   -1.1452
    0.1565

>> dW1 = kron(x', db1)
dW1 =
    0.0475   -0.0475
    0.5726   -0.5726
   -0.0782    0.0782
```

#### % Iteracao #2 - Atualizacoes

```
>> b3 = b3 + eta*db3
b3 =
   -0.2979
    0.5948

>> W3 = W3 + eta*dW3
W3 =
    0.7952    0.7985
   -0.1952   -0.8991

>> b2 = b2 + eta*db2
b2 =
   -0.1995
    0.9068

>> W2 = W2 + eta*dW2
W2 =
   -0.1002    0.5960    0.4972
    0.1966    0.7893   -0.6025

>> b1 = b1 + eta*db1
b1 =
   -0.0989
   -1.0052
    0.5955

>> W1 = W1 + eta*dW1
W1 =
    0.8995   -0.0015
   -0.4974    0.7912
    0.2023    0.5038
```

#### % Iteracao #3 - Feed-Forward

```
>> x = [0.5 ; -0.5]
x =
    0.5000
   -0.5000

>> t = [-1 ; 1]
t =
    -1
     1

>> u1 = W1*x+b1
u1 =
    0.3515
   -1.6495
    0.4447
```

```
>> o1 = tanh(u1)
o1 =
    0.3377
   -0.9288
    0.4175
```

```
>> u2 = W2*o1+b2
u2 =
   -0.5793
   -0.0114
```

```
>> o2 = tanh(u2)
o2 =
   -0.5222
   -0.0114
```

```
>> u3 = W3*o2+b3
u3 =
   -0.7223
    0.7070
```

```
>> o3 = tanh(u3)
o3 =
    -0.6183
    0.6088

>> e = t-o3
e =
    -0.3817
    0.3912
```

### % Iteracao #3 - Backpropagation

```
>> db3 = (eye(2)-diag(o3)^2)*e
db3 =
    -0.2357
    0.2462

>> dW3 = kron(o2', db3)
dW3 =
    0.1231    0.0027
   -0.1286   -0.0028

>> db2 = (eye(2)-diag(o2)^2)*W3'*db3
db2 =
    -0.1713
    -0.4095

>> dW2 = kron(o1', db2)
dW2 =
    -0.0579    0.1591   -0.0715
   -0.1383    0.3804   -0.1710

>> db1 = (eye(3)-diag(o1)^2)*W2'*db2
db1 =
    -0.0561
   -0.0584
    0.1334

>> dW1 = kron(x', db1)
dW1 =
    -0.0281    0.0281
   -0.0292    0.0292
    0.0667   -0.0667
```

### % Iteracao #3 - Atualizacoes

```
>> b1 = b1 + eta*db1
b1 =
    -0.0995
   -1.0058
    0.5968

>> W1 = W1 + eta*dW1
W1 =
    0.8992   -0.0012
   -0.4977    0.7915
    0.2029    0.5032

>> b2 = b2 + eta*db2
b2 =
    -0.2013
    0.9027

>> W2 = W2 + eta*dW2
W2 =
   -0.1008    0.5976    0.4965
    0.1952    0.7931   -0.6042

>> b3 = b3 + eta*db3
b3 =
   -0.3003
    0.5972

>> W3 = W3 + eta*dW3
W3 =
    0.7965    0.7985
   -0.1965   -0.8991
```

## Questao 3, Item (b)

### % Inicializacao W e b

```
>> rand('state',0)
>> W1 = round(20*(rand(3,2)-0.5))/10
W1 =
    0.9000    0
   -0.5000    0.8000
    0.2000    0.5000

>> b1 = round(20*(rand(3,1)-0.5))/10
b1 =
   -0.1000
   -1.0000
    0.6000

>> W2 = round(20*(rand(2,3)-0.5))/10
W2 =
   -0.1000    0.6000    0.5000
    0.2000    0.8000   -0.6000
```

```
>> b2 = round(20*(rand(2,1)-0.5))/10
b2 =
   -0.2000
    0.9000
```

```
>> W3 = round(20*(rand(2,2)-0.5))/10
W3 =
    0.8000    0.8000
   -0.2000   -0.9000
```

