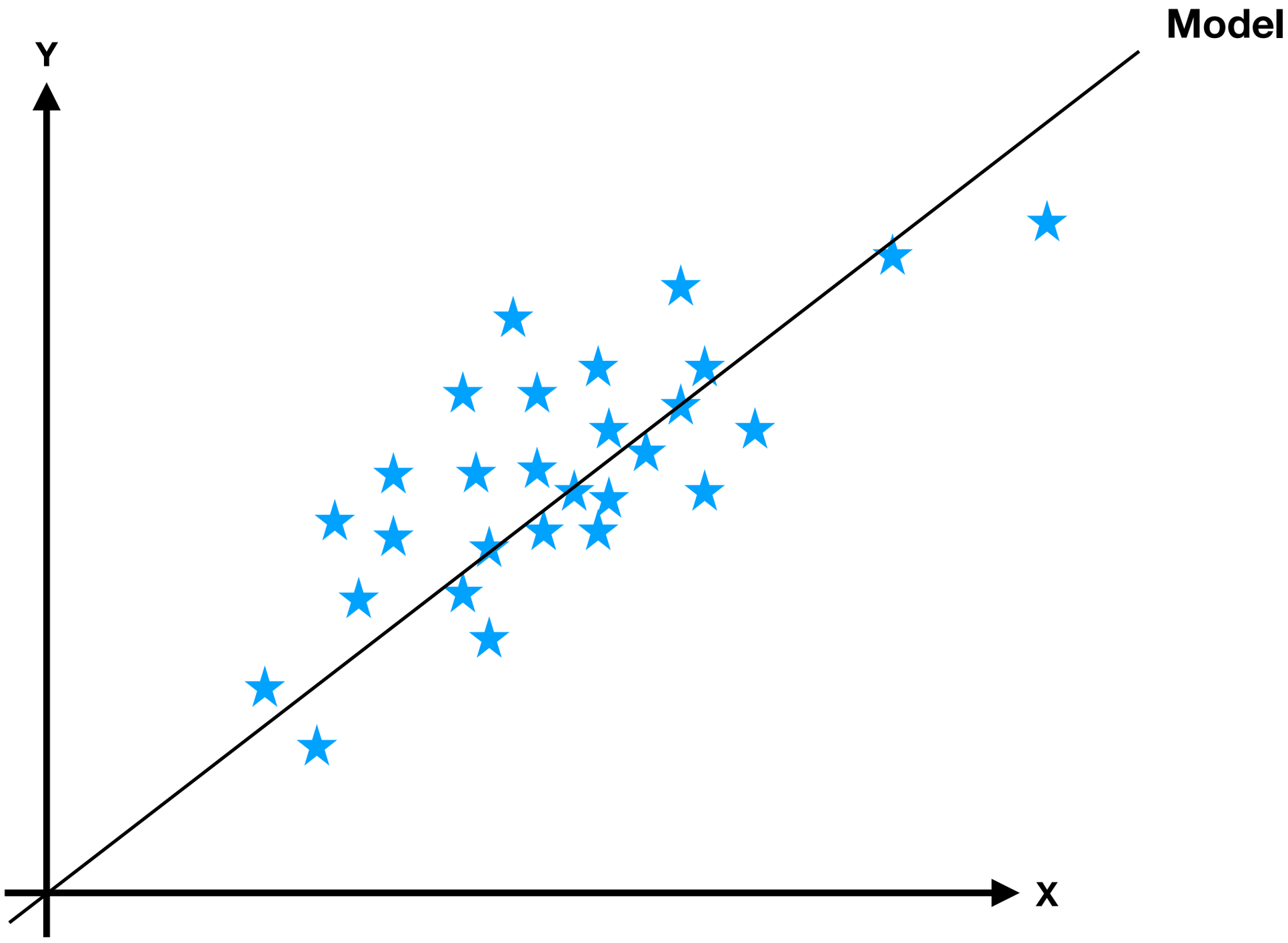
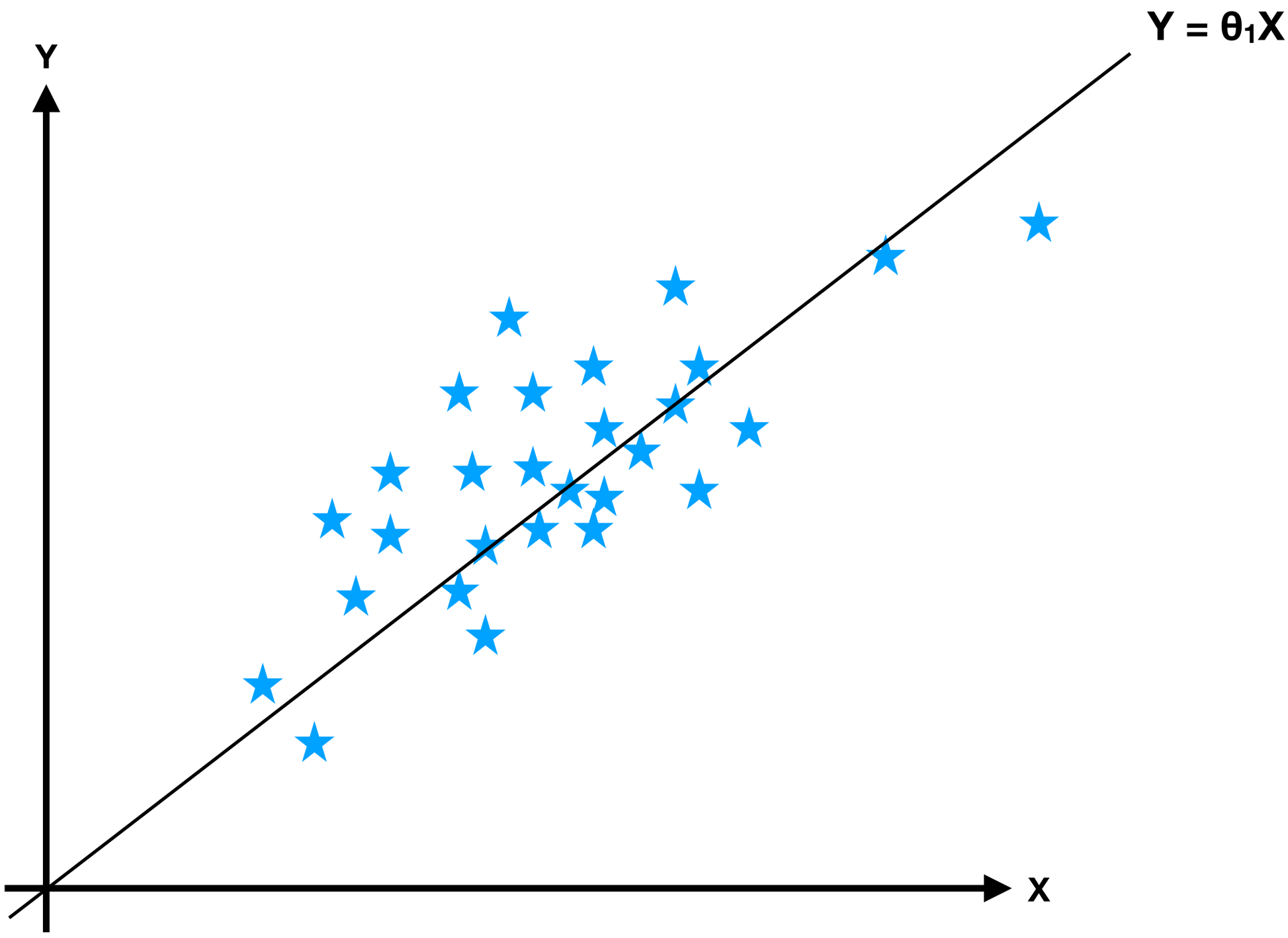
