

LR-GRADIENT-DESCENT($\theta_{init}, \theta_{0init}, \eta, \epsilon$)

1 $\theta^{(0)} = \theta_{init}$

2 $\theta_0^{(0)} = \theta_{0init}$

3 $t = 0$

4 **repeat**

5 $t = t + 1$

6 $\theta^{(t)} = \theta^{(t-1)} - \eta \left(\frac{1}{n} \sum_{i=1}^n \left(\sigma \left(\theta^{(t-1)T} \mathbf{x}^{(i)} + \theta_0^{(t-1)} \right) - y^{(i)} \right) \mathbf{x}^{(i)} + \lambda \theta^{(t-1)} \right)$

7 $\theta_0^{(t)} = \theta_0^{(t-1)} - \eta \left(\frac{1}{n} \sum_{i=1}^n \left(\sigma \left(\theta^{(t-1)T} \mathbf{x}^{(i)} + \theta_0^{(t-1)} \right) - y^{(i)} \right) \right)$

8 **until** $\left| J_{lr}(\theta^{(t)}, \theta_0^{(t)}) - J_{lr}(\theta^{(t-1)}, \theta_0^{(t-1)}) \right| < \epsilon$

9 **return** $\theta^{(t)}, \theta_0^{(t)}$