```
>> b3 = round(20*(rand(2,1)-0.5))/10
b3 =
   -0.3000
    0.6000
```

```
>> Del tab3 = zeros(size(b3))
Del tab3 =
    0
    0
```

```
>> Del taW3 = zeros(size(W3))
Del taW3 =
    0    0
    0    0
```

```
>> Del tab2 = zeros(size(b2))
Del tab2 =
    0
    0
```

```
>> Del taW2 = zeros(size(W2))
Del taW2 =
    0    0    0
    0    0    0
```

```
>> Del tab1 = zeros(size(b1))
Del tab1 =
    0
    0
    0
```

```
>> Del taW1 = zeros(size(W1))
Del taW1 =
    0    0
    0    0
    0    0
```

```
>> eta = 0.01;
```

### % Iteracao #1 - Feed-Forward

```
>> x = [-0.5 ; -0.5]
x =
   -0.5000
   -0.5000
```

```
>> t = [1 ; -1]
t =
    1
   -1
```

```
>> u1 = W1*x+b1
u1 =
   -0.5500
   -1.1500
    0.2500
```

```
>> o1 = tanh(u1)
o1 =
   -0.5005
   -0.8178
    0.2449
```

```
>> u2 = W2*o1+b2
u2 =
   -0.5181
   -0.0013
```

```
>> o2 = tanh(u2)
o2 =
   -0.4763
   -0.0013
```

```
>> u3 = W3*o2+b3
u3 =
   -0.6820
    0.6964
```

```
>> o3 = tanh(u3)
o3 =
   -0.5928
    0.6021
```

```
>> e = t-o3
e =
    1.5928
   -1.6021
```

### % Iteracao #1 - Backpropagation

```
>> db3 = (eye(2)-diag(o3)^2)*e
db3 =
    1.0330
   -1.0213
```

```
>> dW3 = kron(o2', db3)
dW3 =
   -0.4920   -0.0013
```

```
0.4864    0.0013
```

```
>> db2 = (eye(2)-diag(o2)^2)*W3'*db3
db2 =
    0.7969
    1.7456
```

```
>> dW2 = kron(o1', db2)
dW2 =
   -0.3989   -0.6517    0.1952
   -0.8737   -1.4275    0.4275
```

```
>> db1 = (eye(3)-diag(o1)^2)*W2'*db2
db1 =
    0.2019
    0.6210
   -0.6100
```

```
>> dW1 = kron(x', db1)
dW1 =
   -0.1010   -0.1010
   -0.3105   -0.3105
    0.3050    0.3050
```

```
>> Del tab3 = Del tab3 + db3
```

```
Del tab3 =
    1.0330
   -1.0213
```

```
>> Del taW3 = Del taW3 + dW3
```

```
Del taW3 =
   -0.4920   -0.0013
    0.4864    0.0013
```

```
>> Del tab2 = Del tab2 + db2
```

```
Del tab2 =
    0.7969
    1.7456
```

```
>> Del taW2 = Del taW2 + dW2
```

```
Del taW2 =
   -0.3989   -0.6517    0.1952
   -0.8737   -1.4275    0.4275
```

```
>> Del tab1 = Del tab1 + db1
```

```
Del tab1 =
    0.2019
    0.6210
   -0.6100
```

```
>> Del taW1 = Del taW1 + dW1
```

```
Del taW1 =
   -0.1010   -0.1010
   -0.3105   -0.3105
    0.3050    0.3050
```

### % Iteracao #2 - Feed-Forward

```
>> x = [-0.5 ; 0.5]
x =
   -0.5000
    0.5000
```

```
>> t = [-1 ; 1]
t =
   -1
    1
```

```
>> u1 = W1*x+b1
u1 =
   -0.5500
   -0.3500
    0.7500
```

```
>> o1 = tanh(u1)
o1 =
   -0.5005
   -0.3364
    0.6351
```

```
>> u2 = W2*o1+b2
u2 =
   -0.0342
    0.1497
```

```
>> o2 = tanh(u2)
o2 =
   -0.0342
    0.1486
```

```
>> u3 = W3*o2+b3
u3 =
   -0.2085
    0.4731
```

```
>> o3 = tanh(u3)
o3 =
   -0.2055
    0.4407
```

```
>> e = t-o3
e =
   -0.7945
    0.5593
```

### % Iteracao #2 - Backpropagation

