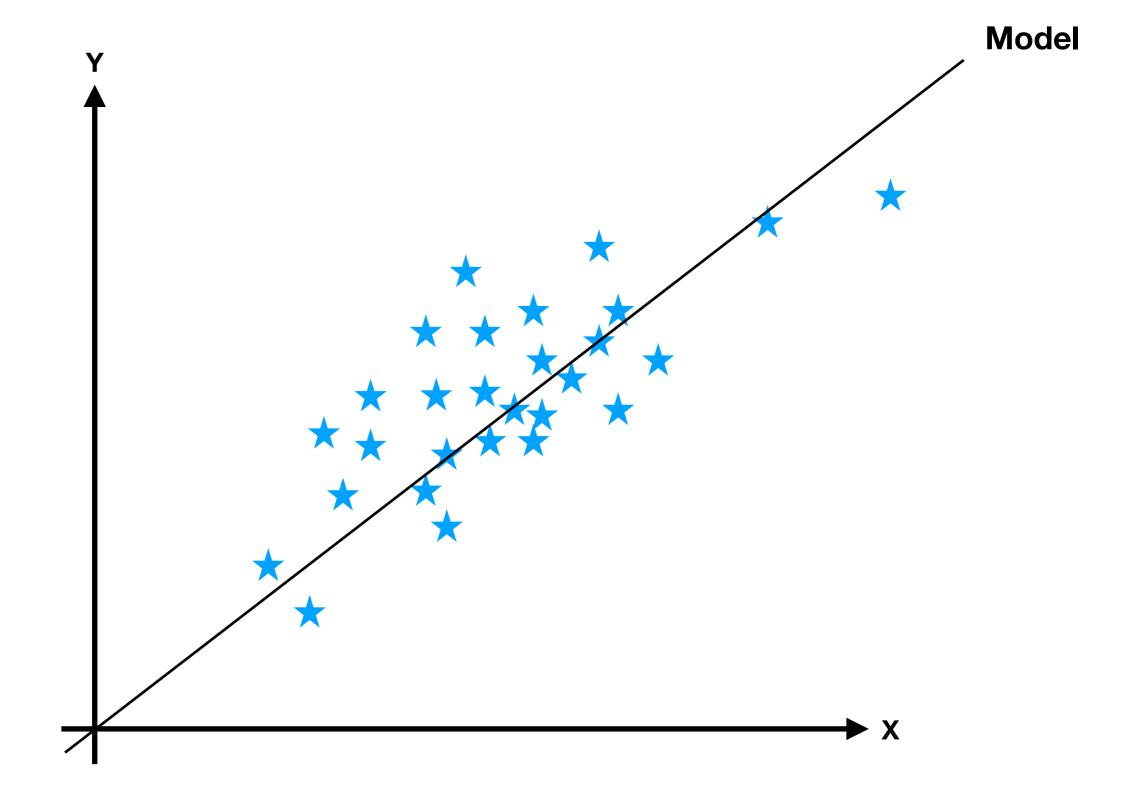
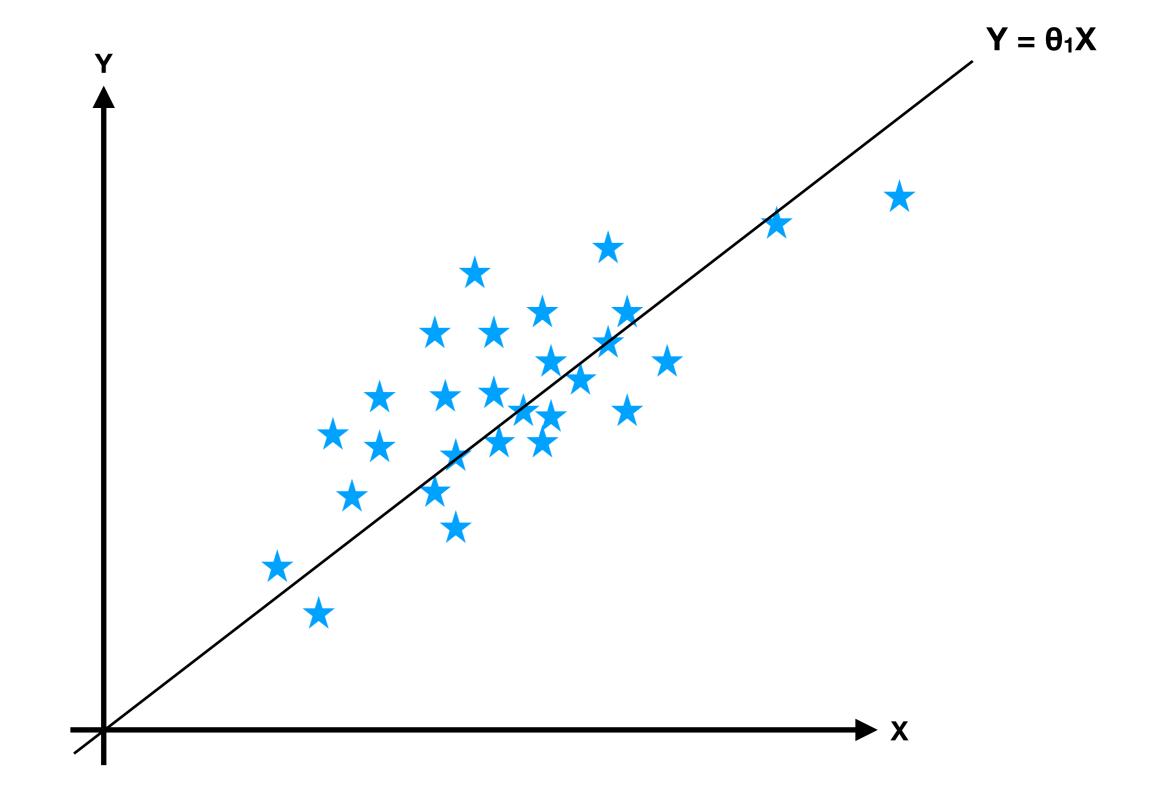
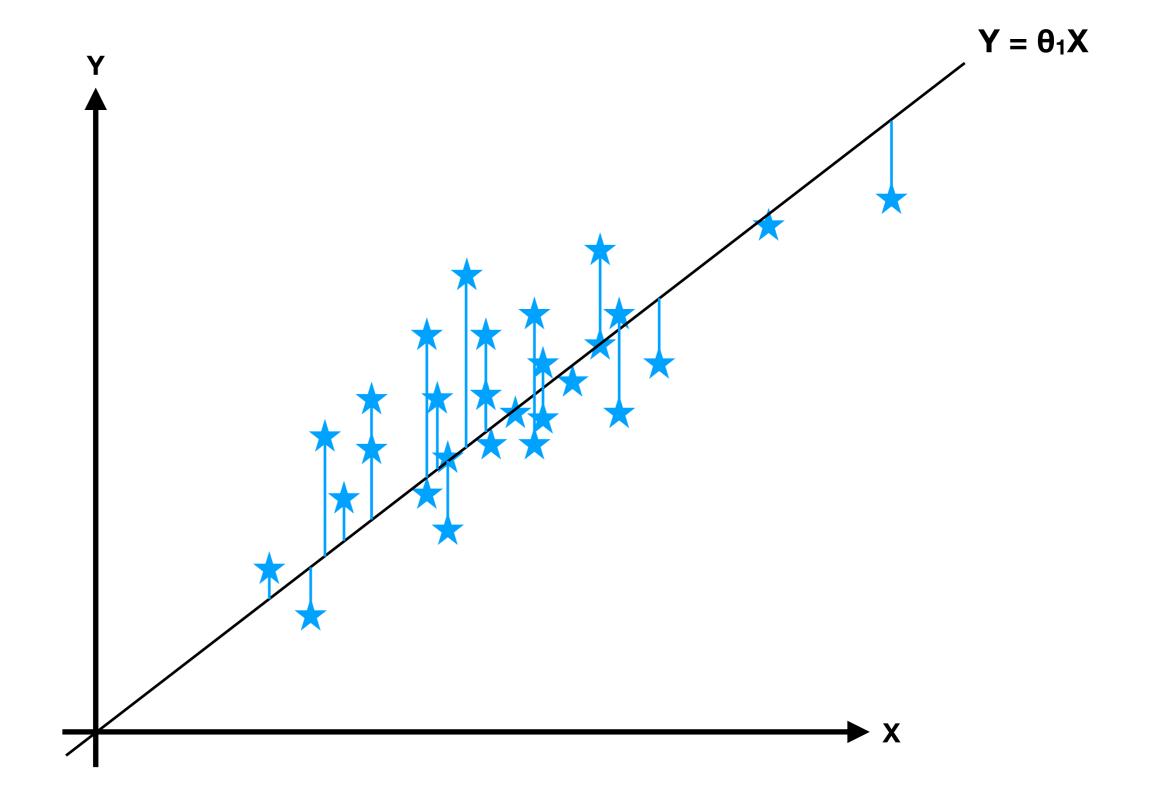
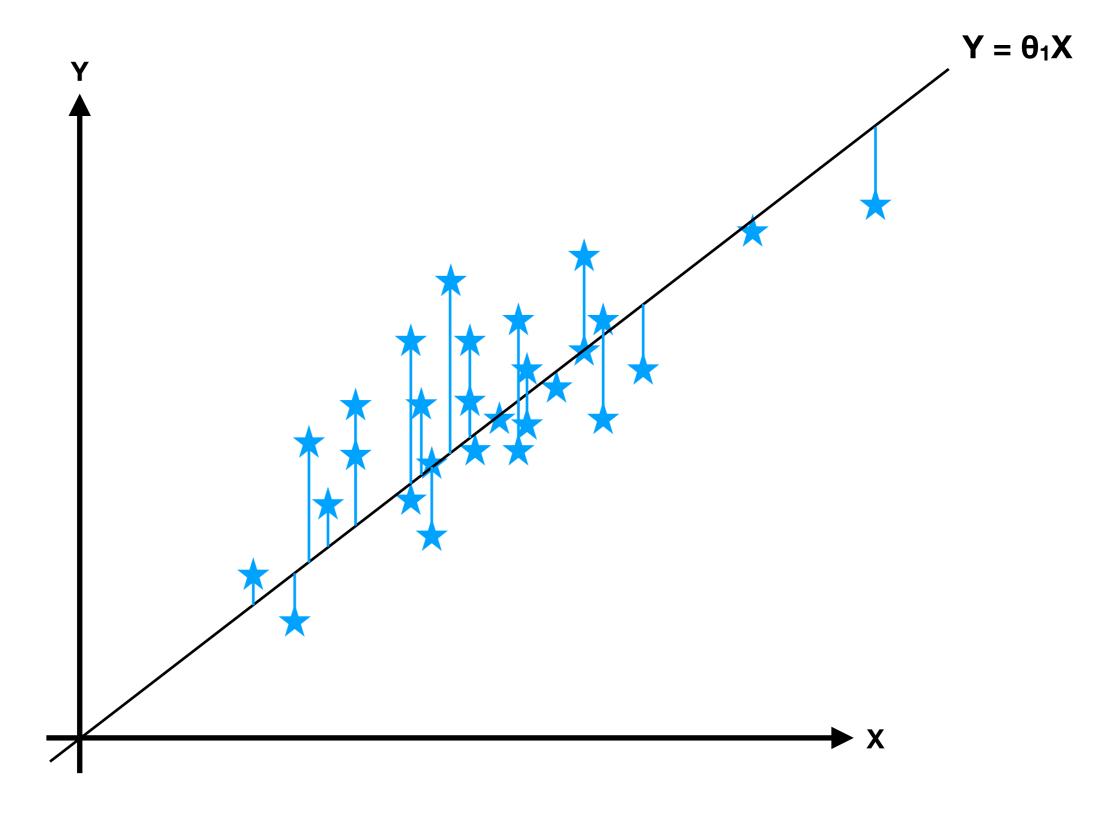
SC201