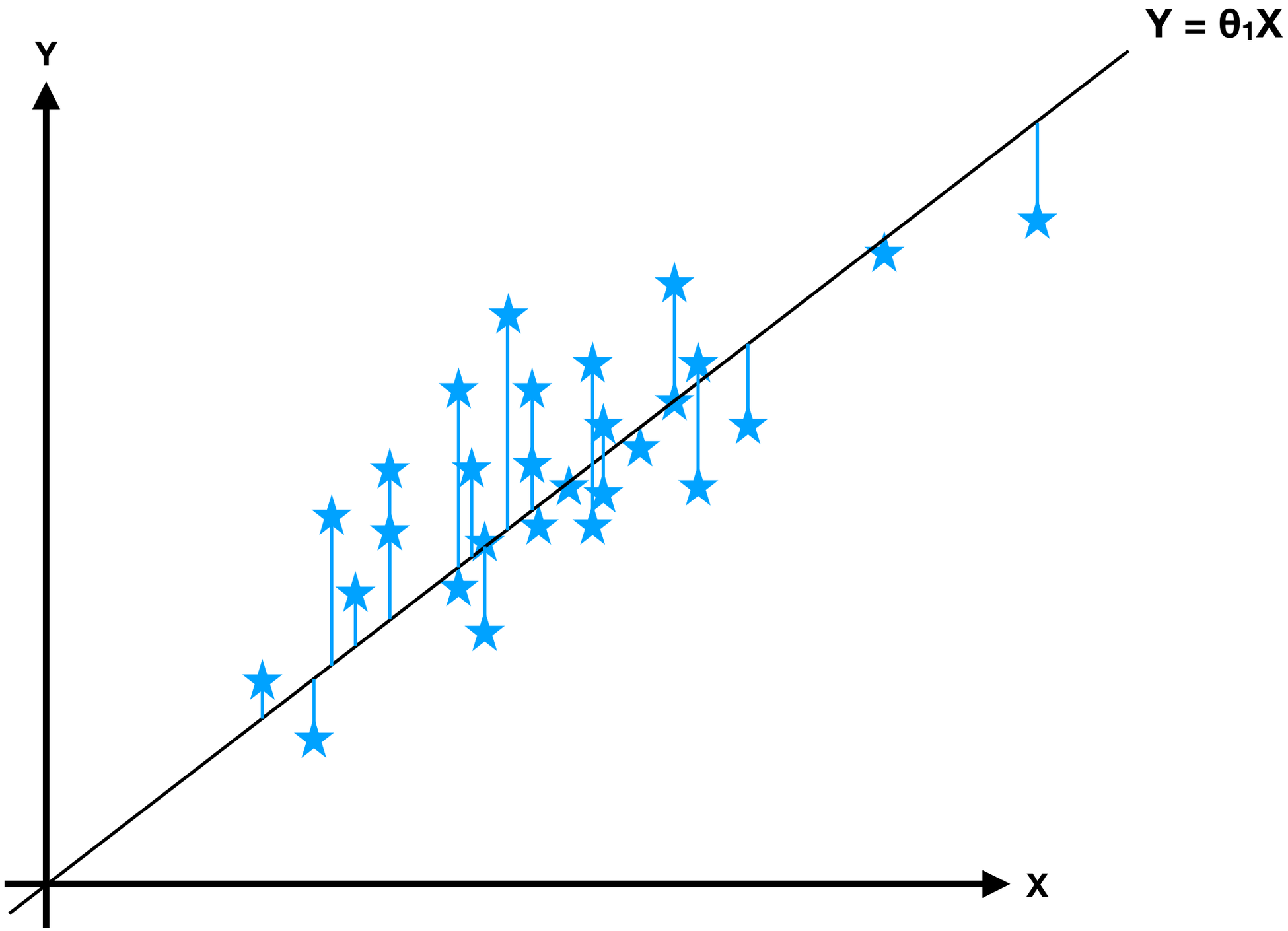
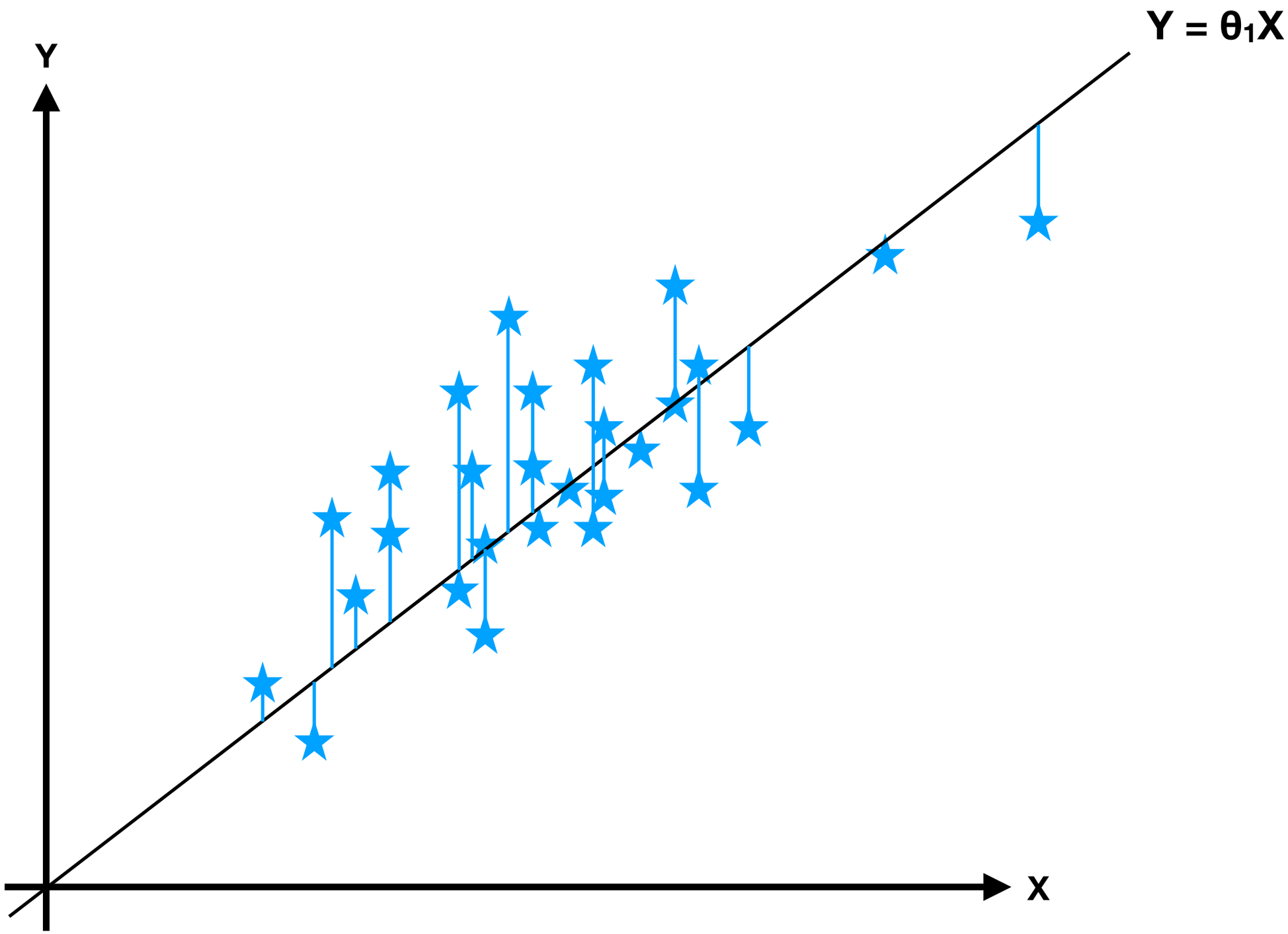
SC201