```
>> db3 = (eye(2)-diag(o3)^2)*e
db3 =
```

```

-0.7609
0.4507
>> dW3 = kron(o2', db3)
dW3 =
    0.0260    -0.1131
   -0.0154     0.0670
>> db2 = (eye(2) -
di ag(o2)^2)*W3'*db3
db2 =
   -0.6981
   -0.9920
>> dW2 = kron(o1', db2)
dW2 =
    0.3494    0.2348   -0.4434
    0.4965    0.3337   -0.6300
>> db1 = (eye(3) -
di ag(o1)^2)*W2'*db2
db1 =
   -0.0964
   -1.0752
    0.1468
>> dW1 = kron(x', db1)
dW1 =
    0.0482   -0.0482
    0.5376   -0.5376
   -0.0734    0.0734
>> Del tab3 = Del tab3 + db3
Del tab3 =
    0.2721
   -0.5707
>> Del taW3 = Del taW3 + dW3
Del taW3 =
   -0.4660   -0.1144
    0.4710    0.0683
>> Del tab2 = Del tab2 + db2
Del tab2 =
    0.0988
    0.7537
>> Del taW2 = Del taW2 + dW2
Del taW2 =
   -0.0495   -0.4169   -0.2482
   -0.3772   -1.0938   -0.2025
>> Del tab1 = Del tab1 + db1
Del tab1 =
    0.1056
   -0.4542
   -0.4632
>> Del taW1 = Del taW1 + dW1
Del taW1 =
   -0.0528   -0.1492
    0.2271   -0.8481
    0.2316    0.3784

```

### % Iteracao #3 - Feed-Forward

```

>> x = [0.5 ; -0.5]
x =
    0.5000
   -0.5000
>> t = [-1 ; 1]
t =
   -1
    1
>> u1 = W1*x+b1
u1 =
    0.3500
   -1.6500
    0.4500
>> o1 = tanh(u1)
o1 =
    0.3364
   -0.9289
    0.4219
>> u2 = W2*o1+b2
u2 =
   -0.5800
   -0.0290
>> o2 = tanh(u2)
o2 =
   -0.5227
   -0.0289
>> u3 = W3*o2+b3
u3 =
   -0.7413
    0.7306
>> o3 = tanh(u3)
o3 =
   -0.6299
    0.6234
>> e = t-o3
e =
   -0.3701
    0.3766

```

### % Iteracao #3 - Backpropagation

```

>> db3 = (eye(2)-di ag(o3)^2)*e
db3 =
   -0.2232
    0.2302
>> dW3 = kron(o2', db3)
dW3 =
    0.1167    0.0065
   -0.1203   -0.0067
>> db2 = (eye(2) -
di ag(o2)^2)*W3'*db3
db2 =
   -0.1633
   -0.3855
>> dW2 = kron(o1', db2)
dW2 =
   -0.0549    0.1516   -0.0689
   -0.1297    0.3580   -0.1626
>> db1 = (eye(3) -
di ag(o1)^2)*W2'*db2
db1 =
   -0.0539
   -0.0558
    0.1230
>> dW1 = kron(x', db1)
dW1 =
   -0.0269    0.0269
   -0.0279    0.0279
    0.0615   -0.0615
>> Del tab3 = Del tab3 + db3
Del tab3 =
    0.0489
   -0.3404
>> Del taW3 = Del taW3 + dW3
Del taW3 =
   -0.3493   -0.1079
    0.3507    0.0616
>> Del tab2 = Del tab2 + db2
Del tab2 =
   -0.0644
    0.3682
>> Del taW2 = Del taW2 + dW2
Del taW2 =
   -0.1044   -0.2652   -0.3171
   -0.5069   -0.7358   -0.3651
>> Del tab1 = Del tab1 + db1
Del tab1 =
    0.0517
   -0.5100
   -0.3401
>> Del taW1 = Del taW1 + dW1
Del taW1 =
   -0.0797   -0.1222
    0.1992   -0.8203
    0.2931    0.3169

```

### % Iteracao #4 - Feed-Forward

```

>> x = [0.5 ; 0.5]
x =
    0.5000
    0.5000
>> t = [1 ; -1]
t =
    1
   -1
>> u1 = W1*x+b1
u1 =
    0.3500
   -0.8500
    0.9500
>> o1 = tanh(u1)
o1 =
    0.3364
   -0.6911
    0.7398
>> u2 = W2*o1+b2
u2 =
   -0.2784
   -0.0295
>> o2 = tanh(u2)
o2 =
   -0.2714
   -0.0294
>> u3 = W3*o2+b3
u3 =
   -0.5407
    0.6808
>> o3 = tanh(u3)
o3 =
   -0.4935
    0.5920

```