Lecture 3

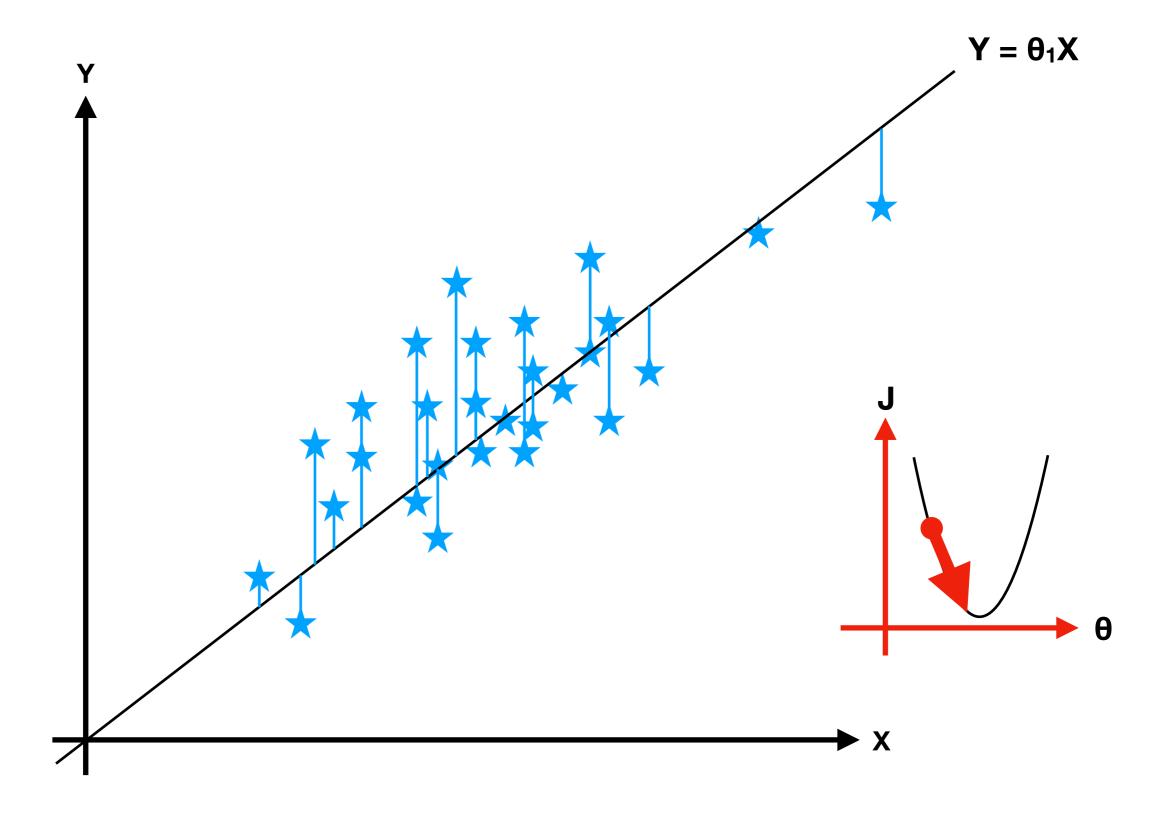




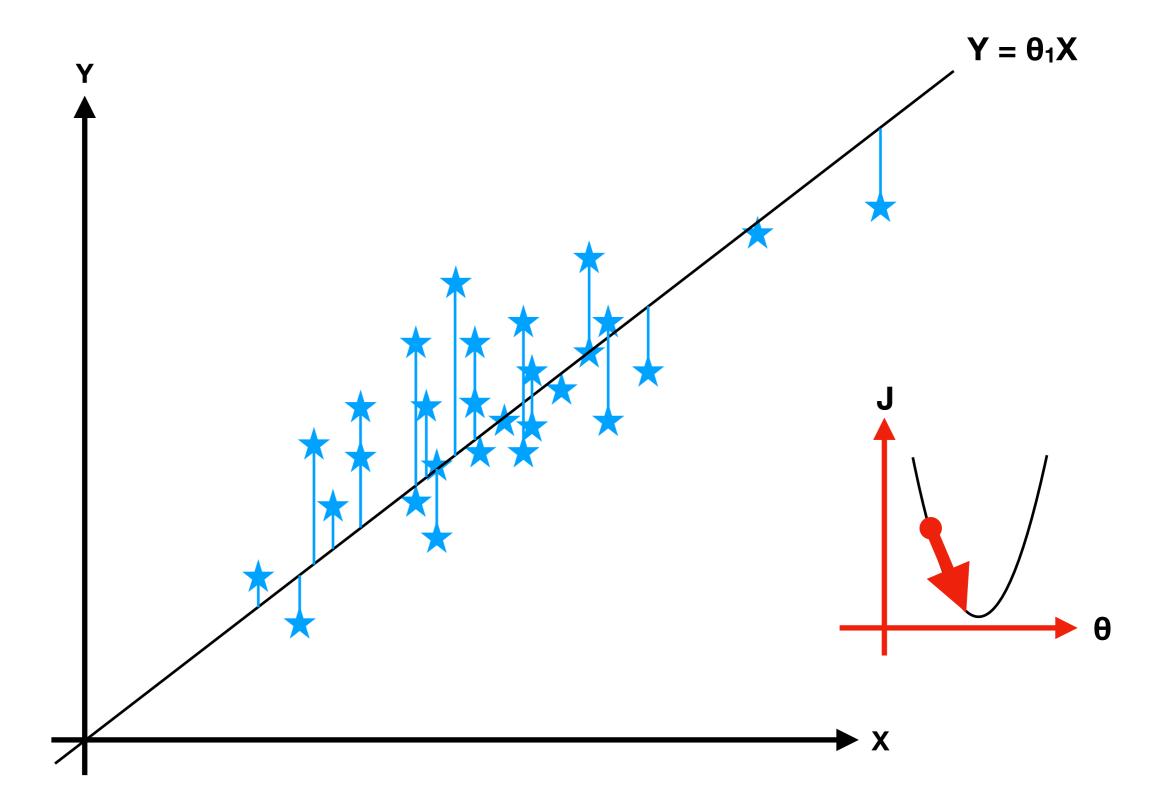




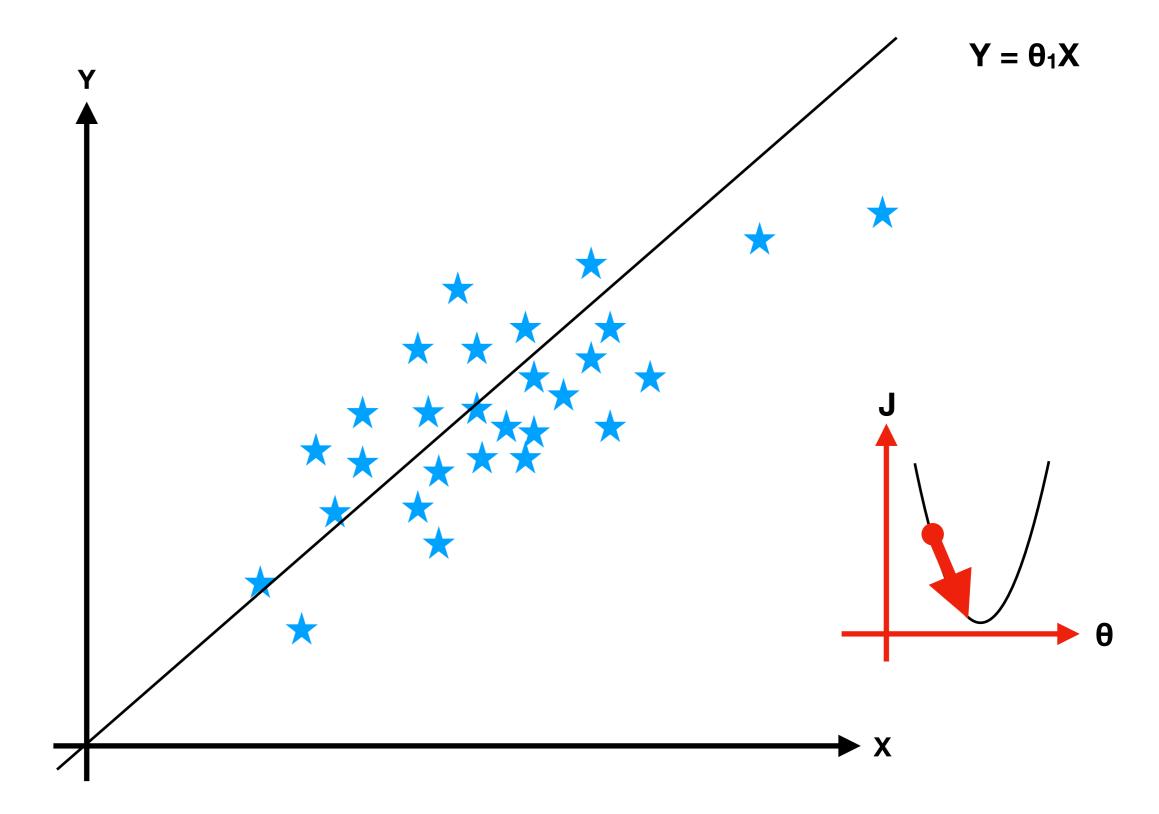
$$J = \Sigma (Y' - Y_i)^2 / 2m$$



 $J = \Sigma (Y' - Y_i)^2 / 2m$

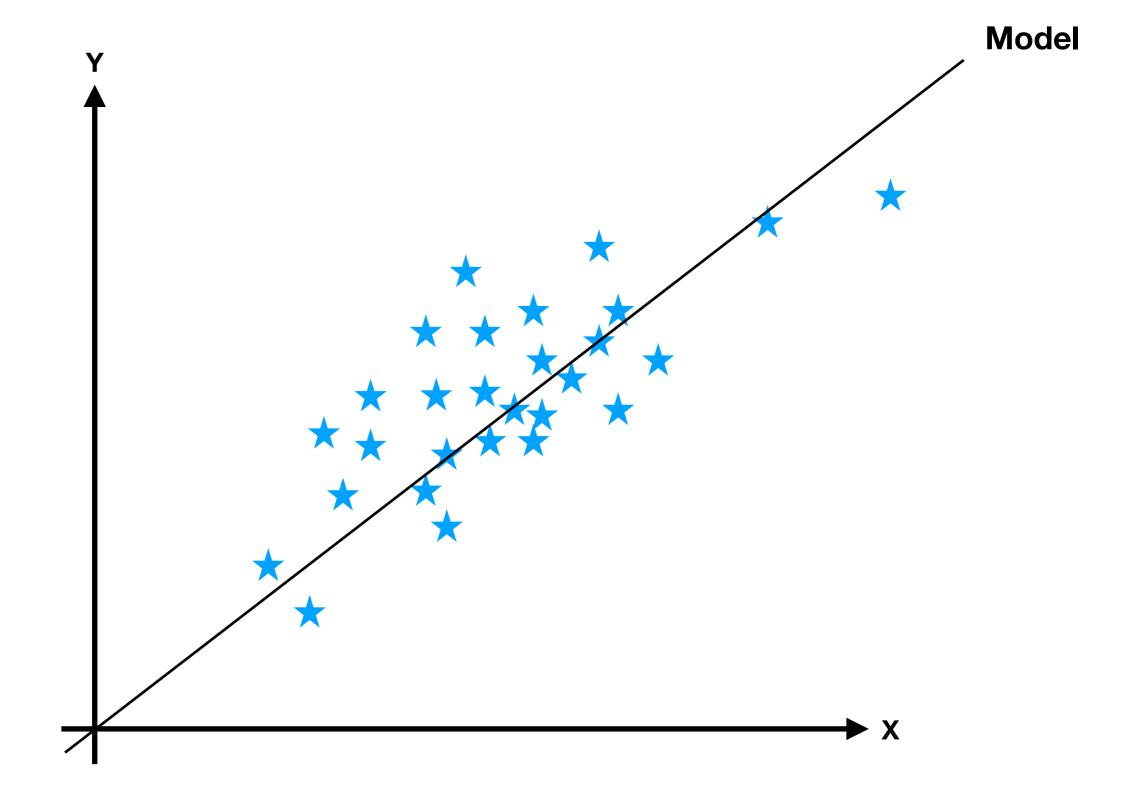


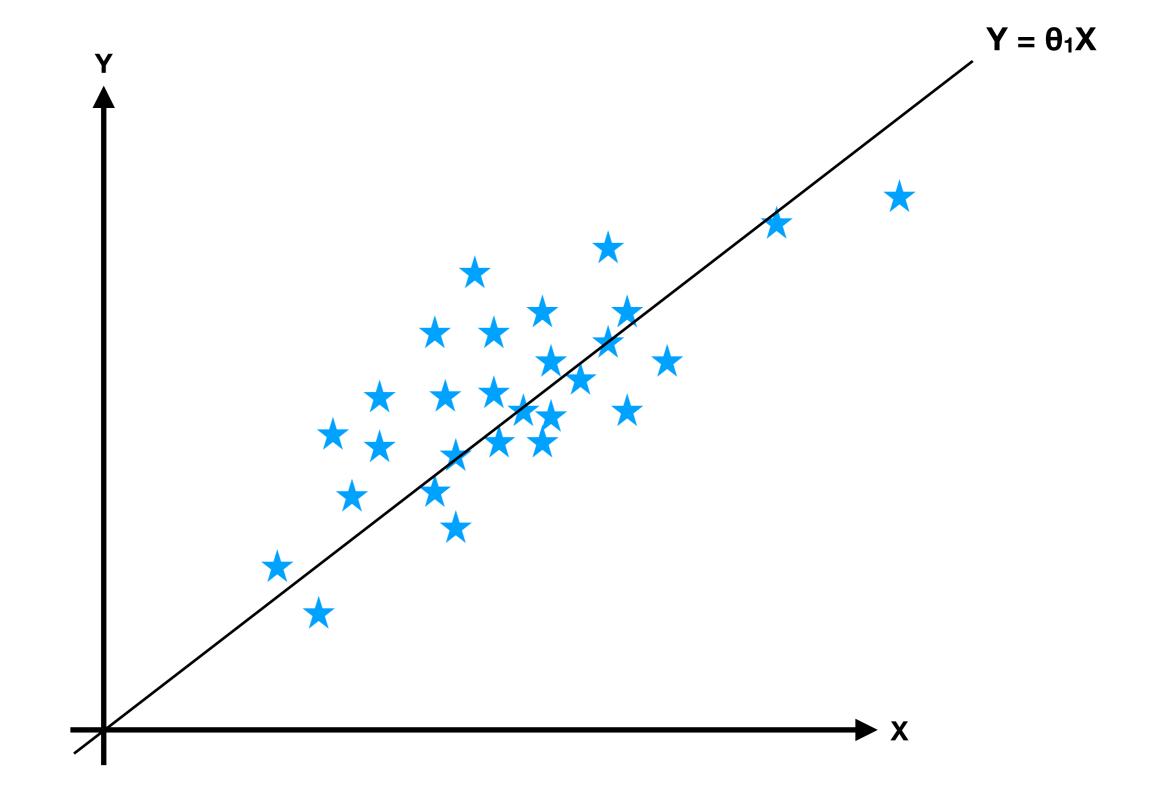
$$\theta = \theta - \alpha (\Sigma (Y' - Y_i)X_i / m)$$