Lecture 3

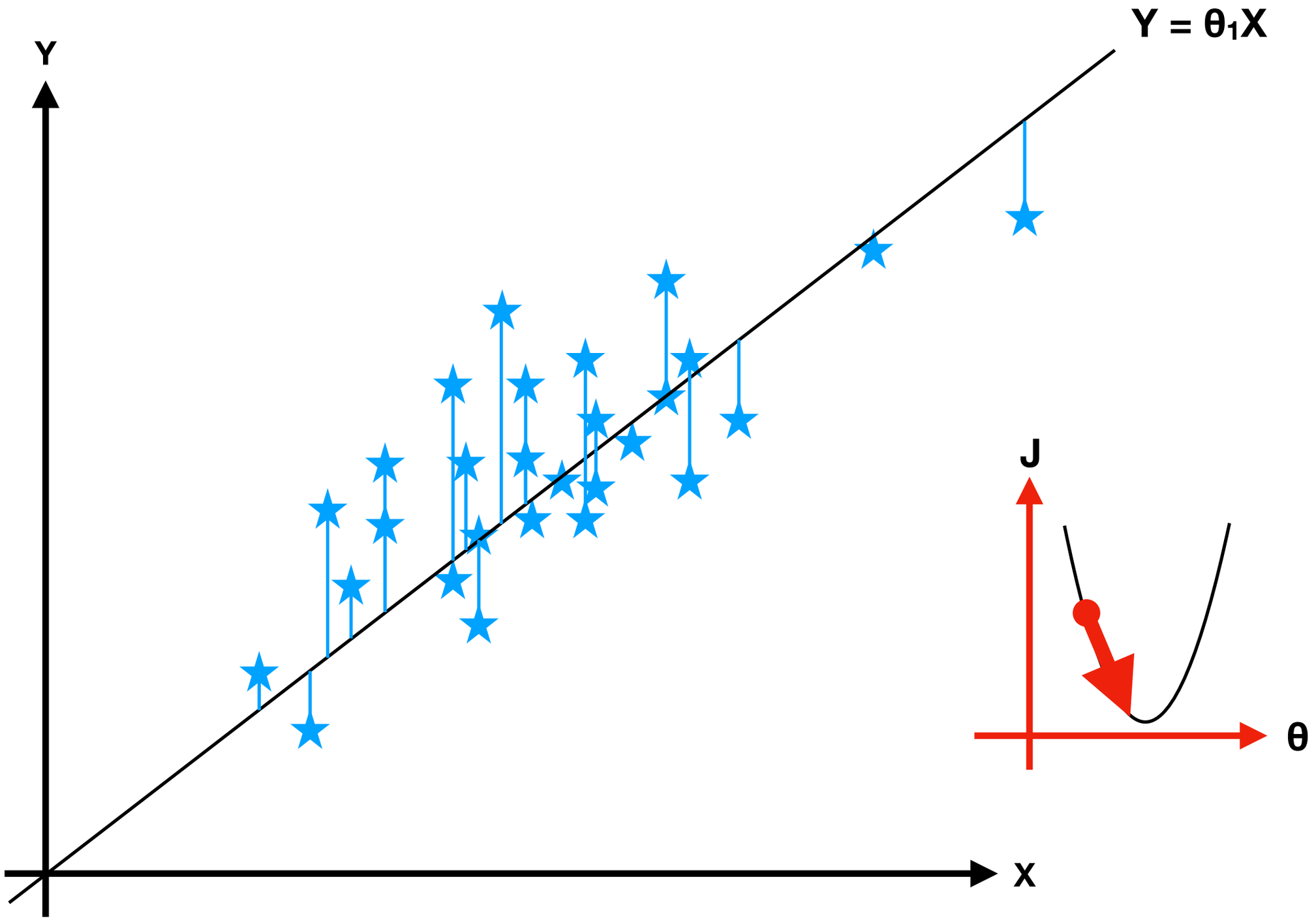

Model
Y
X


Y = θ1X
Y
X

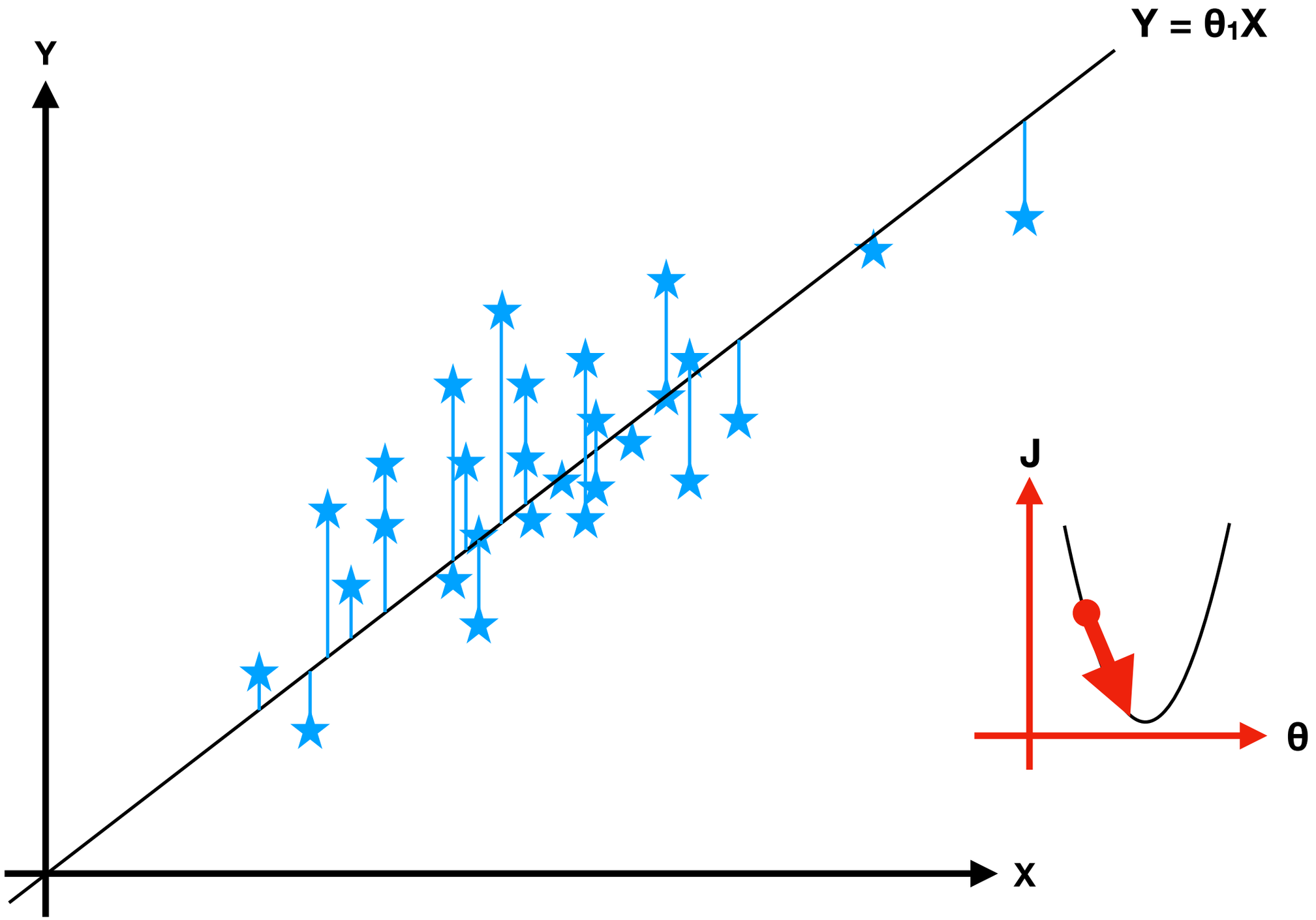

Y = θ1X
Y
X