```

>> e = t-o3
e =
    1.4935
   -1.5920

```

### % Iteracao #4 - Backpropagation

```

>> db3 = (eye(2)-di ag(o3)^2)*e
db3 =
    1.1298
   -1.0340
>> dW3 = kron(o2', db3)
dW3 =
   -0.3066   -0.0333
    0.2806    0.0304
>> db2 = (eye(2) -
di ag(o2)^2)*W3'*db3
db2 =
    1.0288
    1.8328
>> dW2 = kron(o1', db2)
dW2 =
    0.3461   -0.7110    0.7611
    0.6165   -1.2666    1.3559
>> db1 = (eye(3) -
di ag(o1)^2)*W2'*db2
db1 =
    0.2339
    1.0885
   -0.2650
>> dW1 = kron(x', db1)
dW1 =
    0.1169    0.1169
    0.5443    0.5443
   -0.1325   -0.1325
>> Del tab3 = Del tab3 + db3
Del tab3 =
    1.1786
   -1.3745
>> Del taW3 = Del taW3 + dW3
Del taW3 =
   -0.6559   -0.1412
    0.6313    0.0920
>> Del tab2 = Del tab2 + db2
Del tab2 =
    0.9644
    2.2011
>> Del taW2 = Del taW2 + dW2
Del taW2 =
    0.2417   -0.9762    0.4440
    0.1096   -2.0024    0.9908
>> Del tab1 = Del tab1 + db1
Del tab1 =
    0.2855
    0.5785
   -0.6051
>> Del taW1 = Del taW1 + dW1
Del taW1 =
    0.0372   -0.0053
    0.7435   -0.2760
    0.1606    0.1844

```

### % Fim da Epoca - Atualizacoes

```

>> b3 = b3 + eta*Del tab3
b3 =
   -0.2882
    0.5863
>> W3 = W3 + eta*Del taW3
W3 =
    0.7934    0.7986
   -0.1937   -0.8991
>> b2 = b2 + eta*Del tab2
b2 =
   -0.1904
    0.9220
>> W2 = W2 + eta*Del taW2
W2 =
   -0.0976    0.5902    0.5044
    0.2011    0.7800   -0.5901
>> b1 = b1 + eta*Del tab1
b1 =
   -0.0971
   -0.9942
    0.5939
>> W1 = W1 + eta*Del taW1
W1 =
    0.9004   -0.0001
   -0.4926    0.7972
    0.2016    0.5018

```

## Questao #5

```
% 070709 gabriel@pads.ufrj.br

close all; clear all;

% [A] Data

X = [-0.5 -0.5 0.5 0.5 ; -0.5 0.5 -0.5 0.5];
t = [1 -1 -1 1 ; -1 1 1 -1];

% [B] Network Init

% [B1] Parameters

K = 3; % Number of Layers
Delta = 1e-5; % Stop Criterion
N = size(X,2); % Number of Input Vectors
E = 1; % Number of Feed-Forward Iterations per Epoch

eta = 0.02; alpha = 1;

% [B2] Layers

rand('state',0);
L(1).W = round(20*(rand(3,2)-0.5))/10;
L(1).b = round(20*(rand(3,1)-0.5))/10;
L(2).W = round(20*(rand(2,3)-0.5))/10;
L(2).b = round(20*(rand(2,1)-0.5))/10;
L(3).W = round(20*(rand(2,2)-0.5))/10;
L(3).b = round(20*(rand(2,1)-0.5))/10;

% [C] Batch Error Backpropagation Training

n=1; i=1; fim=0;
while not(fim),

    for k=1:K,
        L(k).db = zeros(size(L(k).b));
        L(k).dW = zeros(size(L(k).W));
    end;
    J(i) = 0;
    for ep=1:E,

        % [C1] Feed-Forward

        L(1).x = X(:,n);
        for k = 1:K,
            L(k).u = L(k).W*L(k).x + L(k).b;
            L(k).o = tanh(L(k).u);
            L(k+1).x = L(k).o;
        end;
        e = t(:,n) - L(K).o;
```

```

J(i) = J(i) + (e'*e)/2;

% [C2] Error Backpropagation

L(K+1).alpha = e; L(K+1).W = eye(length(e));
for k = fliplr(1:K),
    L(k).M = eye(length(L(k).o)) - diag(L(k).o)^2;
    L(k).alpha = L(k).M*L(k+1).W'*L(k+1).alpha;
    L(k).db = L(k).db + L(k).alpha;
    L(k).dW = L(k).dW + kron(L(k).x',L(k).alpha);
end;
n = n+1; if n>N, n=1; end;

end;