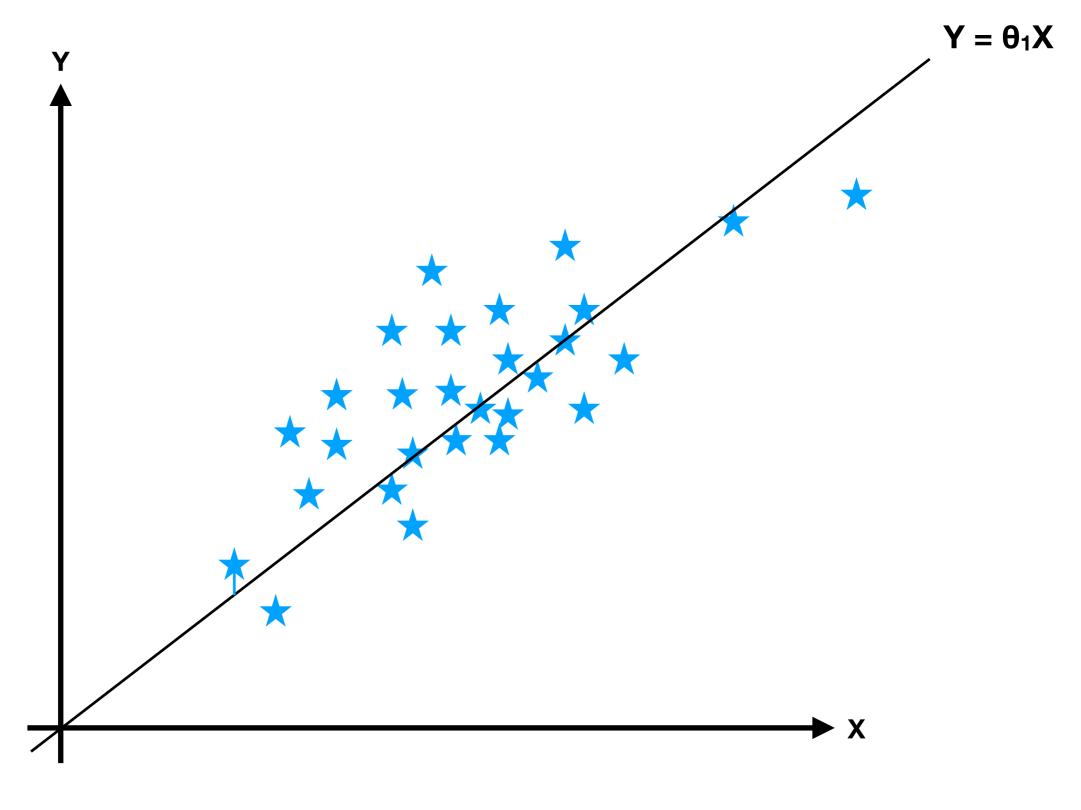


$$\theta = \theta - \alpha (\Sigma (Y' - Y_i)X_i / m)$$

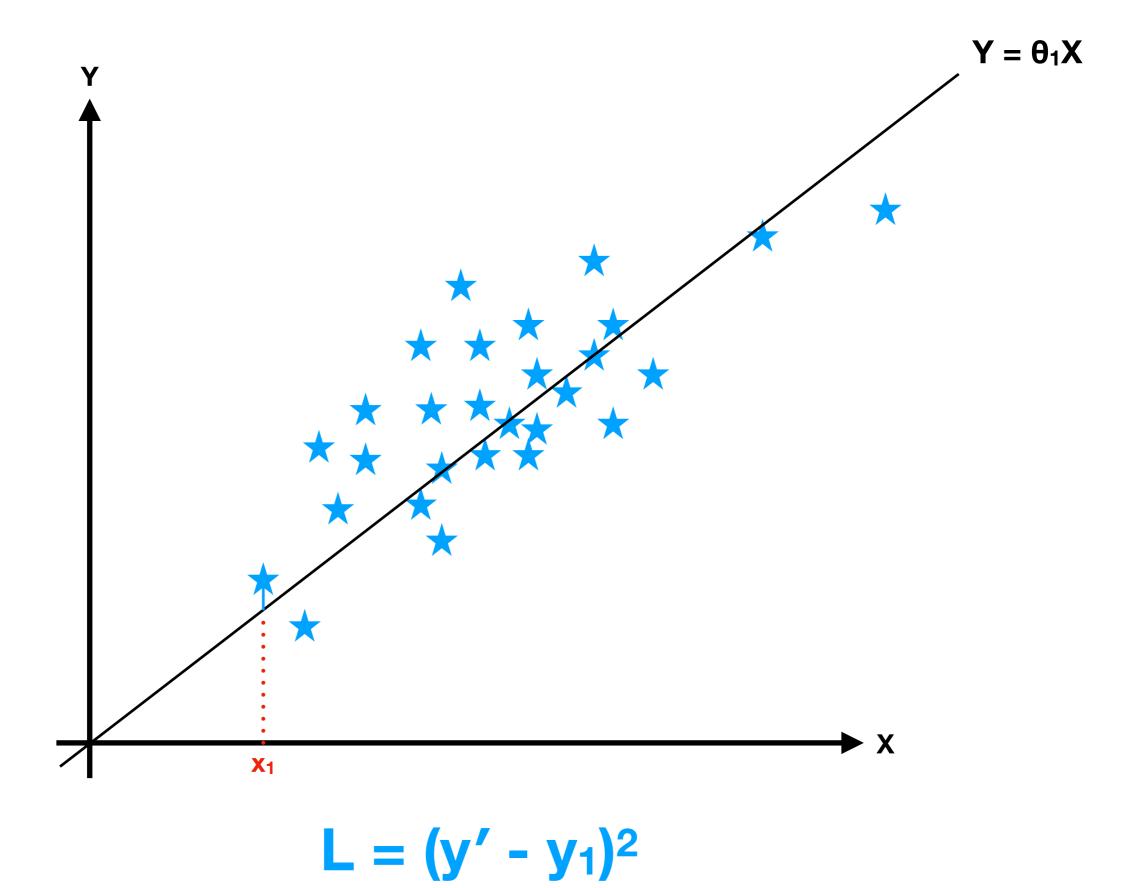
Batch Gradient Descent

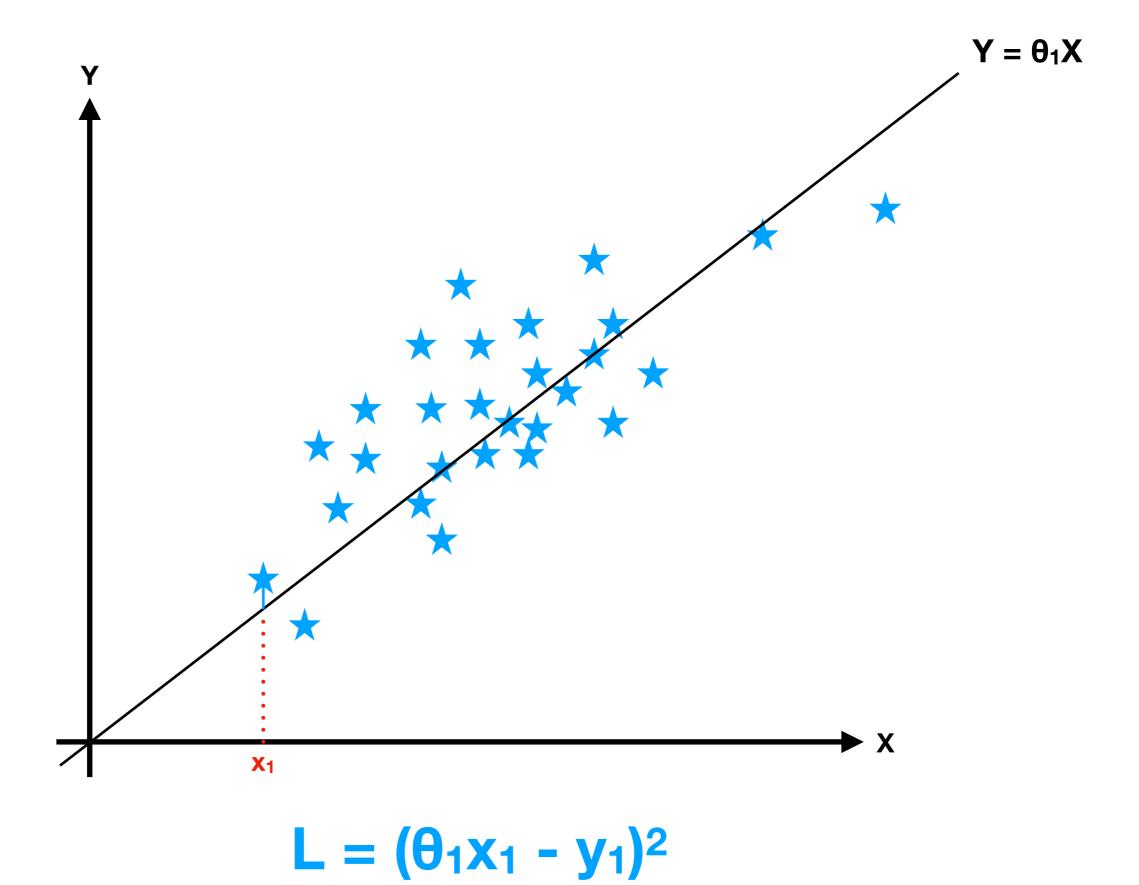


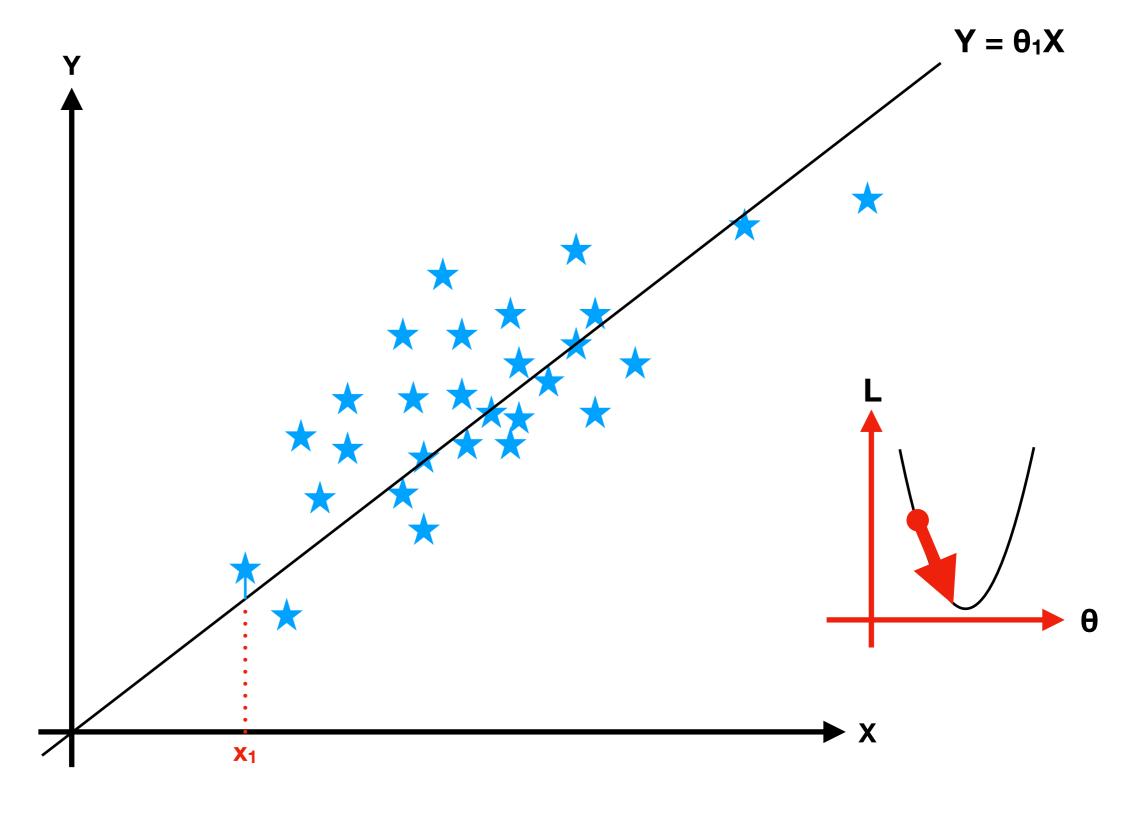




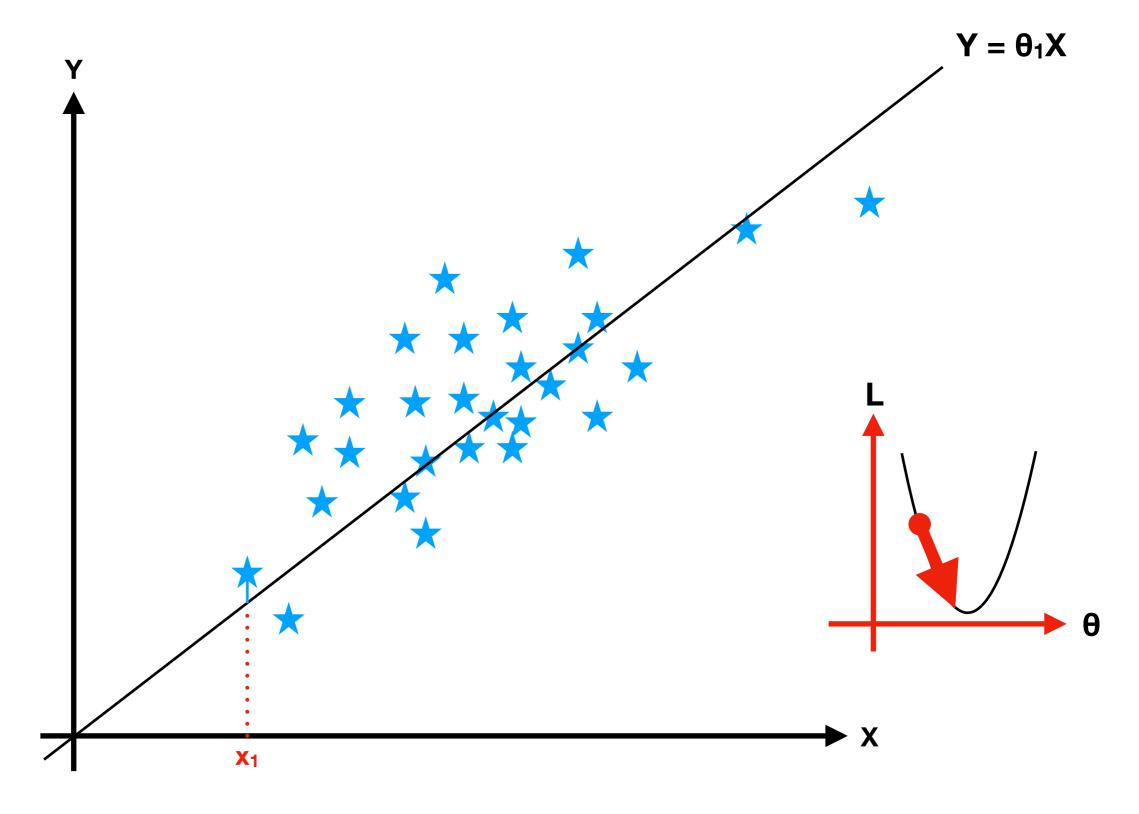