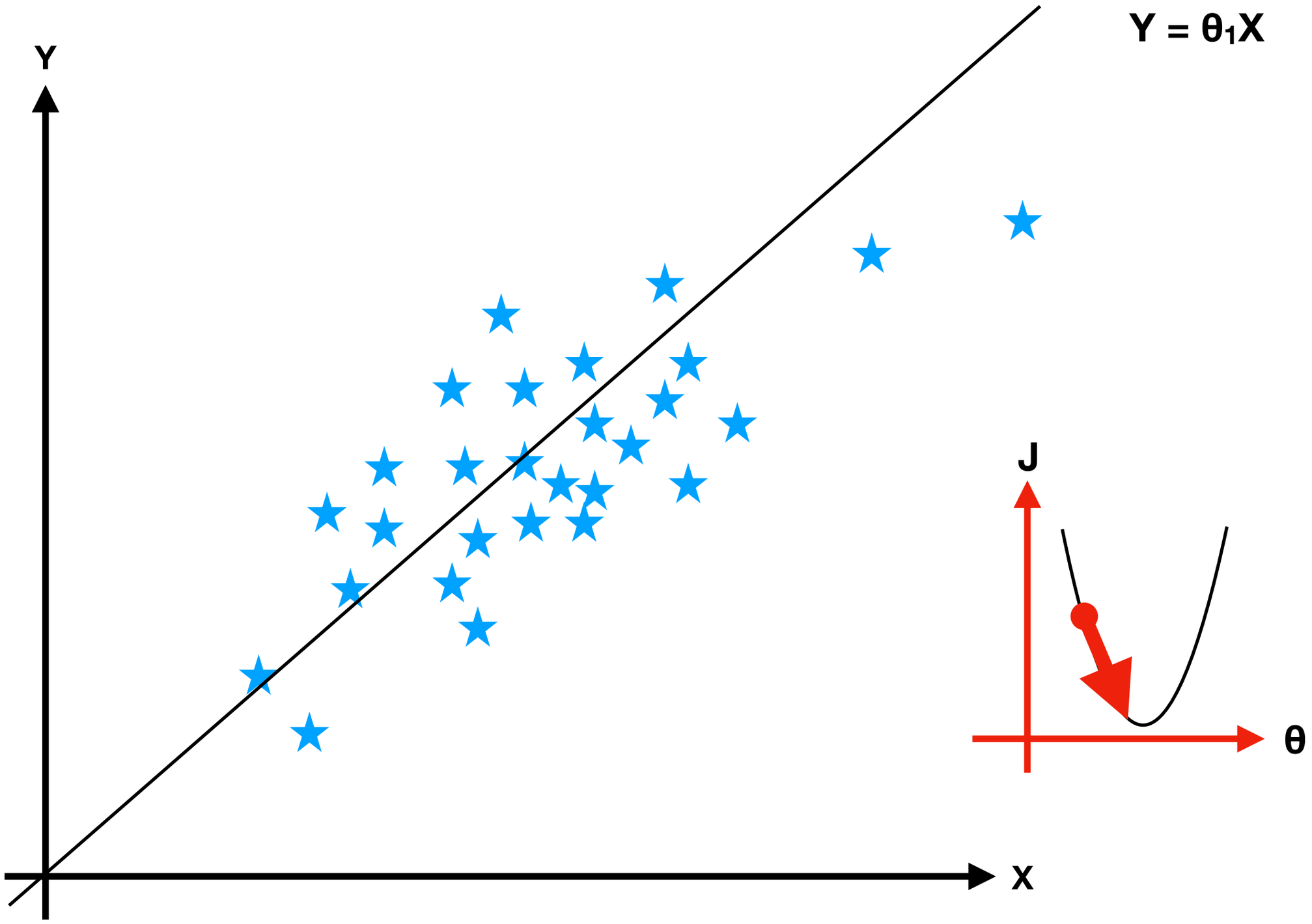
$$J = \Sigma (Y' - Y_i)^2 / 2m$$



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