% [C3] Updates

for k = 1:K,
    L(k).b = L(k).b + eta*L(k).db;
    L(k).W = L(k).W + eta*L(k).dW;
end;
J(i) = J(i)/E;

% [C4] Stop criterion

if (i>1),
    if (abs(J(i)-J(i-1))/J(i) < Delta)|(i>1000),
        fim = 1;
    end;
end;
if not(fim)
    i = i+1; if n>N, n=1; end; eta = eta*alpha;
end;

end;

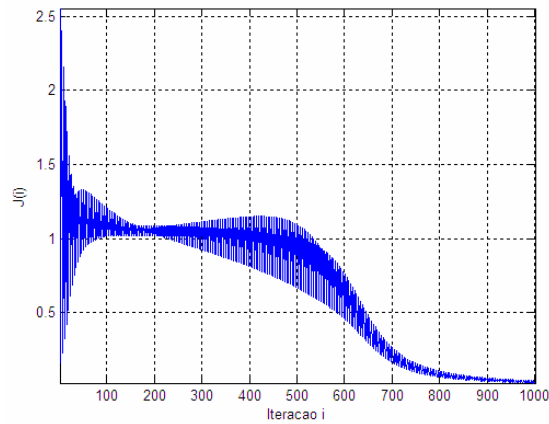
% [D] Test

for n = 1:size(X,2),
    L(1).x = X(:,n);
    for k = 1:K,
        L(k).u = L(k).W*L(k).x + L(k).b;
        L(k).o = tanh(L(k).u);
        L(k+1).x = L(k).o;
    end;
    [X(:,n) L(K).o]
end;
figure; plot(J); axis tight;

```



Resultados em modo sequencial:



Entradas e Saidas

```
-0.5000  0.7877
-0.5000 -0.8275

-0.5000 -0.8404
 0.5000  0.8475

 0.5000 -0.8299
-0.5000  0.8362

 0.5000  0.8582
 0.5000 -0.8118
```

```
>> L(1).W
ans =
    0.9330    0.2343
    1.3015    1.4955
    1.5721    1.5051

>> L(1).b
ans =
   -0.1186
   -0.8131
    0.8907

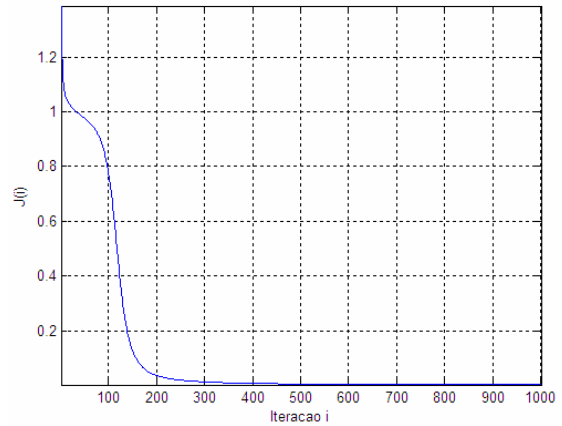
>> L(2).W
ans =
    0.0801    0.7409    0.2331
    0.3629    1.3747   -1.7343

>> L(2).b
ans =
   -0.2018
    1.4611

>> L(3).W
ans =
    0.3986    1.7967
   -0.1856   -1.8501

>> L(3).b
ans =
    0.1089
   -0.0147
```

Resultados em modo Batch (basta substituir E = 1 por E = 4 no código acima)



Entradas e Saidas

```
-0.5000  0.9438
-0.5000 -0.9512

-0.5000 -0.9587
 0.5000  0.9598

 0.5000 -0.9572
-0.5000  0.9583

 0.5000  0.9650
 0.5000 -0.9542
```

```
>> L(1).W
ans =
    0.9299    0.2892
    1.5128    1.6586
    1.7624    1.6978

>> L(1).b
ans =
   -0.1426
   -0.8752
    0.9415

>> L(2).W
ans =
    0.0905    0.8291    0.1338
    0.3278    1.6361   -2.0063

>> L(2).b
ans =
   -0.1877
    1.6381

>> L(3).W
ans =
    0.4775    2.3948
   -0.3173   -2.4261

>> L(3).b
ans =
    0.2152
   -0.1305
```