$$L = (y' - y_1)^2$$



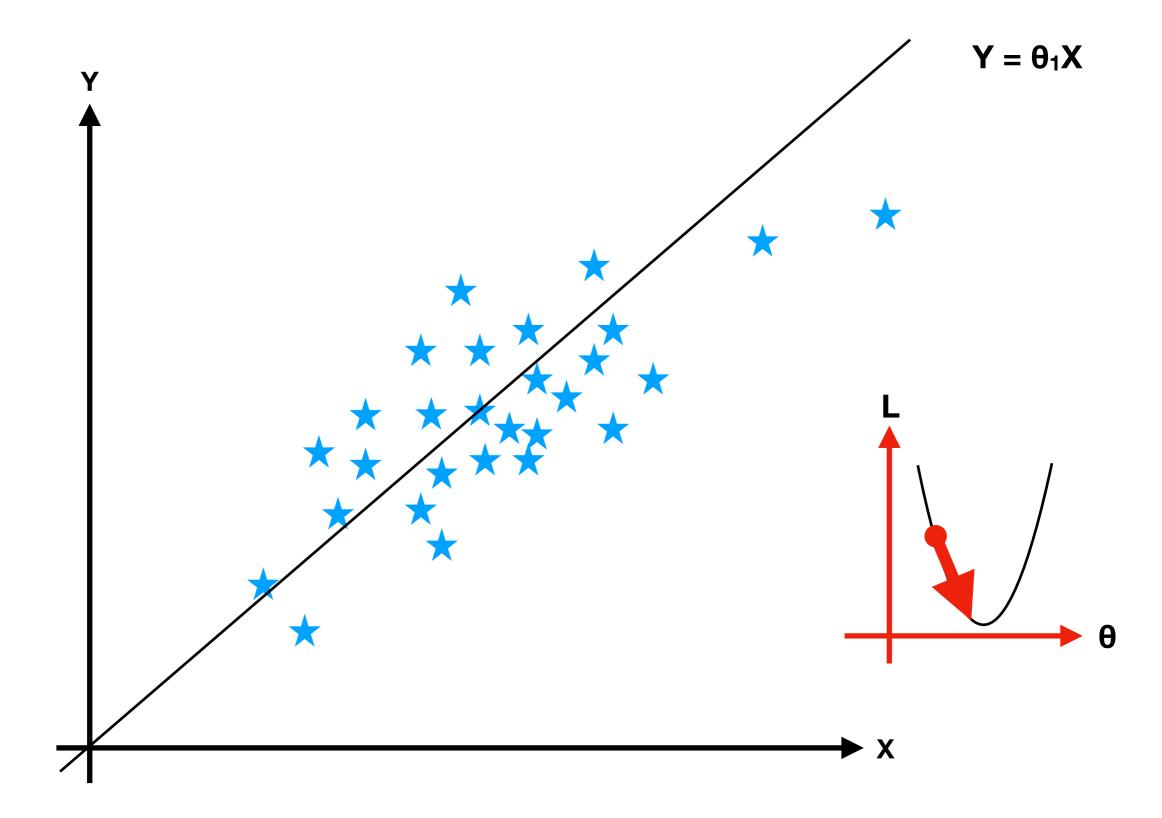




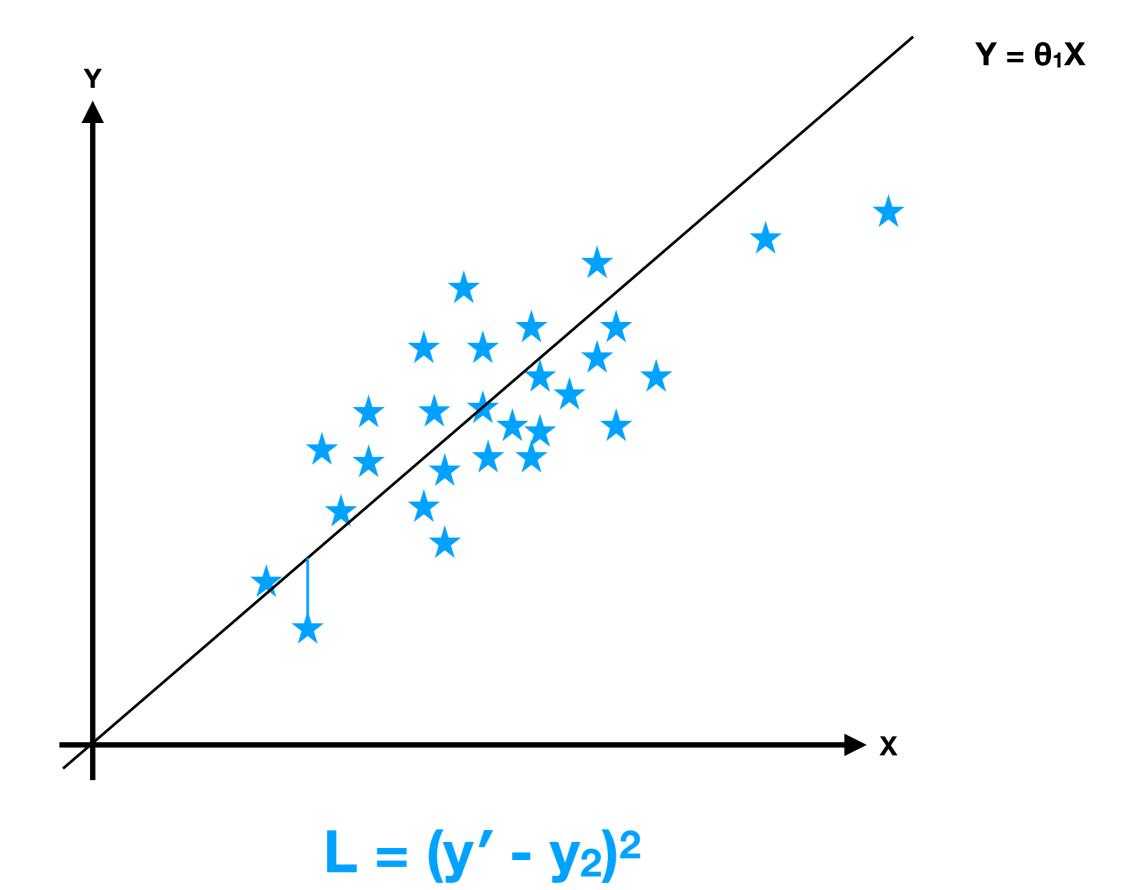
$$L = (\theta_1 x_1 - y_1)^2$$

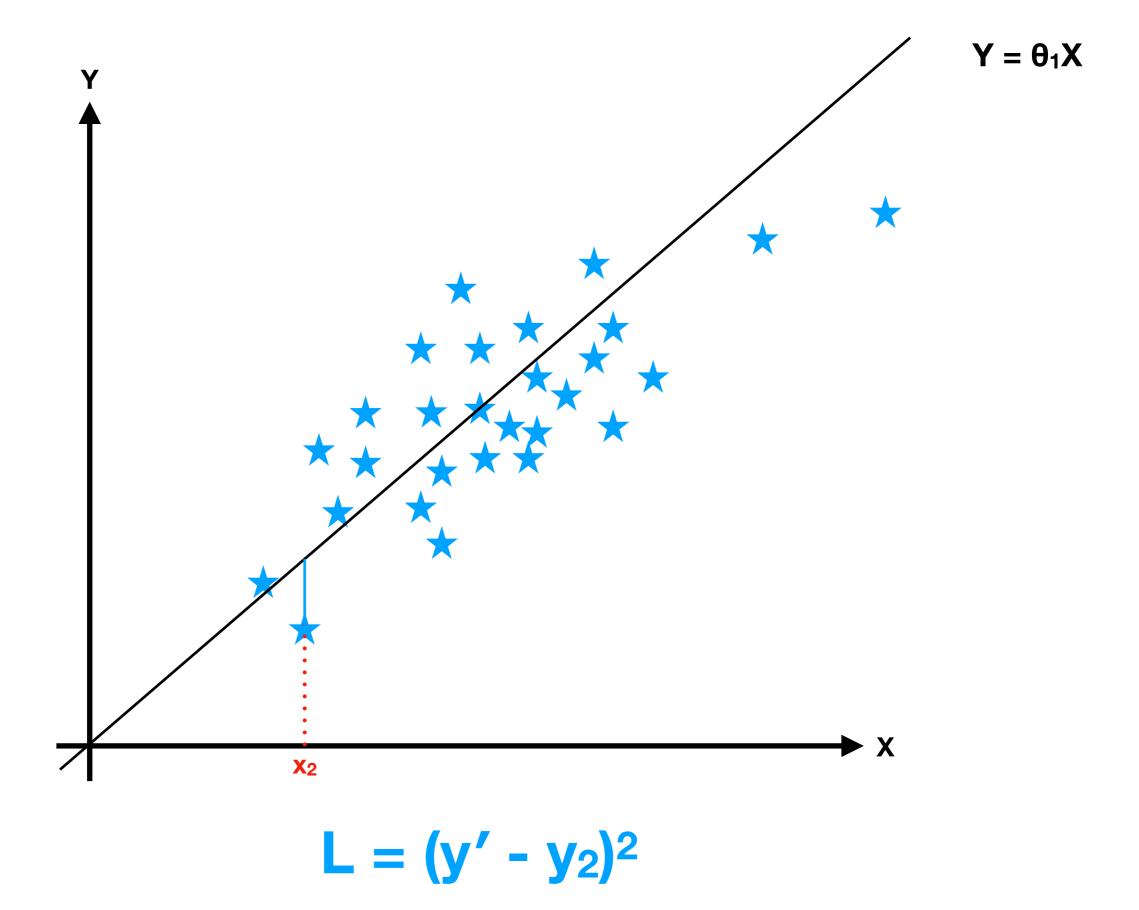


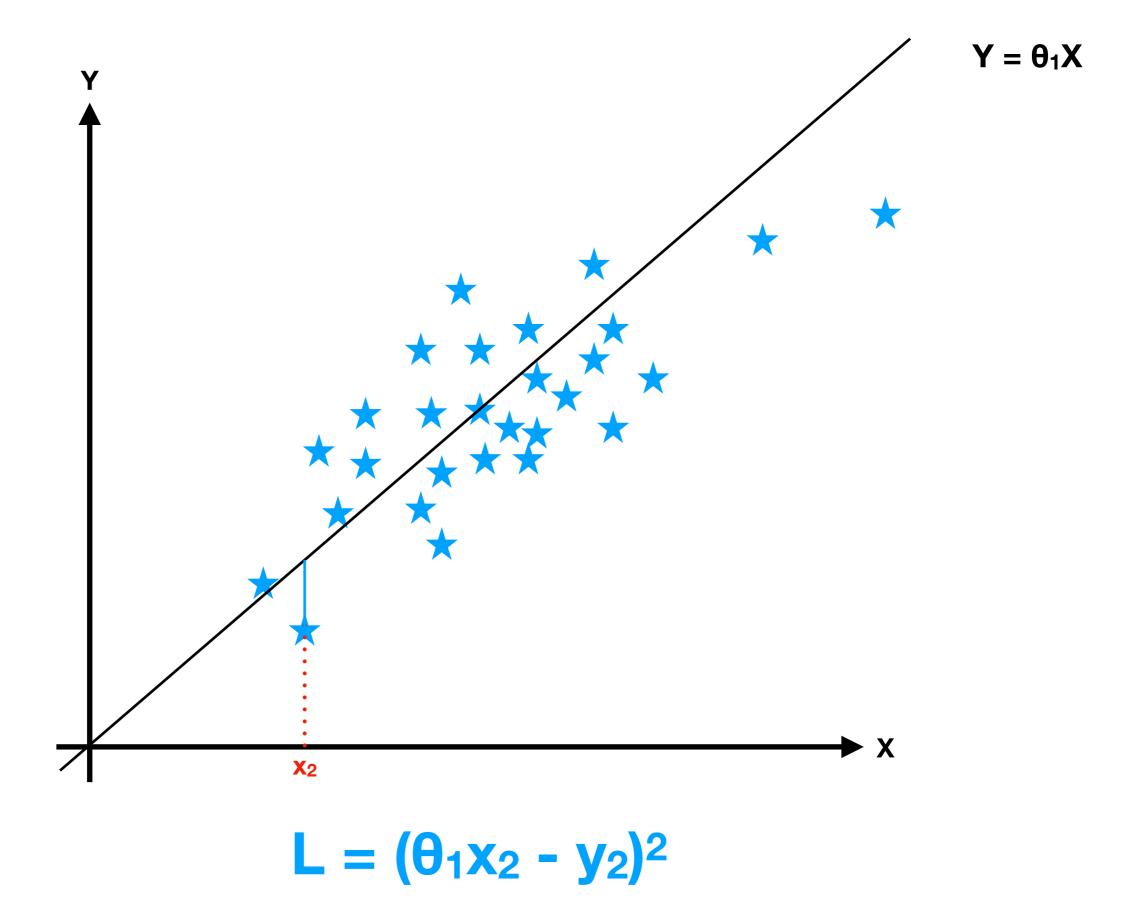
 $\theta = \theta - \alpha \left(2(\theta_1 x_1 - y_1) x_1 \right)$

