Repeat {

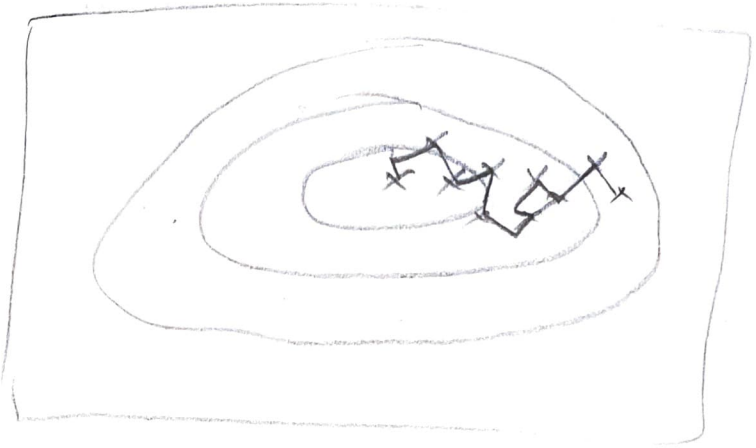
for every j

for $i = 1$ to m {

$$\theta_j := \theta_j - \alpha (h_\theta(x^{(i)}) - y^{(i)}) \cdot x_j^{(i)}$$

}

} $\Rightarrow$ Stochastic gradient descent



$\theta = (\dots)$

18008969999

$$\nabla_\theta J(\theta) = \begin{bmatrix} \frac{\partial J}{\partial \theta_0} \\ \frac{\partial J}{\partial \theta_1} \\ \frac{\partial J}{\partial \theta_2} \end{bmatrix}$$

$\theta \in \mathbb{R}^{n+1}$

$A \in \mathbb{R}^{2 \times 2}$ $\quad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$