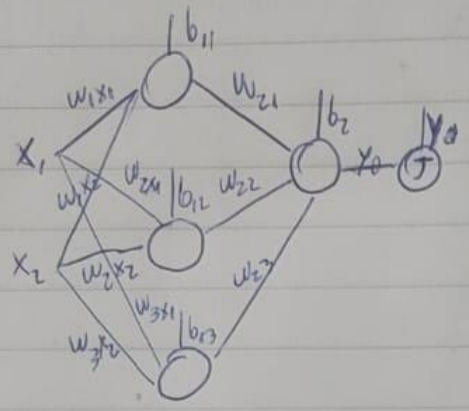
Análisis Matemático

EJERCICIO 4.5



$$z_1 = w_1 x + b_1$$

$$y_1 = G(w_1 x + b_1)$$

$$z_2 = w_2 x_{(y_1)} + b_2$$

$$J = \frac{1}{2}(y_0 - y)^2 \qquad y_0 = G(w_2 \cdot x_{(y_1)} + b_2)$$

$$\frac{\partial J}{\partial w_2} = \frac{\partial J}{\partial y_0} \cdot \frac{\partial y_0}{\partial z_2} \cdot \frac{\partial z_2}{\partial w_2} = (y_0 - y) \cdot \left[G(w_2 \cdot x_{(y_1)} + b_2) \cdot \left(1 - G(w_2 x_{(y_1)} + b_2)\right) \right] \cdot x_{(y_1)}$$

$$\frac{\partial J}{\partial b_2} = \frac{\partial J}{\partial y_0} \cdot \frac{\partial y_0}{\partial z_2} \cdot \frac{\partial z_2}{\partial b_2} = (y_0 - y) \left[G(w_2 x_{(y_1)} + b_2) \cdot \left(1 - G(w_2 x_{(y_1)} + b_2)\right) \right] \cdot 1$$

$$\frac{\partial J}{\partial x_{(y_1)}} = \frac{\partial J}{\partial y_0} \cdot \frac{\partial y_0}{\partial z_2} \cdot \frac{\partial z_2}{\partial x_{(y_1)}} = (y_0 - y) \cdot \left[G(w_2 x_{(y_1)} + b_2) \cdot \left(1 - G(w_2 x_{(y_1)} + b_2)\right) \right] \cdot w_2$$

$$\frac{\partial J}{\partial w_1} = \frac{\partial J}{\partial y_0} \cdot \frac{\partial y_0}{\partial z_2} \cdot \frac{\partial z_2}{\partial x_{(y_1)}} \cdot \frac{\partial x_{(y_1)}}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_1} = \frac{\partial J}{\partial x_{(y_1)}} \cdot \left[G(z_1) \cdot \left(1 - G(z_1)\right) \right] \cdot x$$

$$\frac{\partial J}{\partial x_1} = \frac{\partial J}{\partial y_0} \cdot \frac{\partial y_0}{\partial z_2} \cdot \frac{\partial z_2}{\partial x_{(y_1)}} \cdot \frac{\partial x_{(y_1)}}{\partial z_1} \cdot \frac{\partial z_1}{\partial x_1} = \frac{\partial J}{\partial x_{(y_1)}} \cdot \left[G(z_1) \cdot \left(1 - G(z_1)\right) \right] \cdot w_1$$

$$\frac{\partial J}{\partial b_1} = \frac{\partial J}{\partial y_0} \cdot \frac{\partial y_0}{\partial z_2} \cdot \frac{\partial z_2}{\partial x_{(y_1)}} \cdot \frac{\partial x_{(y_1)}}{\partial z_1} \cdot \frac{\partial z_1}{\partial b_1} = \frac{\partial J}{\partial x_{(y_1)}} \cdot \left[G(z_1) \cdot \left(1 - G(z_1)\right) \right] \cdot 1$$

Funciones auxiliares para aplicar regla de la cadena, algunas calculadas manualmente en la hoja.

```
[29] def sig(x):
       return(1/(1+np.power((math.e),-x)))
     def dsig(x):
       return sig(x) * (1 - sig(x))
     def dJy0(y0,y):
       return (y0-y)
     def dyz(w,x,b):
       return dsig(w.dot(x)+b)
     def dzw(x):
       return(x)
     def dzb():
       return(1)
     def dzx(w):
       return(w)
```

we,be,x: Son los vectores de pesos, bias y entrada, respectivamente de la primer capa.

w2,b2,y1: Son los vectores de pesos, bias y entrada, respectivamente de la segunda capa.

y0: es la salida de la red predicha.

y: es la salida esperada o etiqueta.

```
[4]  we=np.array([[0.1, -0.5],[-0.3, -0.9],[0.8, 0.02]])
     be=np.array([0.1, 0.5, 0.8])
     w2=np.array([-0.4, 0.2, -0.5])
     b2=0.7
     x=np.array([1.8, -3.4])
```

Funcion de costo de la red.

```
[31] def loss(y0,y):
         loss=1/2*np.square(y0-y)
         return loss
```

Se calcula la salida de la red

```
[32] y1=sig(we.dot(x)+be)
     y0=sig(w2.dot(y1)+b2)
     y=5
     y0
```

```
0.5225766515999583
```

Se calcula el costo para esta entrada particular x

```
[33] loss(y0,y)
```

```
10.023659920398922
```

Calculo de derivadas de J respecto a los parametros we,be,w2,b2

```
[34] dJw2=dJy0(y0,y)*dyz(w2,y1,b2)*dzw(y1)
     dJy1=dJy0(y0,y)*dyz(w2,y1,b2)*dzx(w2) #Funcion auxiliar para calcular dJw1 y dJb1
     dJb2=dJy0(y0,y)*dyz(w2,y1,b2)*dzb()

     dJw1=dJy1*dyz(we,x,be)*dzw(x[0])+dJy1*dyz(we,x,be)*dzw(x[1])
     #dJx1=dJy1*dyz(we,x,be)*dzx(we[:,0])+dJy1*dyz(we,x,be)*dzx(we[:,1]) #No lo ejecutamos ya que no es un parametro
     dJb1=dJy1*dyz(we,x,be)*dzb()
```

```
print(dJw2,'\n',dJb2,'\n',dJy1,'\n',dJw1,'\n',dJb1)
```

```
[-0.98155159 -1.0650957  -1.0028046 ]
 -1.1170736711480758
 [ 0.44682947 -0.22341473  0.55853684]
 [-0.07621165  0.01585901 -0.0820641 ]
 [ 0.04763228 -0.00991188  0.05129006]
```