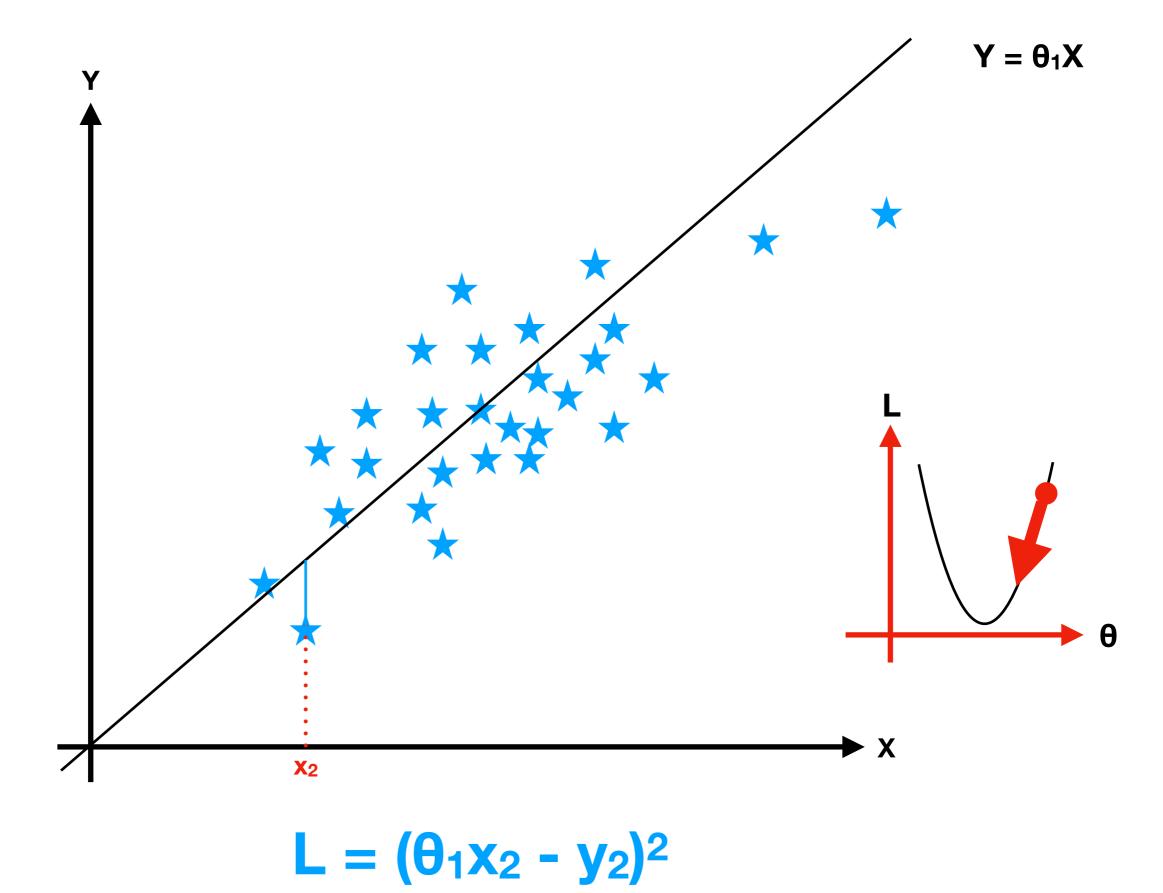


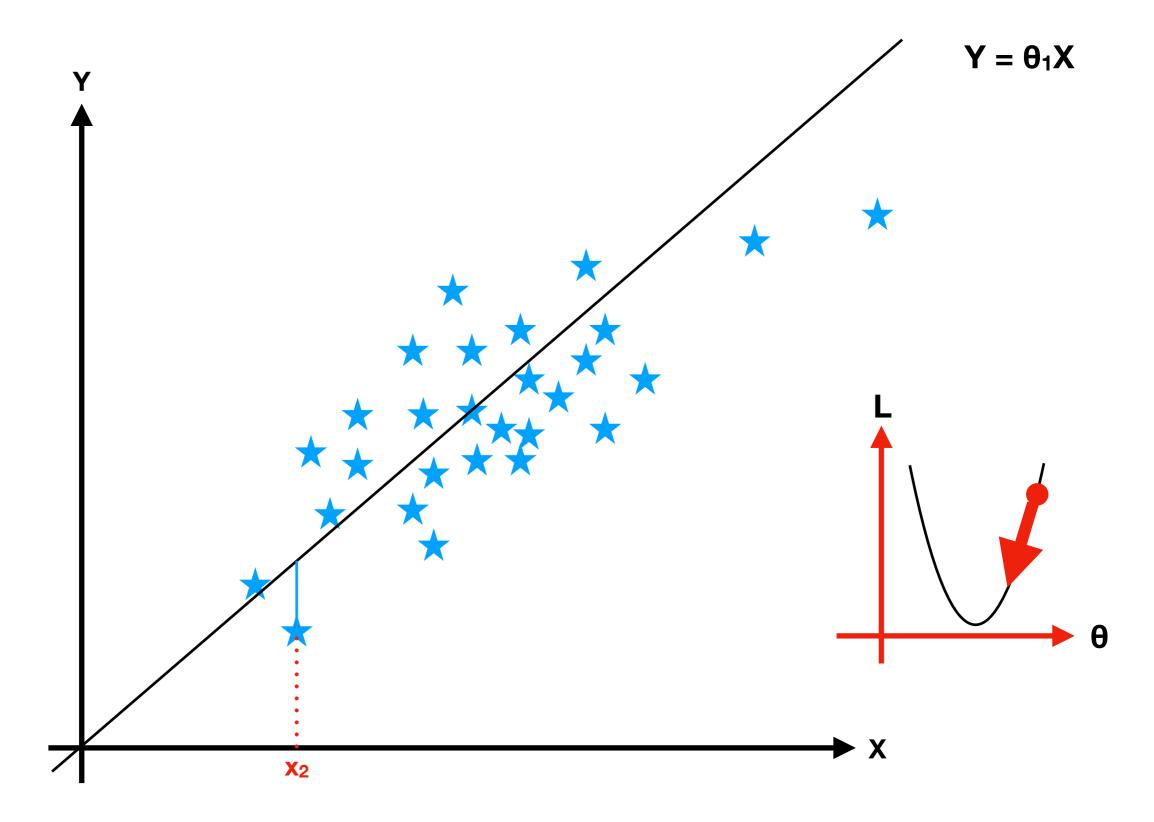
$$\theta = \theta - \alpha \left(2(\theta_1 x_1 - y_1) x_1 \right)$$



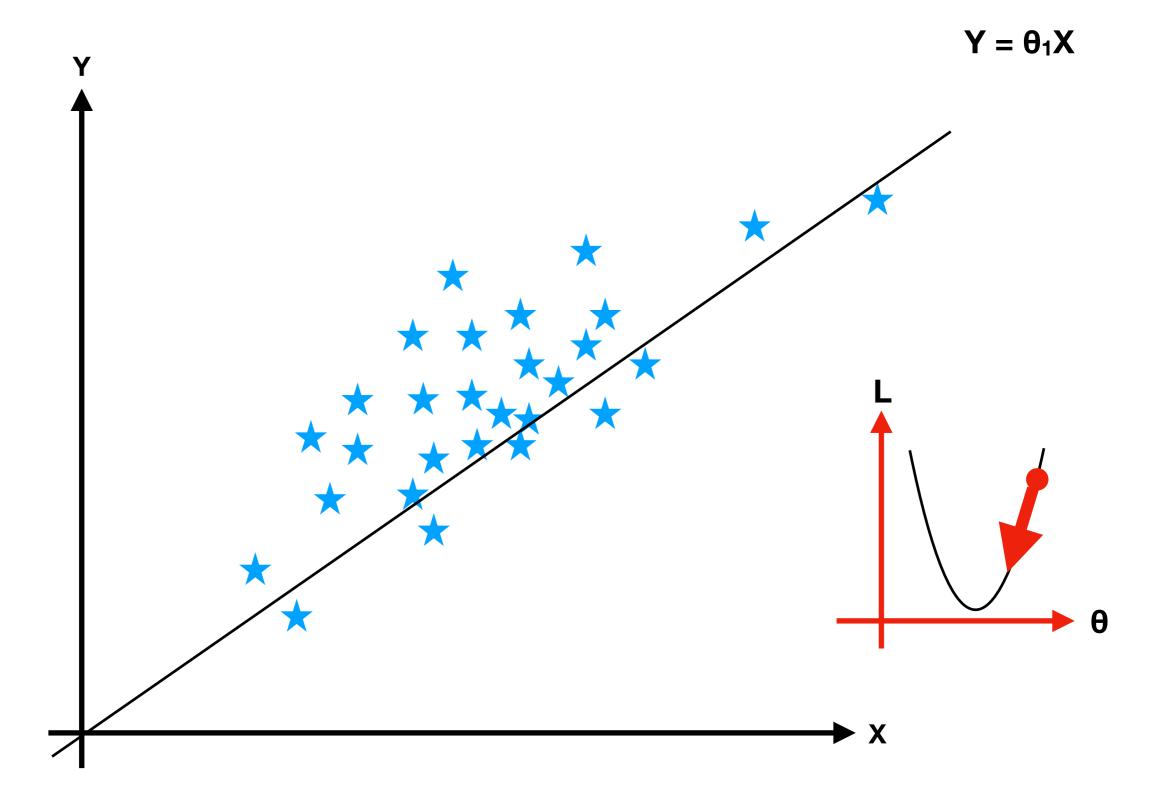




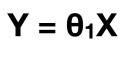


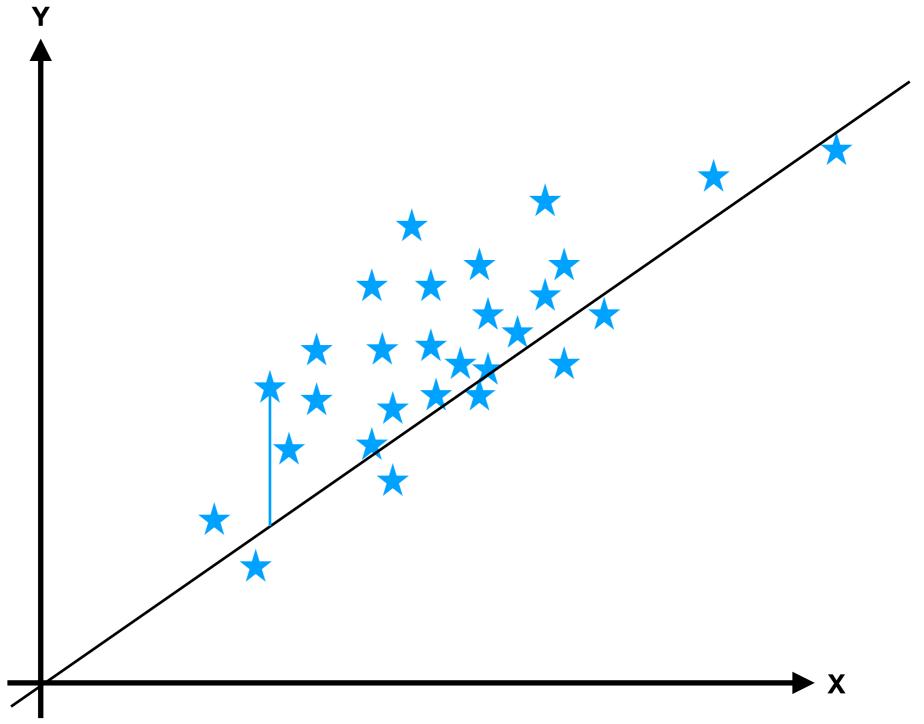


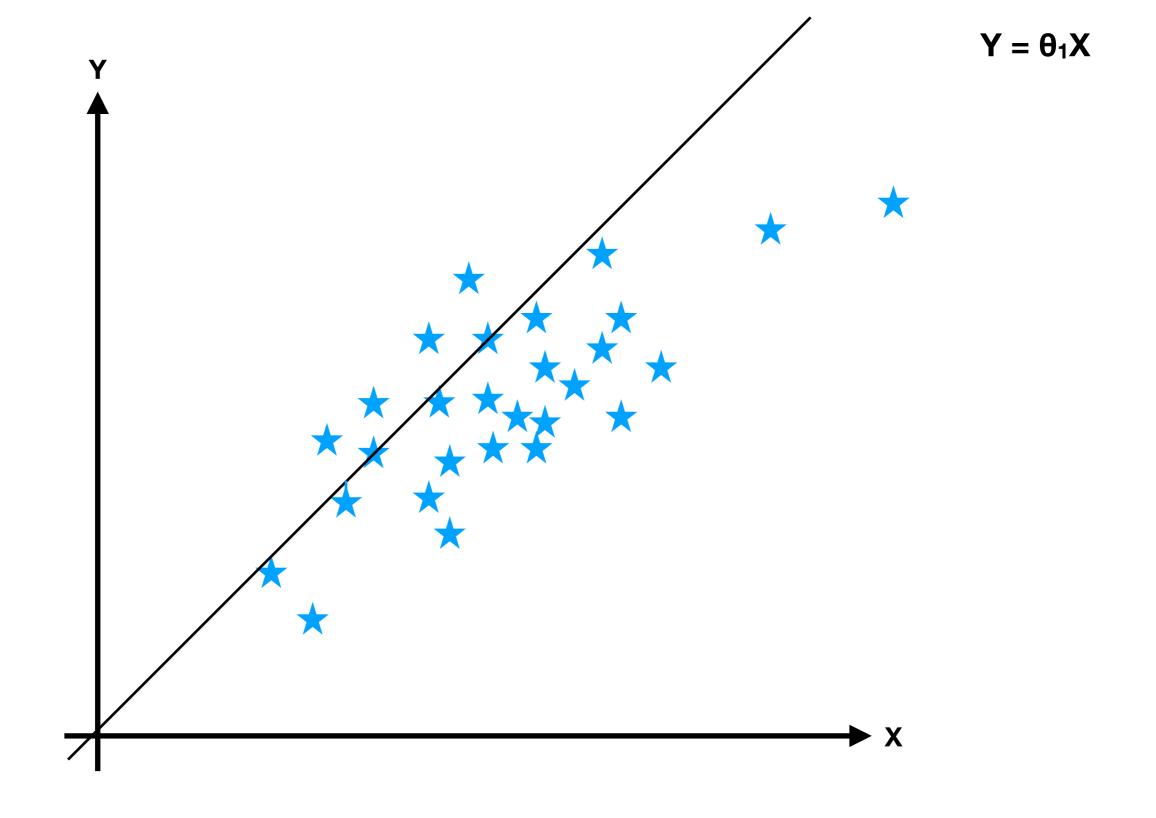
 $\theta = \theta - \alpha (2(\theta_1x_2 - y_2)x_2)$