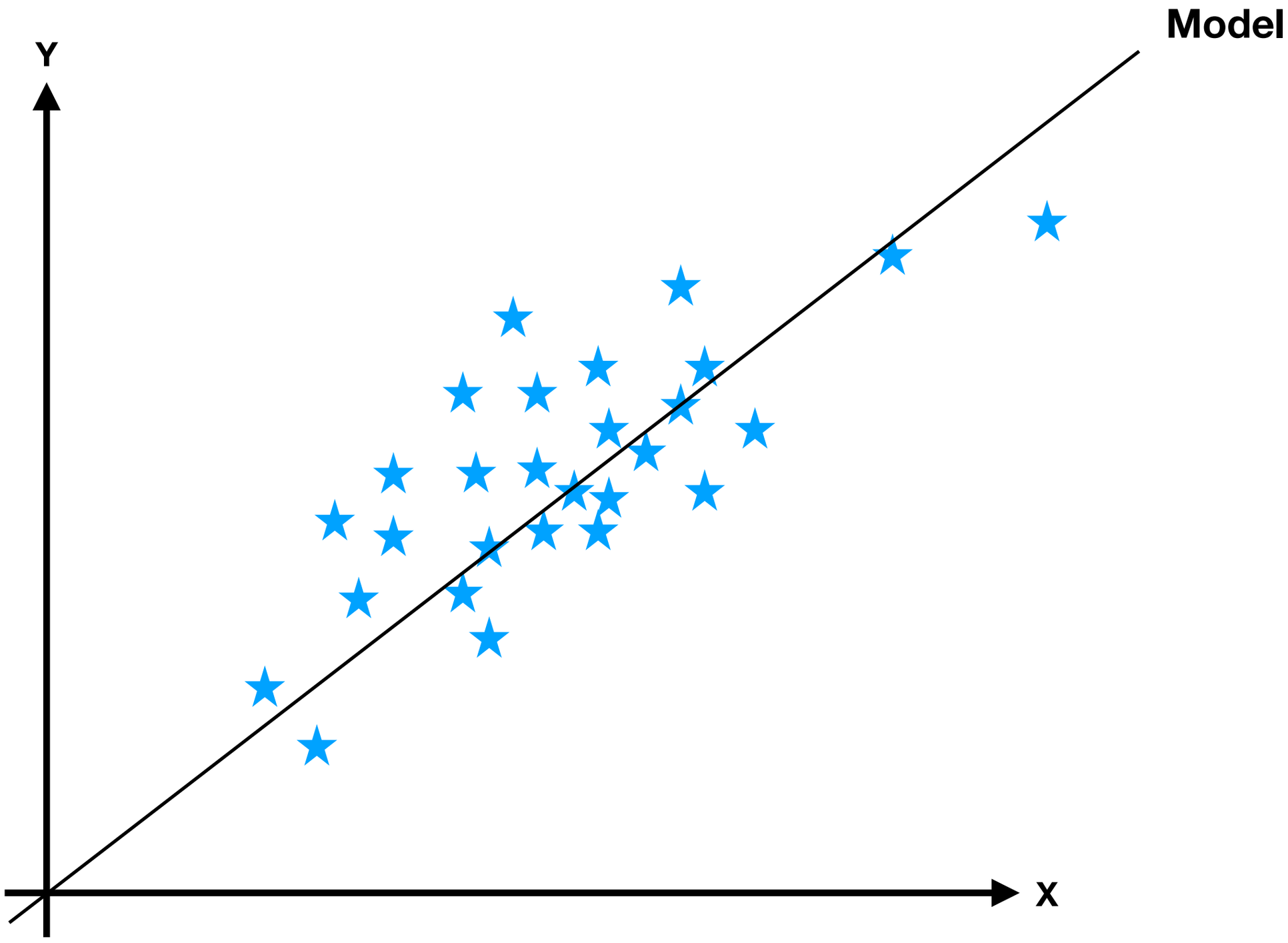


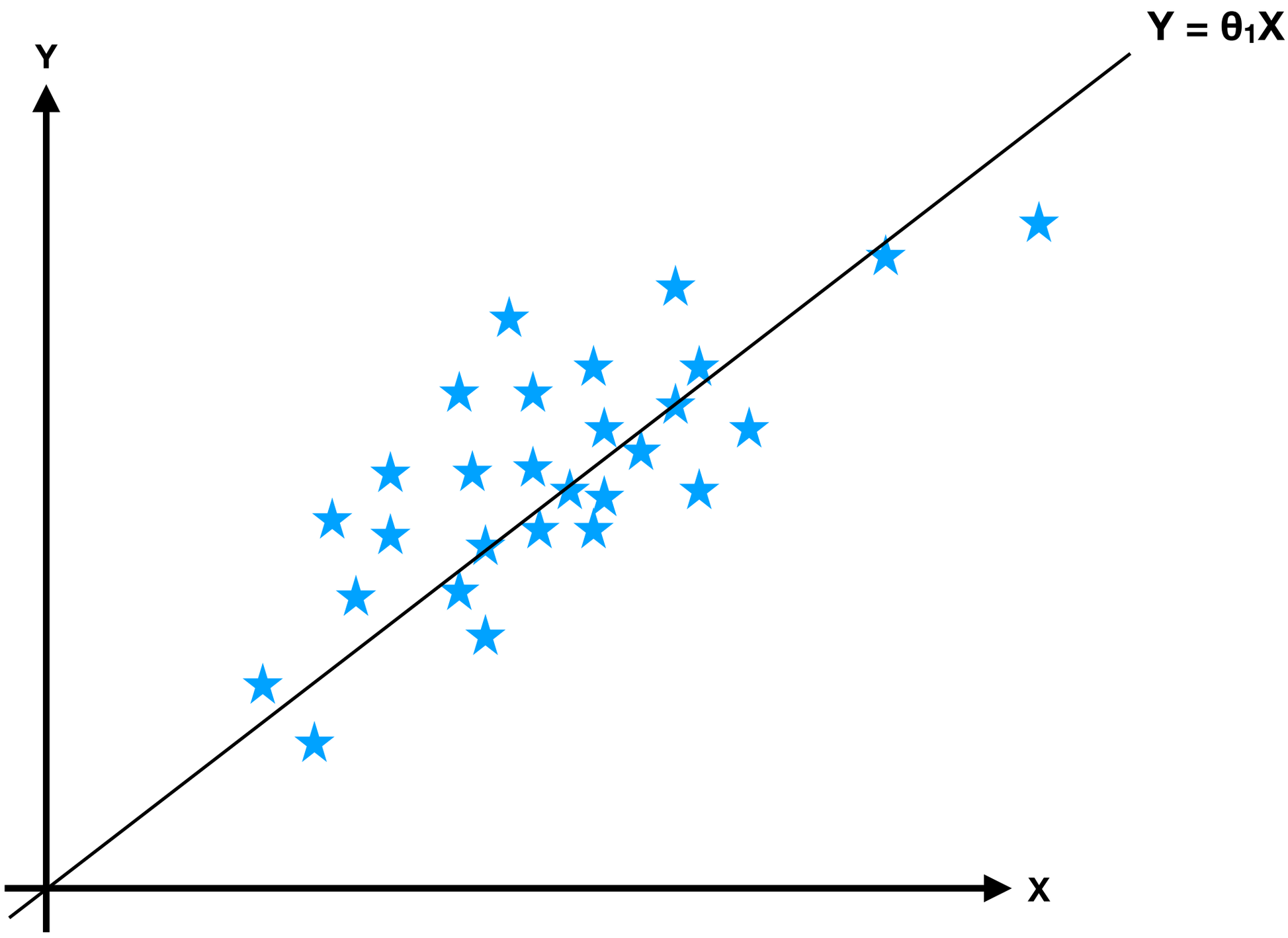
$$\theta = \theta - \alpha \left(\Sigma \left(Y' - Y_i\right)X_i / m\right)$$

