

```

#BGD
#initialize w0 randomly with value in [-1,1]
#b0=0
#For i=0...max_iter:
#    delta w=(0,0,...,0)
#.    delta b=0
#.    For every training example in x: 699 j from 0 to 699
#        delta w += -(y-sigmoid(xj))*xj
#        delta b +=sigmoid(xj)-y
#    if |delta w|=0
#.        break
#    w^(i+1)=w^i-lr/m delta w
#    b^(i+1)=b^i-1/m delta b

```

$$\Delta w_i = \alpha \sum_{d=1}^D (y_d - \mathbf{w}^i \cdot \mathbf{x}_d) x_{di}$$