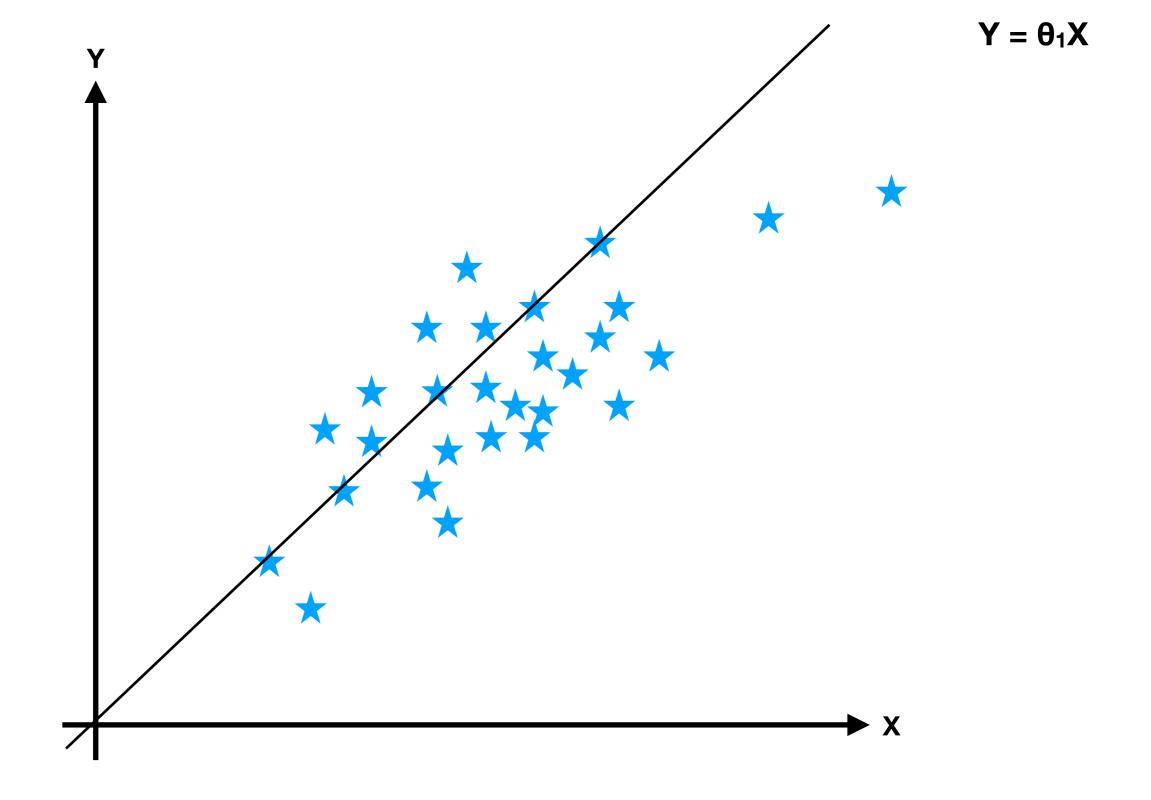


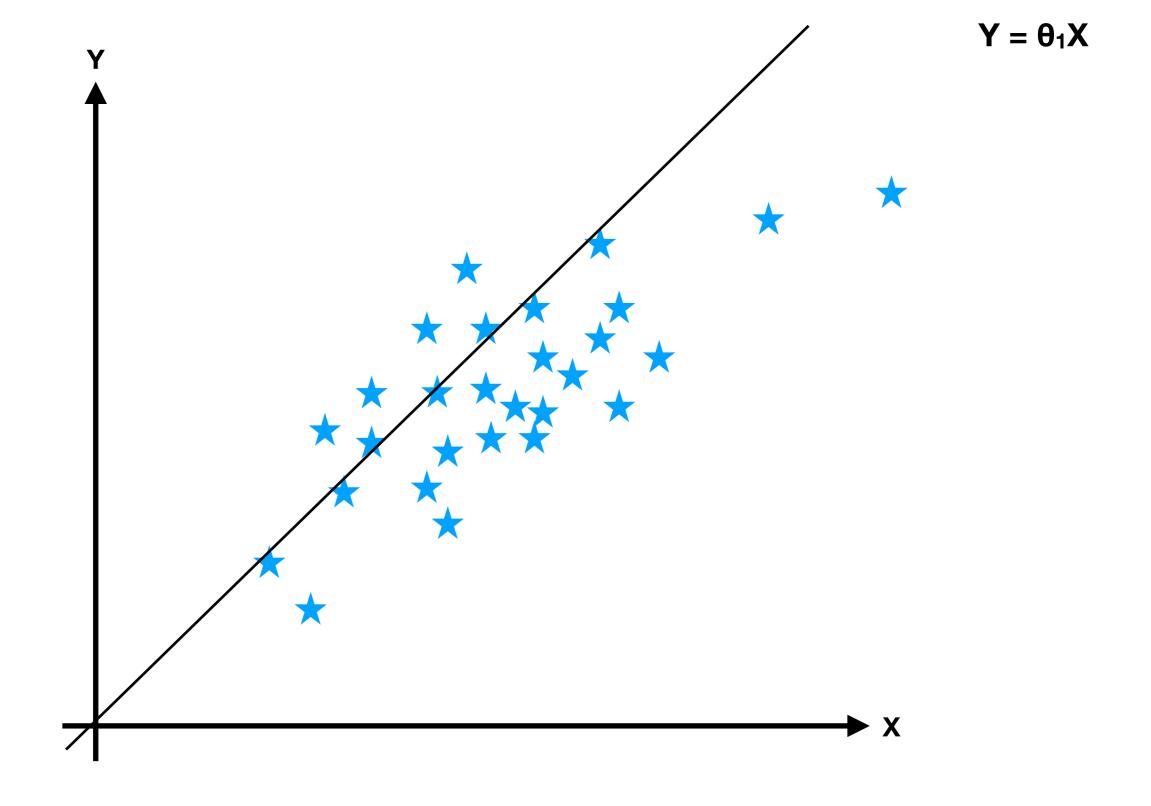
$$\theta = \theta - \alpha \left(2(\theta_1 x_2 - y_2) x_2 \right)$$