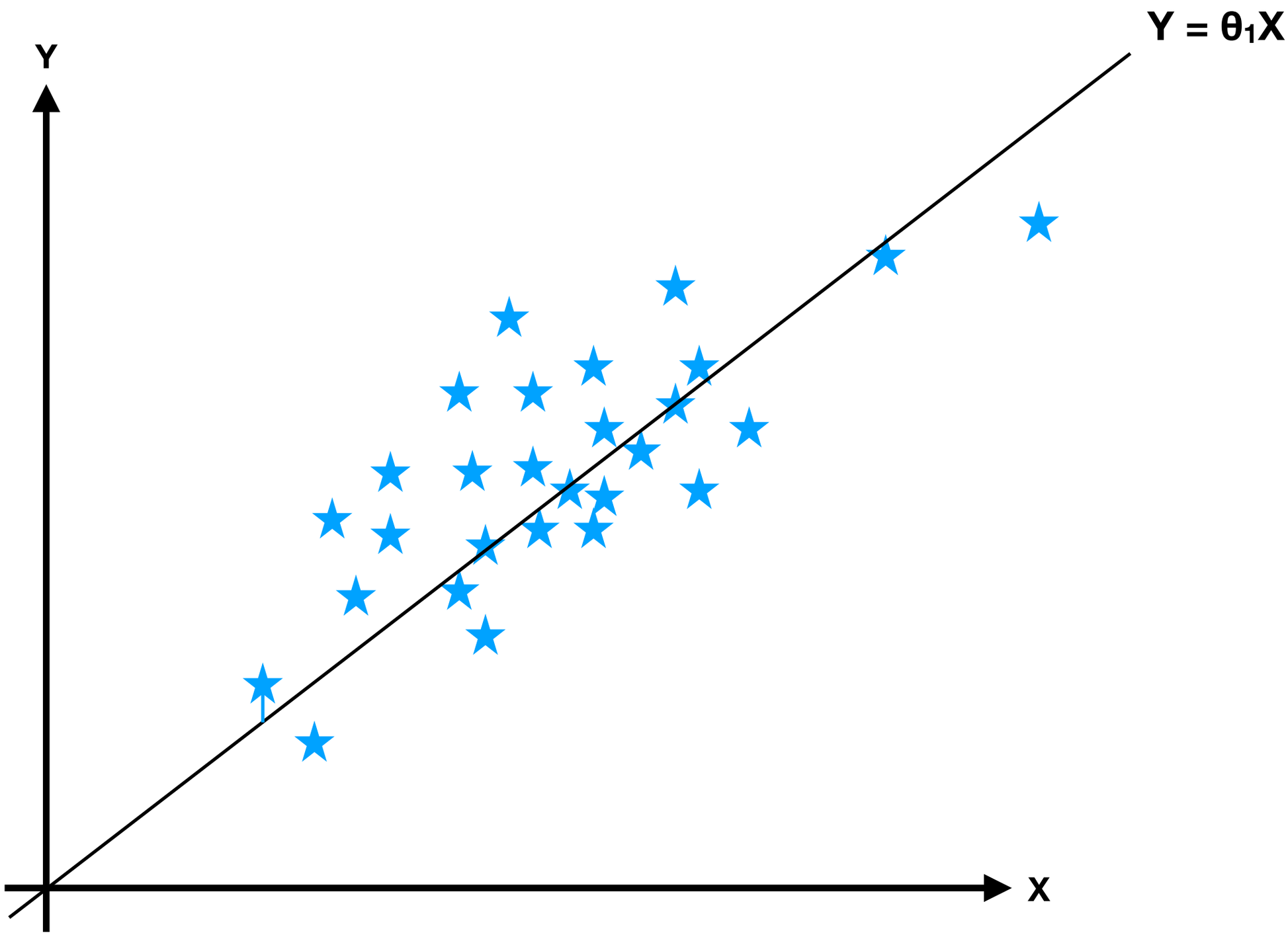


$$\theta = \theta - \alpha \left(\Sigma \left(Y' - Y_i\right)X_i / m\right)$$

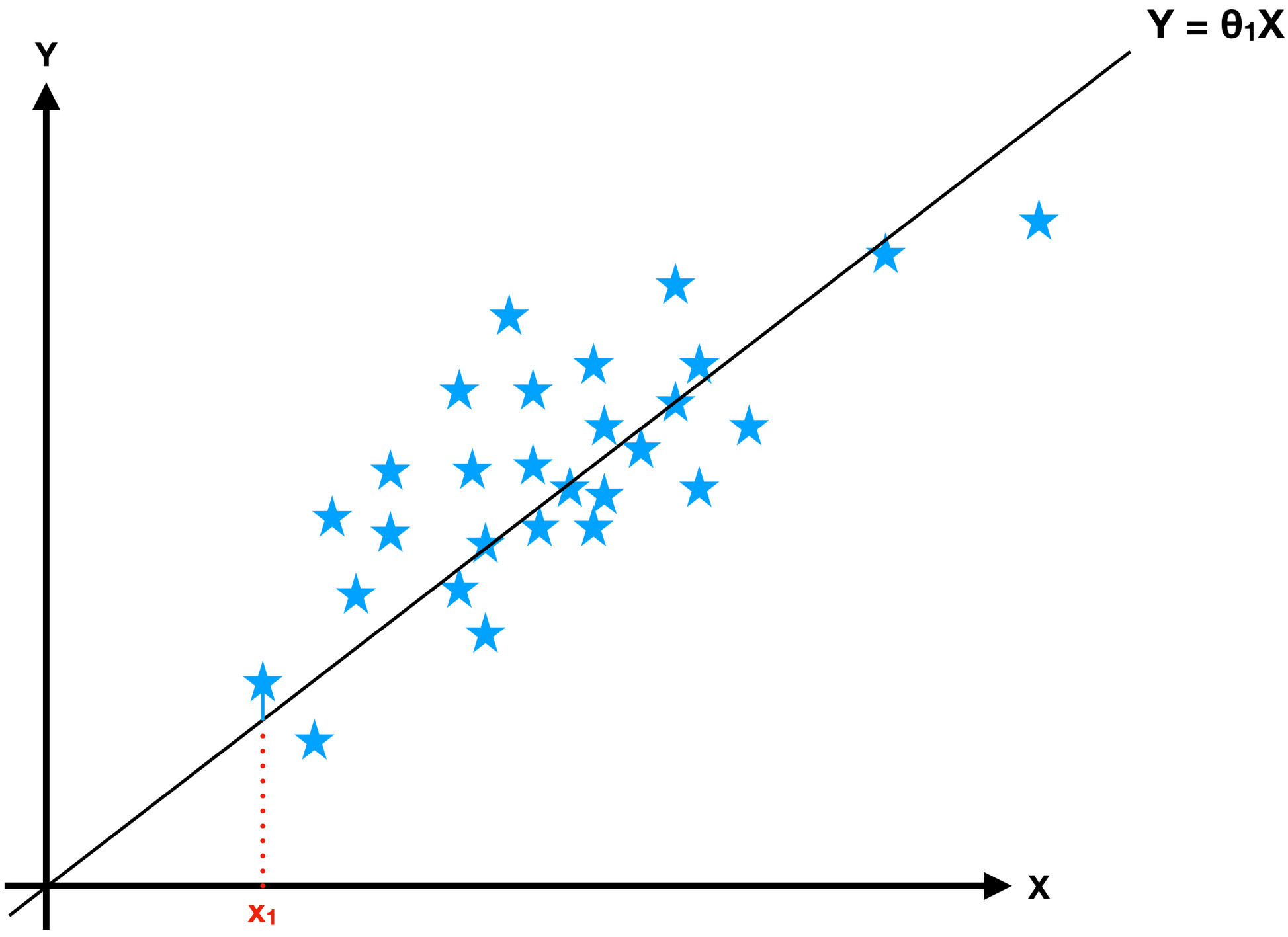