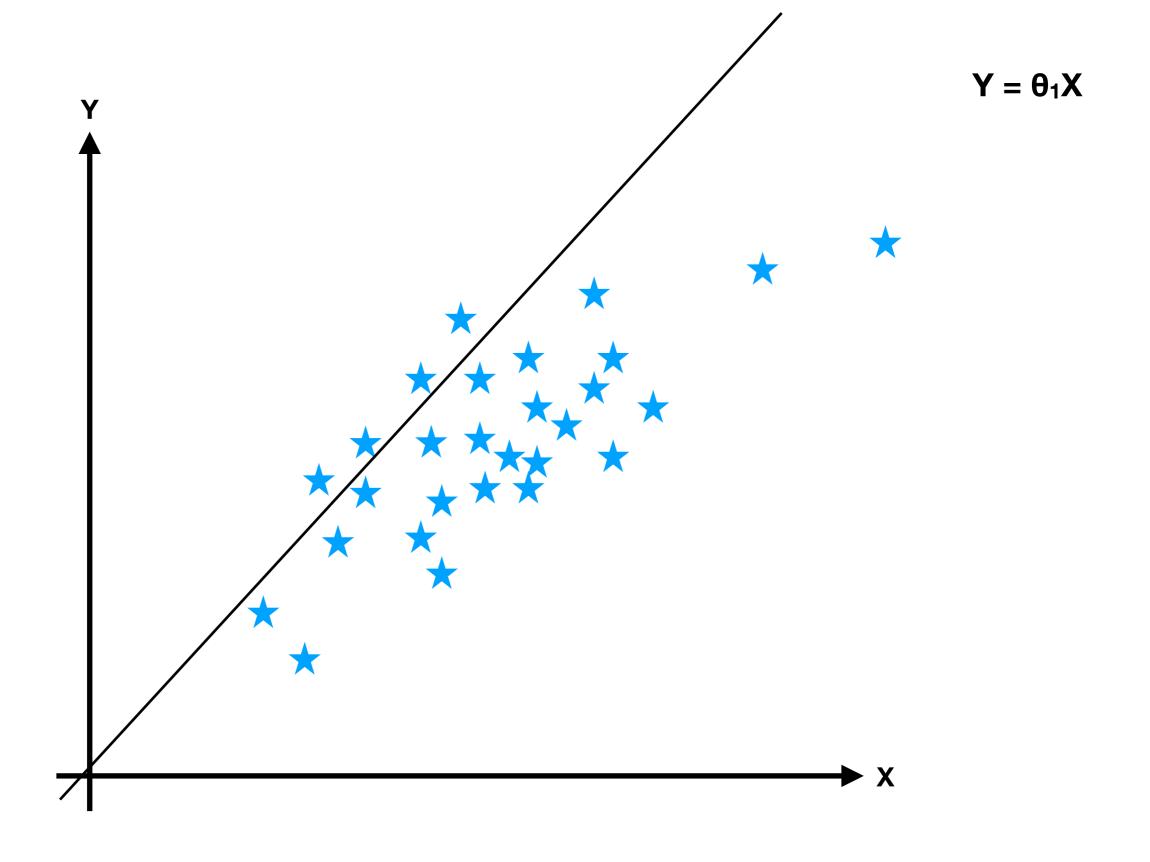


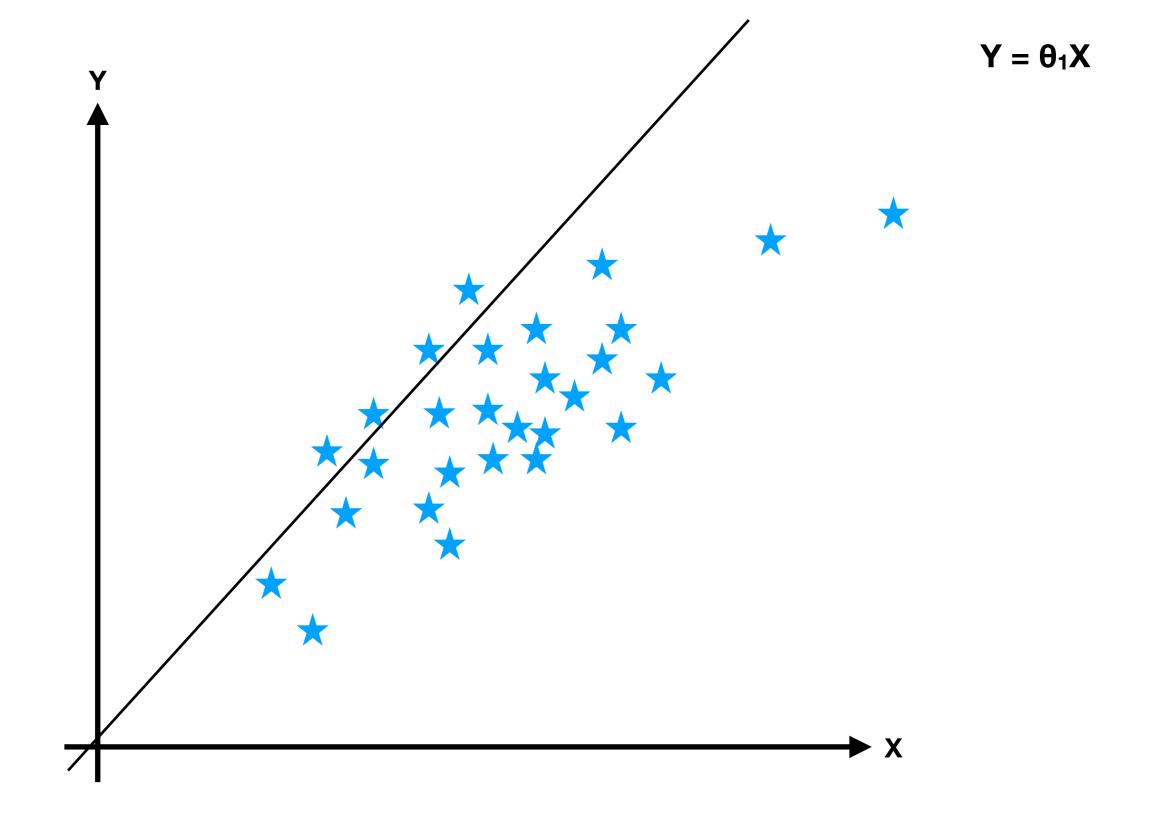


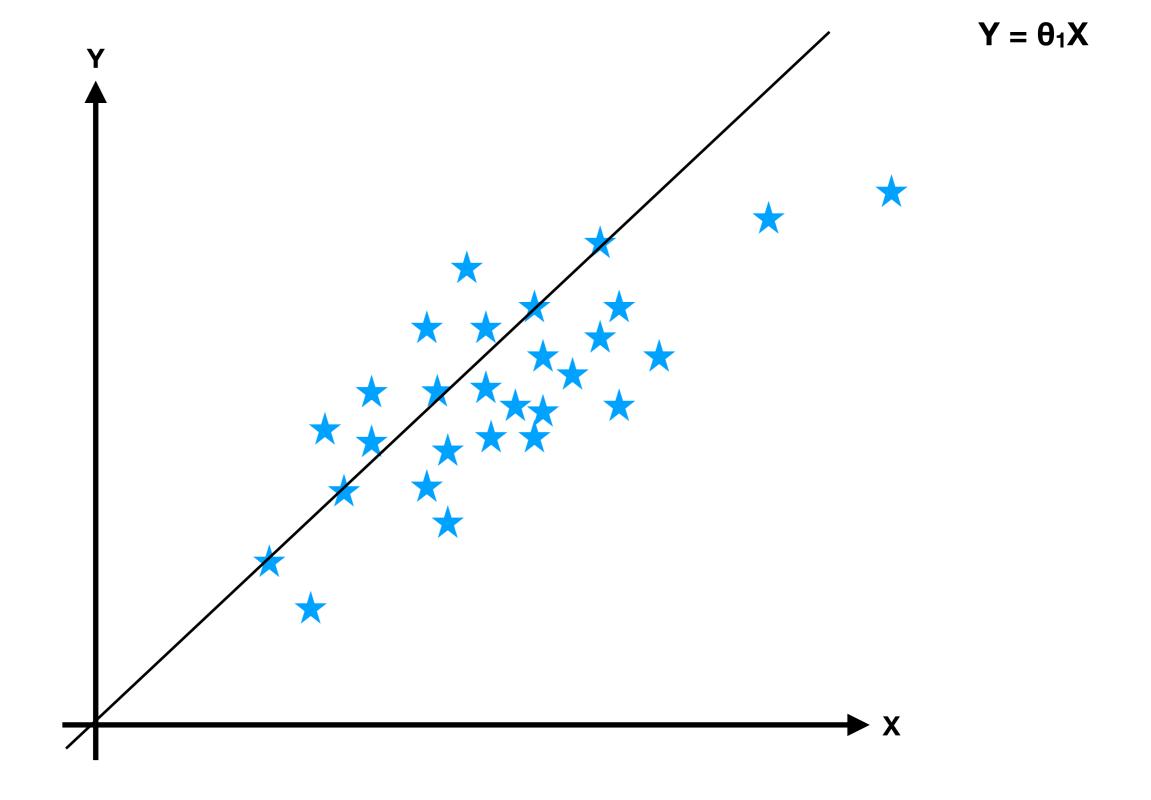


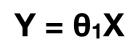


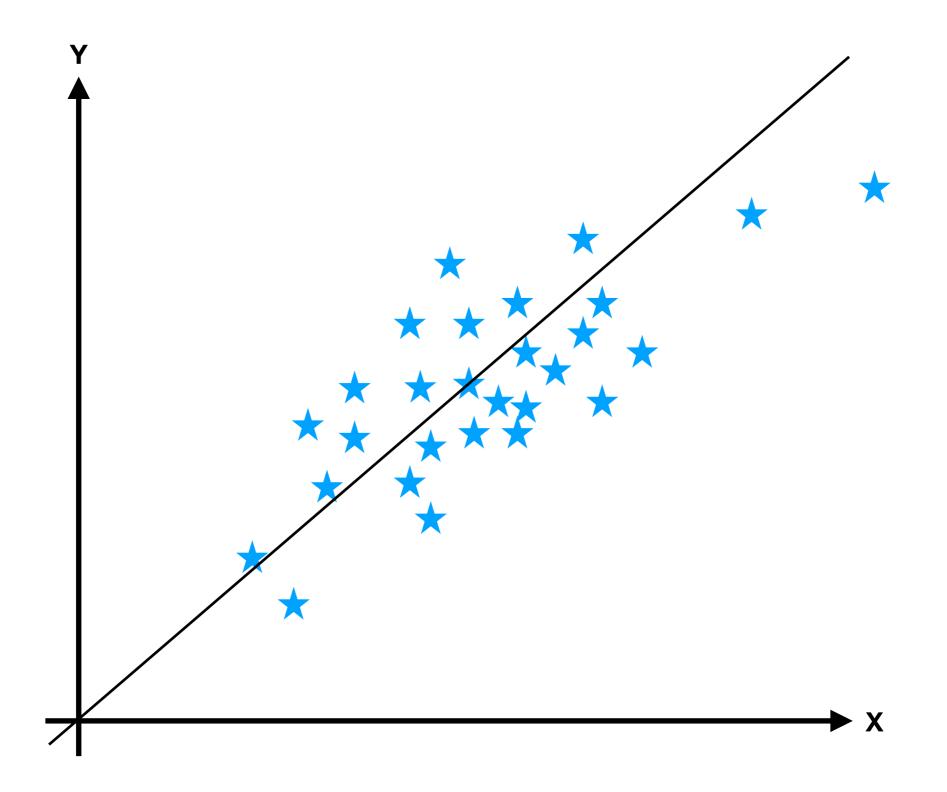


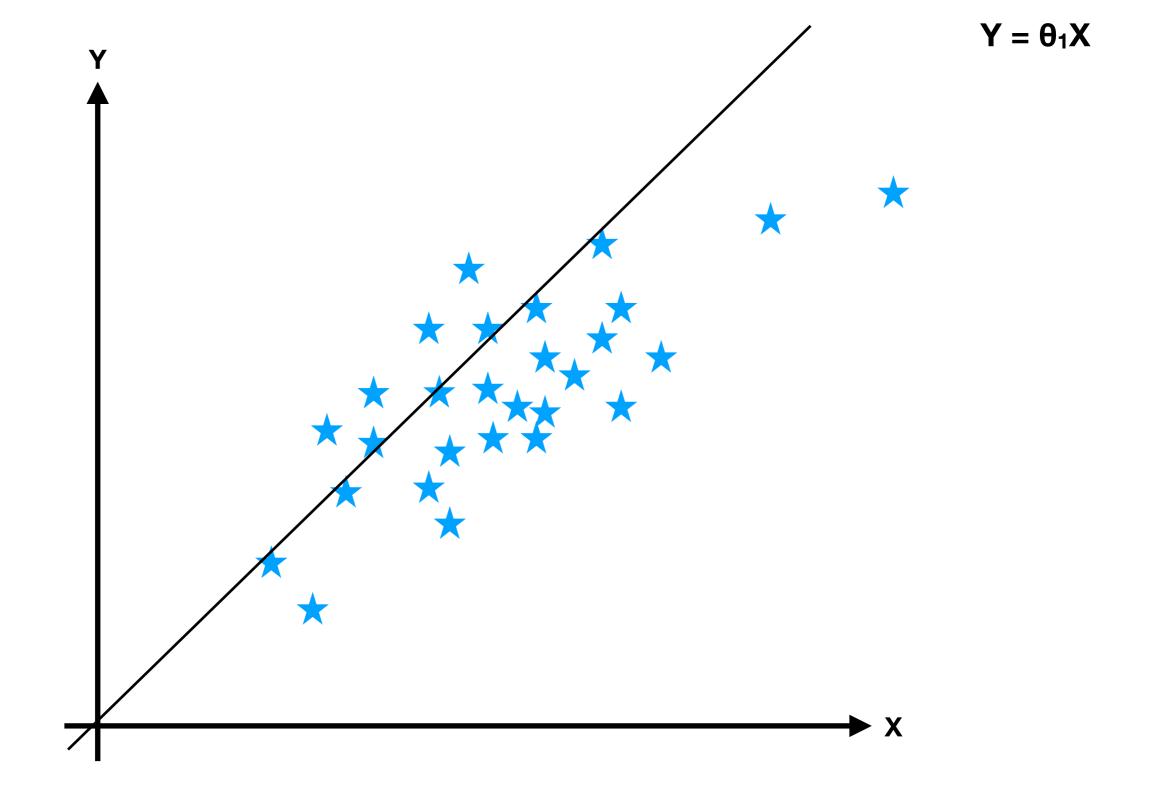


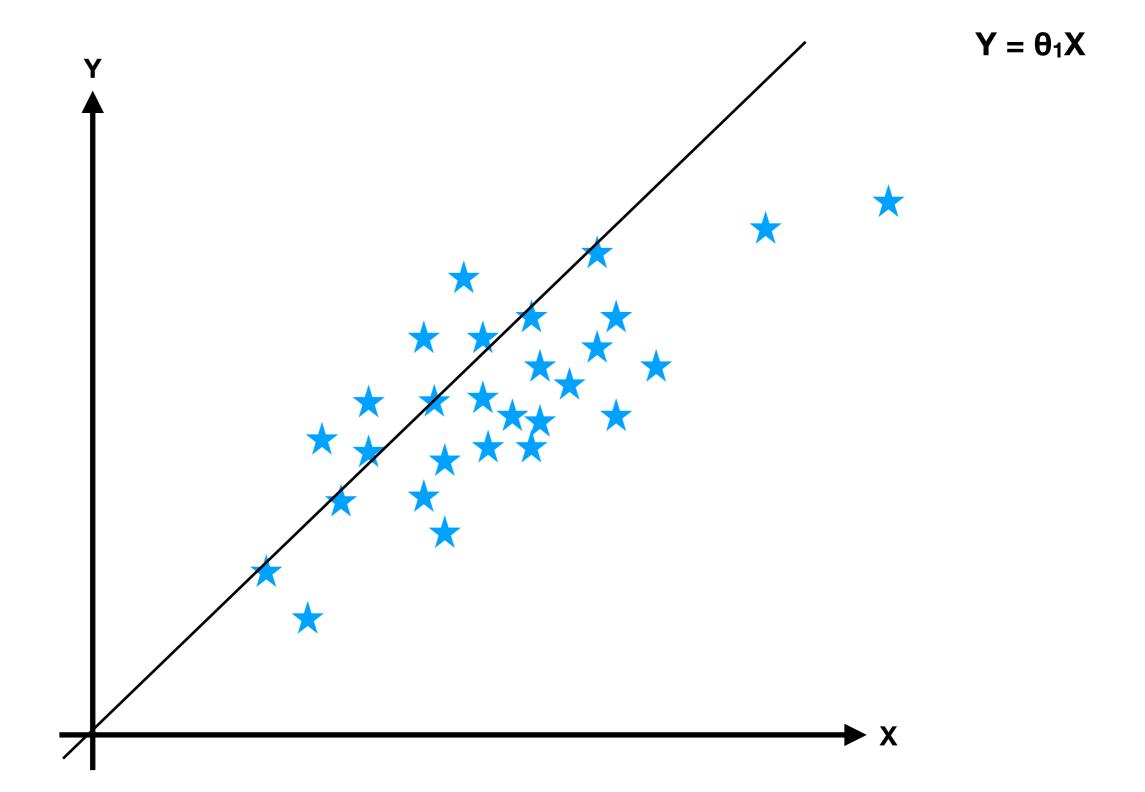


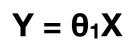


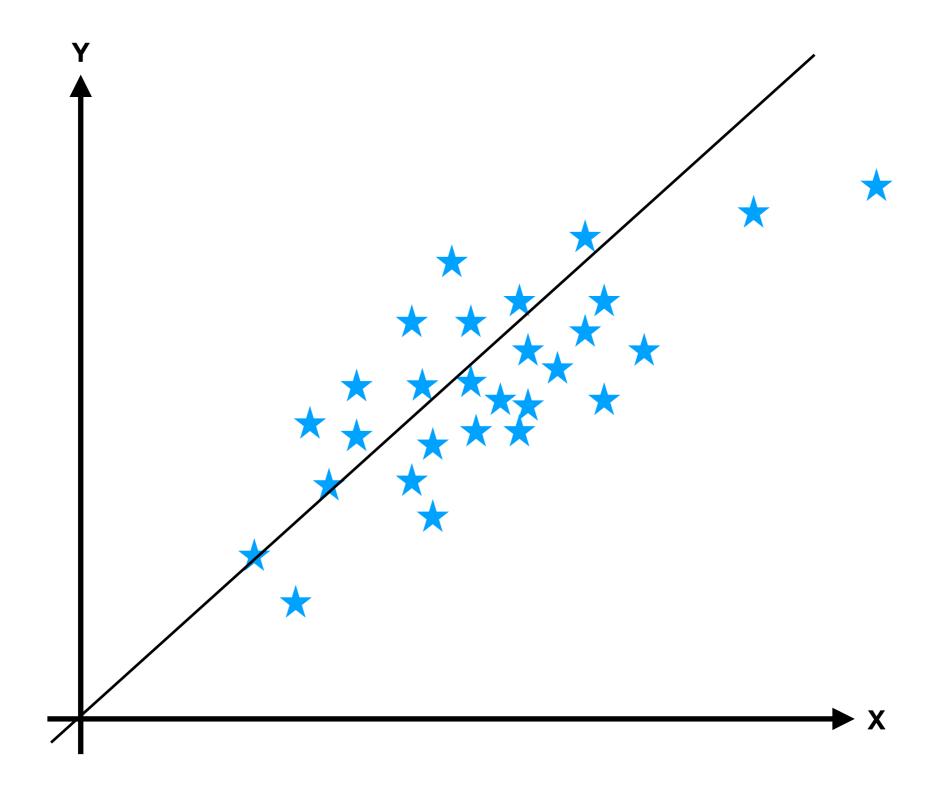


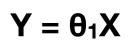


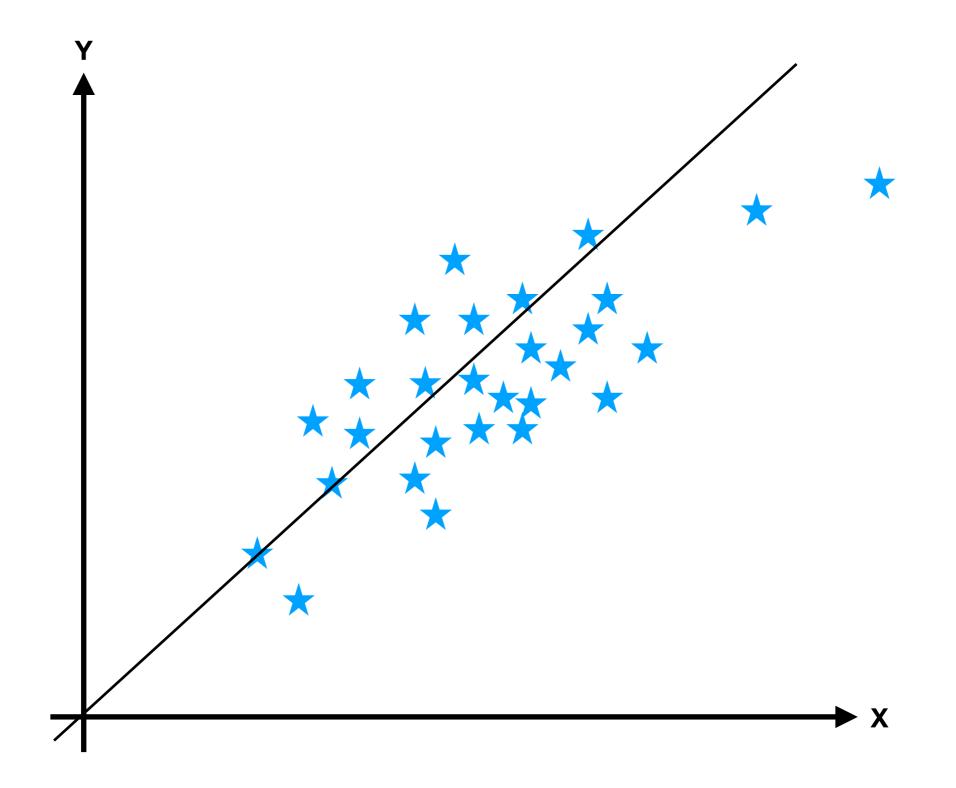


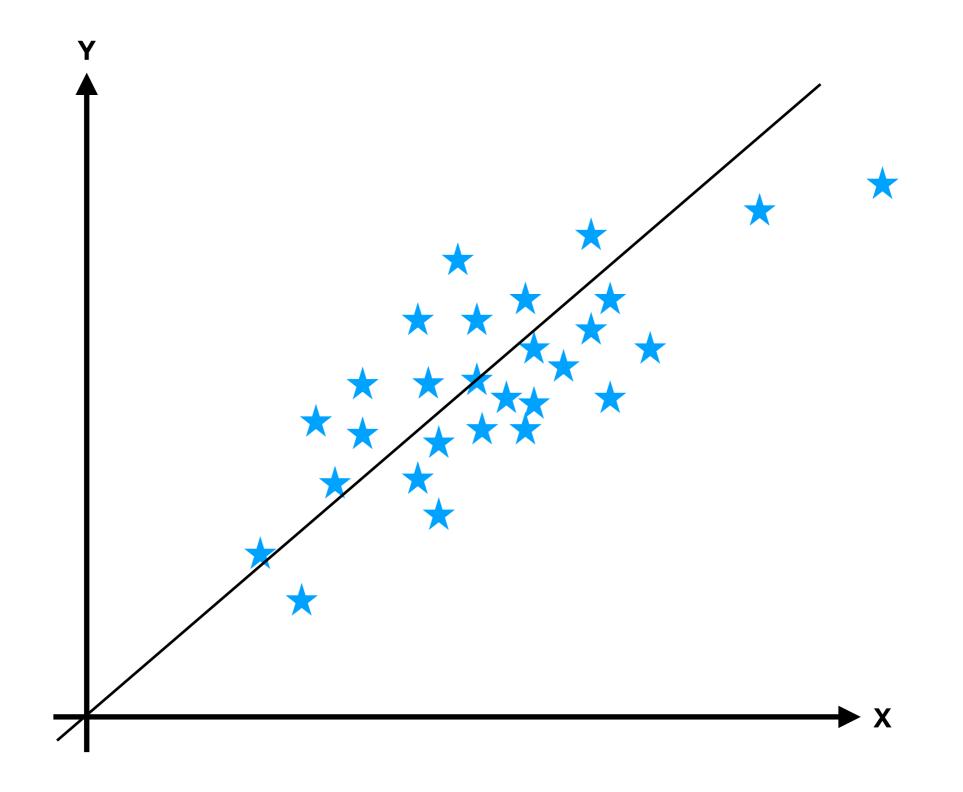


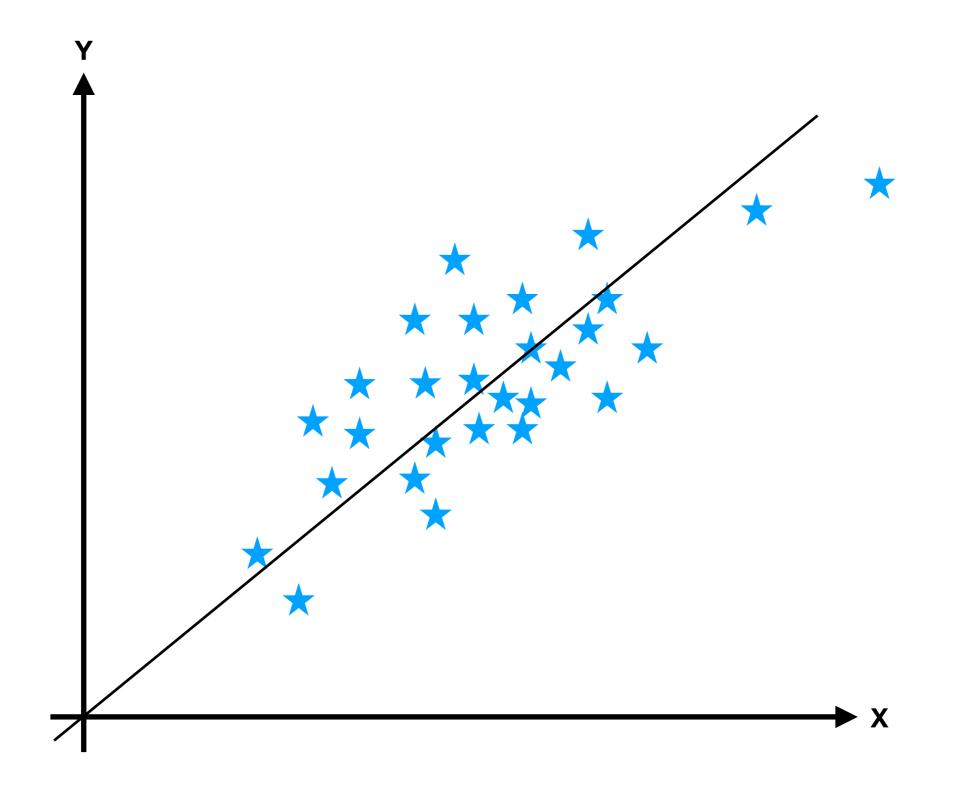


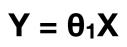


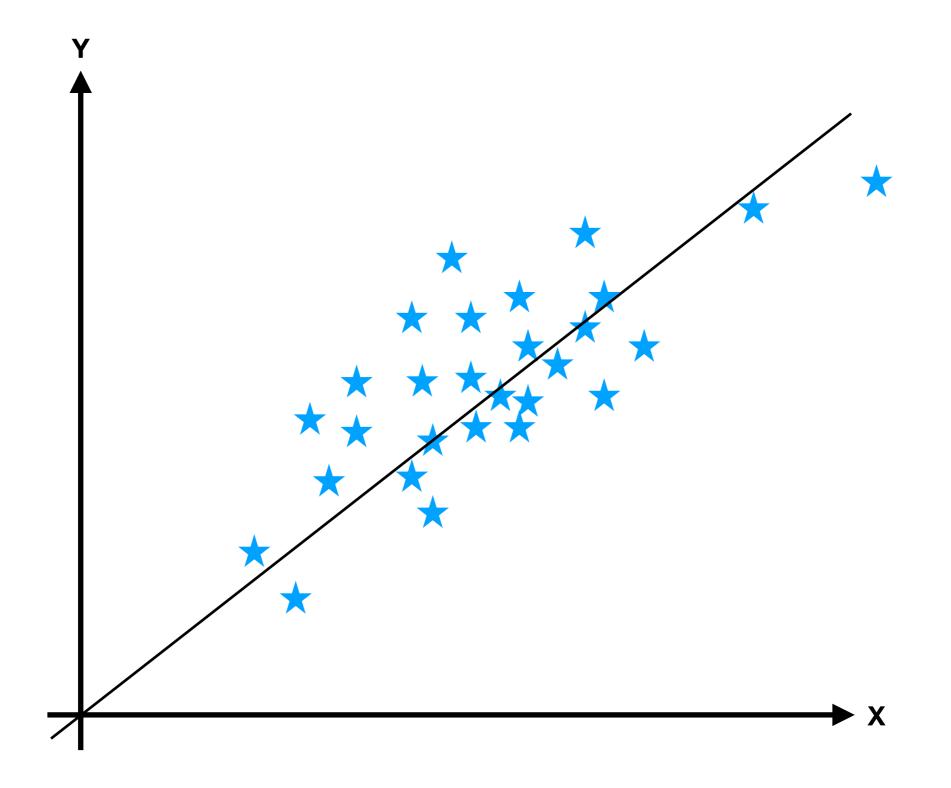






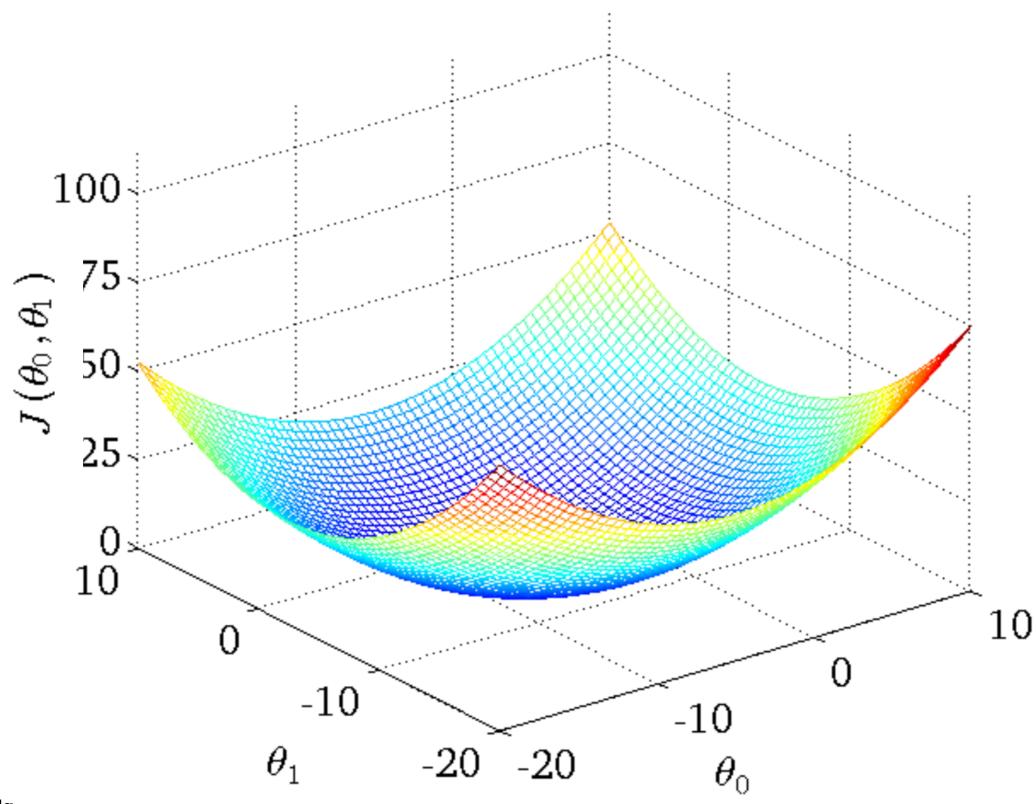






Stochastic Gradient Descent

Contour Plot