Batch Gradient Descent


Model
Y
X

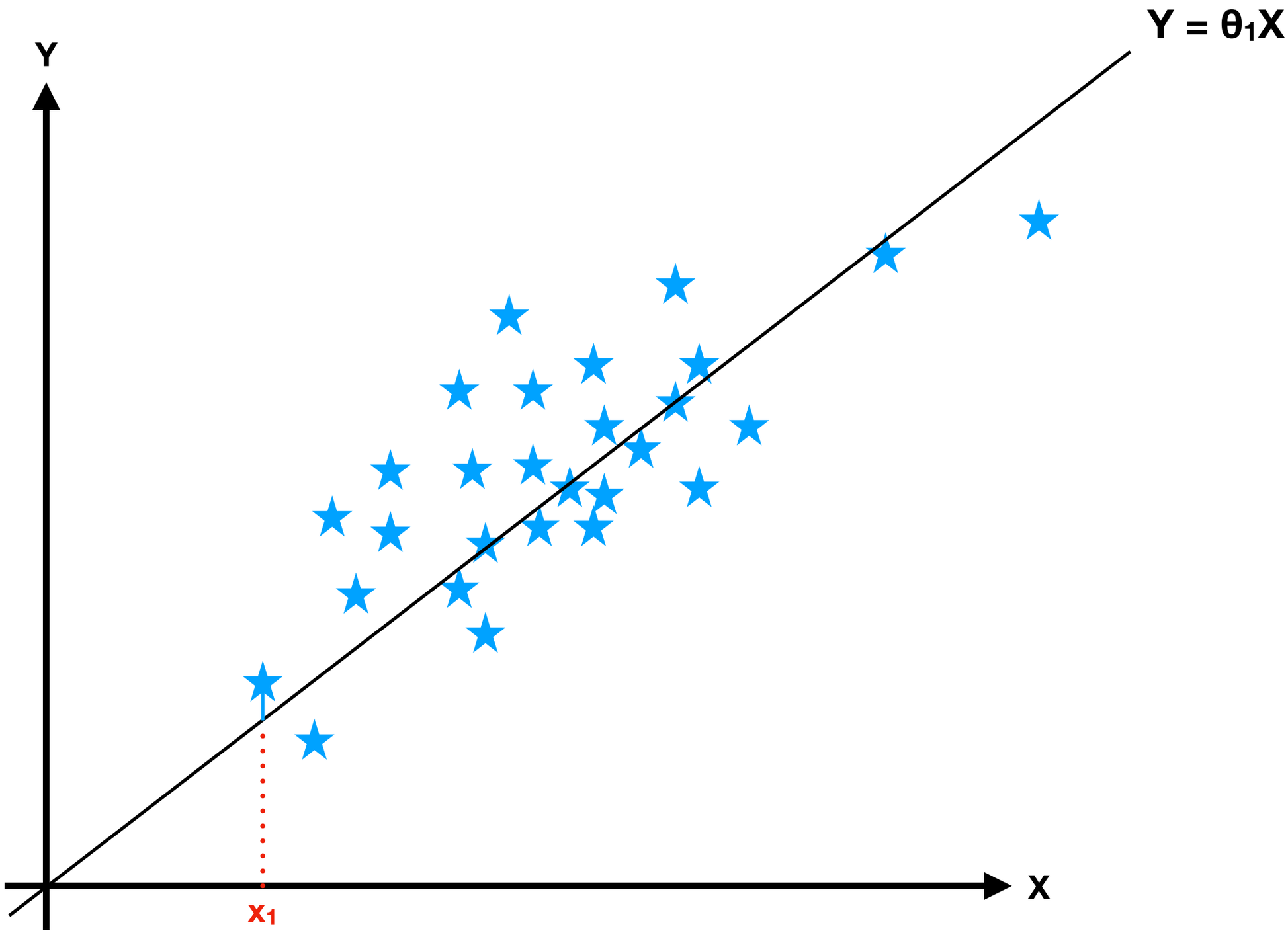

Y = θ1X
Y
X



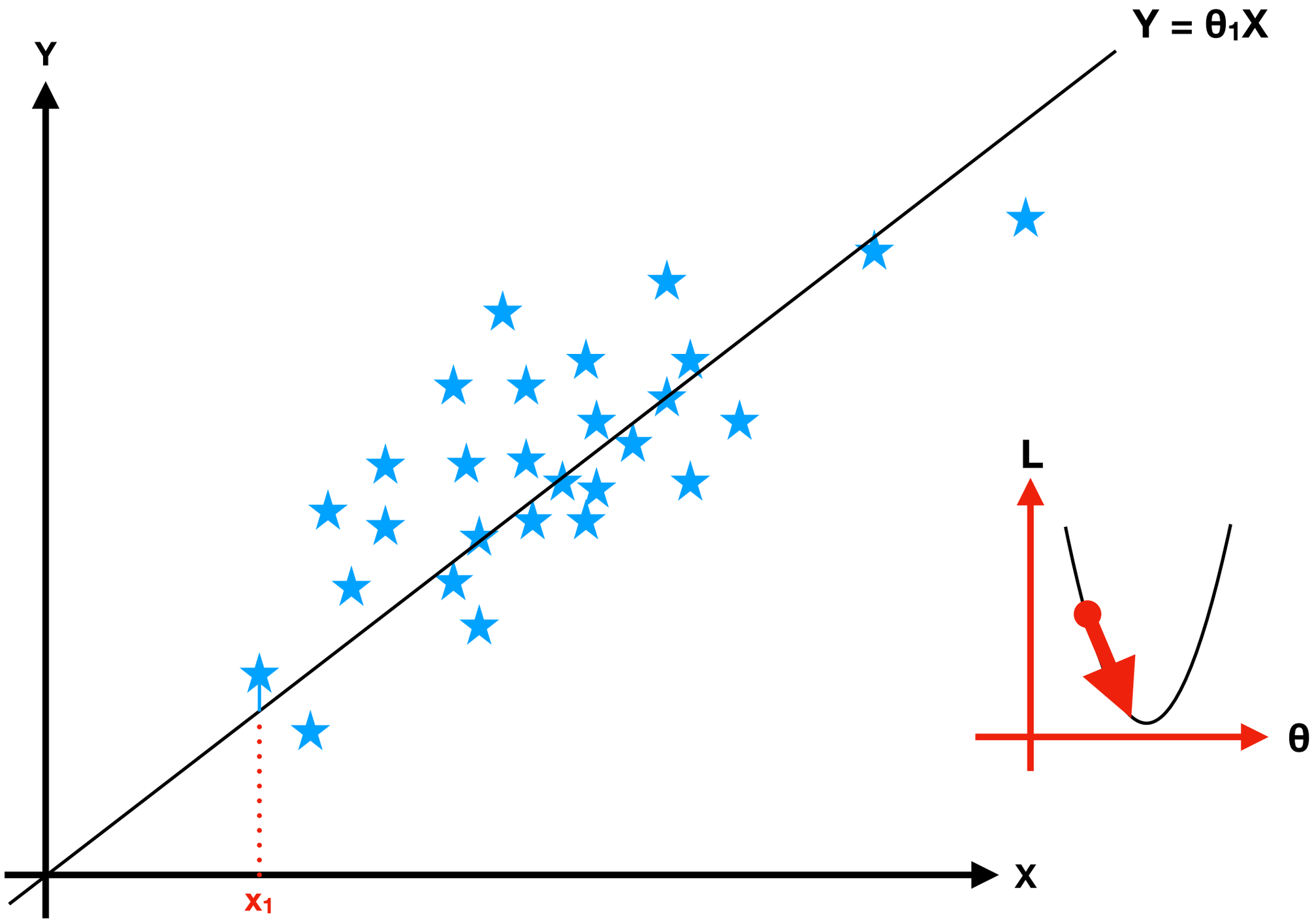