 $h_{\theta}(x)$ $J(\theta_0,\theta_1)$ (for fixed θ_0, θ_1 , this is a function of x) (function of the parameters θ_0, θ_1) 700 0.5 0.4 600 0.3 Price \$ (in 1000s) 200 200 200 200 200 500 0.2 0.1 -0.1 -0.2 -0.3 100 Training data -0.4 Current hypothesis -0.5 -1000 1000 2000 3000 4000 -500 1000 1500 0 500 2000 Size (feet²)

 θ_0

 $h_{\theta}(x)$ $J(\theta_0,\theta_1)$ (for fixed θ_0 , θ_1 , this is a function of x) (function of the parameters θ_0, θ_1) 700 0.5 0.4 600 0.3 Price \$ (in 1000s) 000 \$ 300 000 \$ 300 500 0.2 0.1 -0.1 -0.2 -0.3 100 Training data -0.4 Current hypothesis -0.5 -1000 1000 2000 3000 4000 -500 1000 1500 0 500 2000 Size (feet²) θ_0

 $h_{\theta}(x)$ $J(\theta_0, \theta_1)$ (for fixed θ_0 , θ_1 , this is a function of x) (function of the parameters θ_0, θ_1) 700 0.5 0.4 600 0.3 Price \$ (in 1000s) 200 200 200 200 200 500 0.2 0.1 -0.1 -0.2 -0.3 100 Training data -0.4 Current hypothesis 0 -0.5 -1000 1000 2000 3000 4000 -500 500 1000 1500 2000 0 Size (feet²) θ_0