$$L = (y' - y_1)^2$$



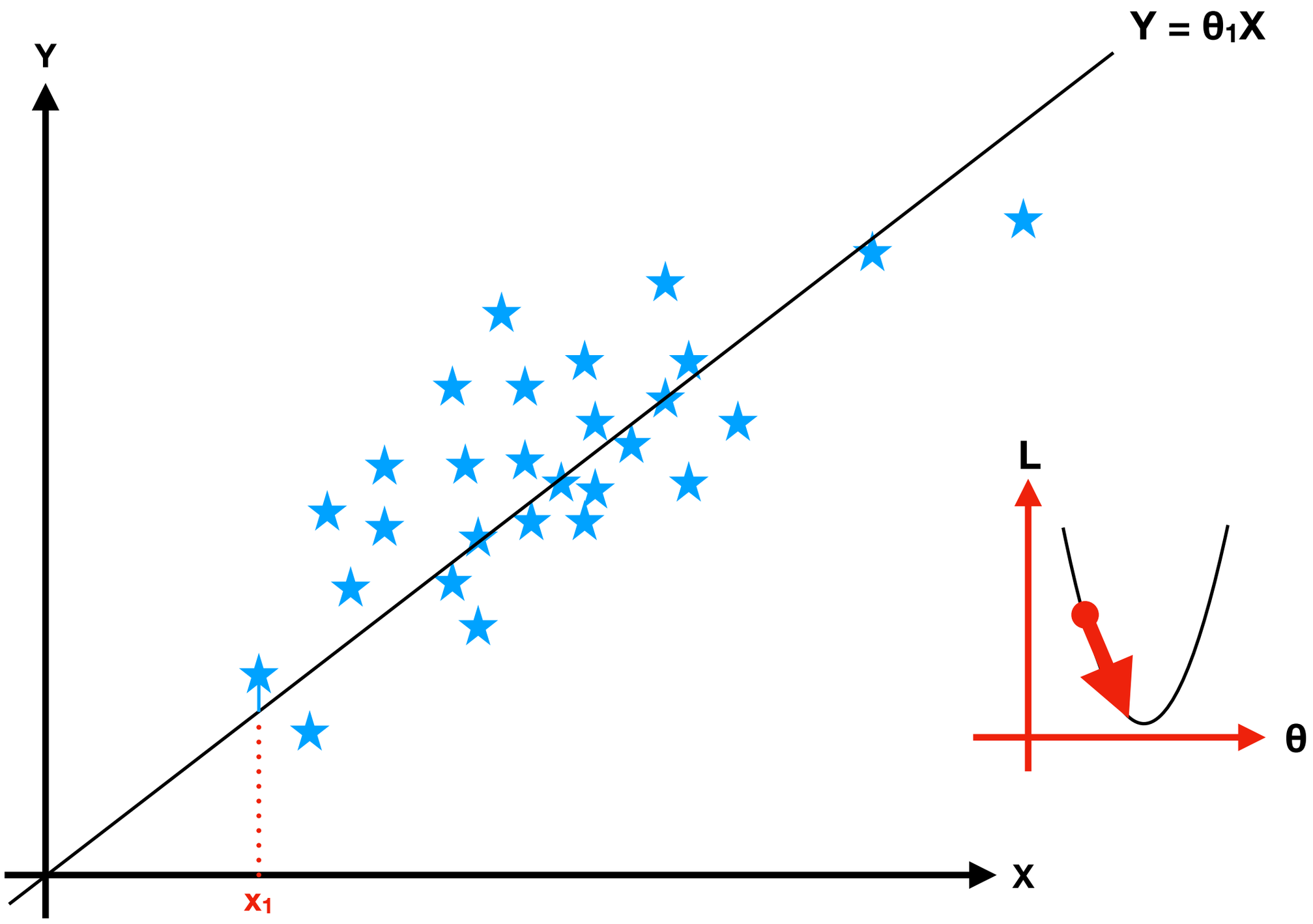
$$L = (y' - y_1)^2$$



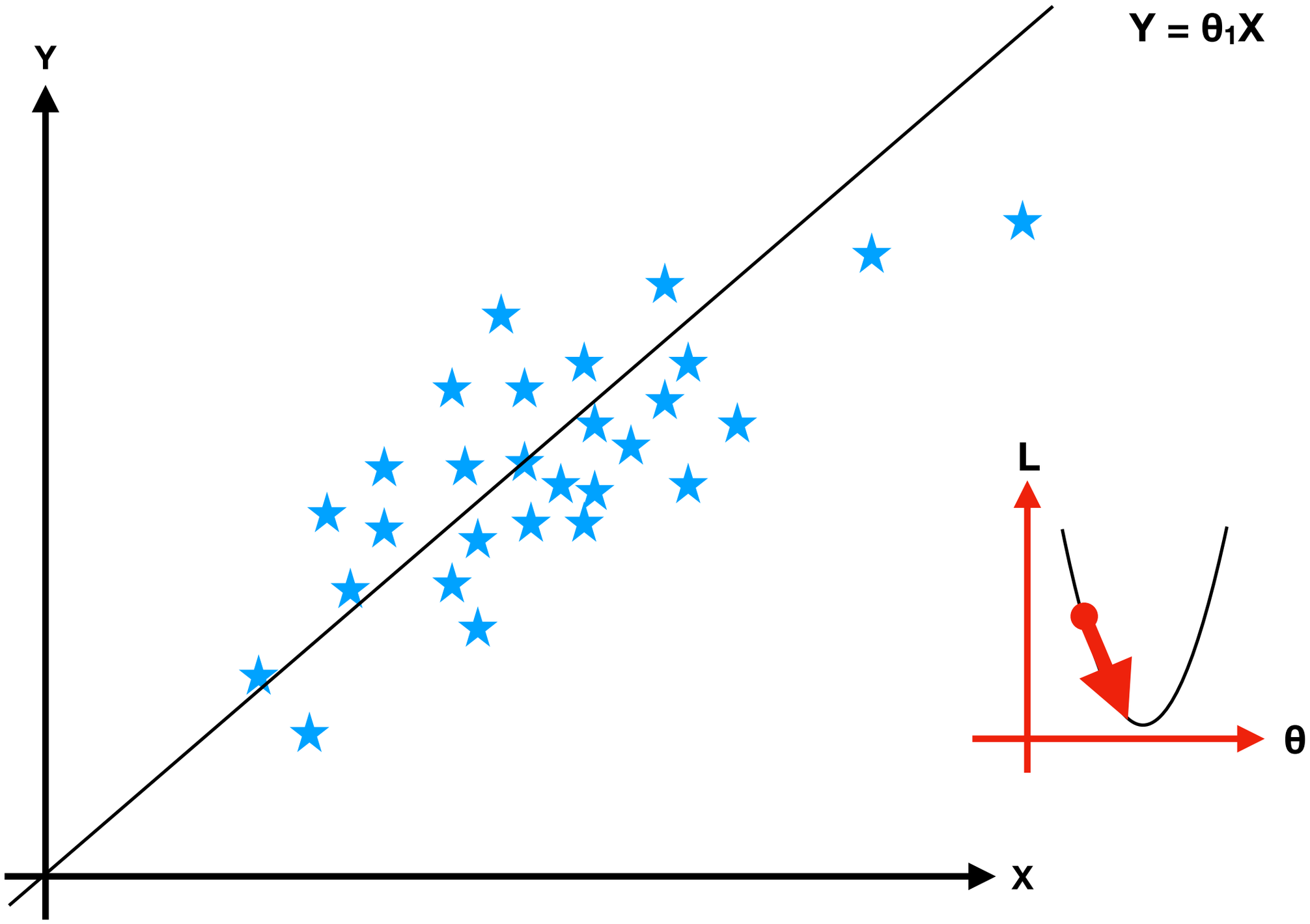
$$L = (\theta_1 x_1 - y_1)^2$$



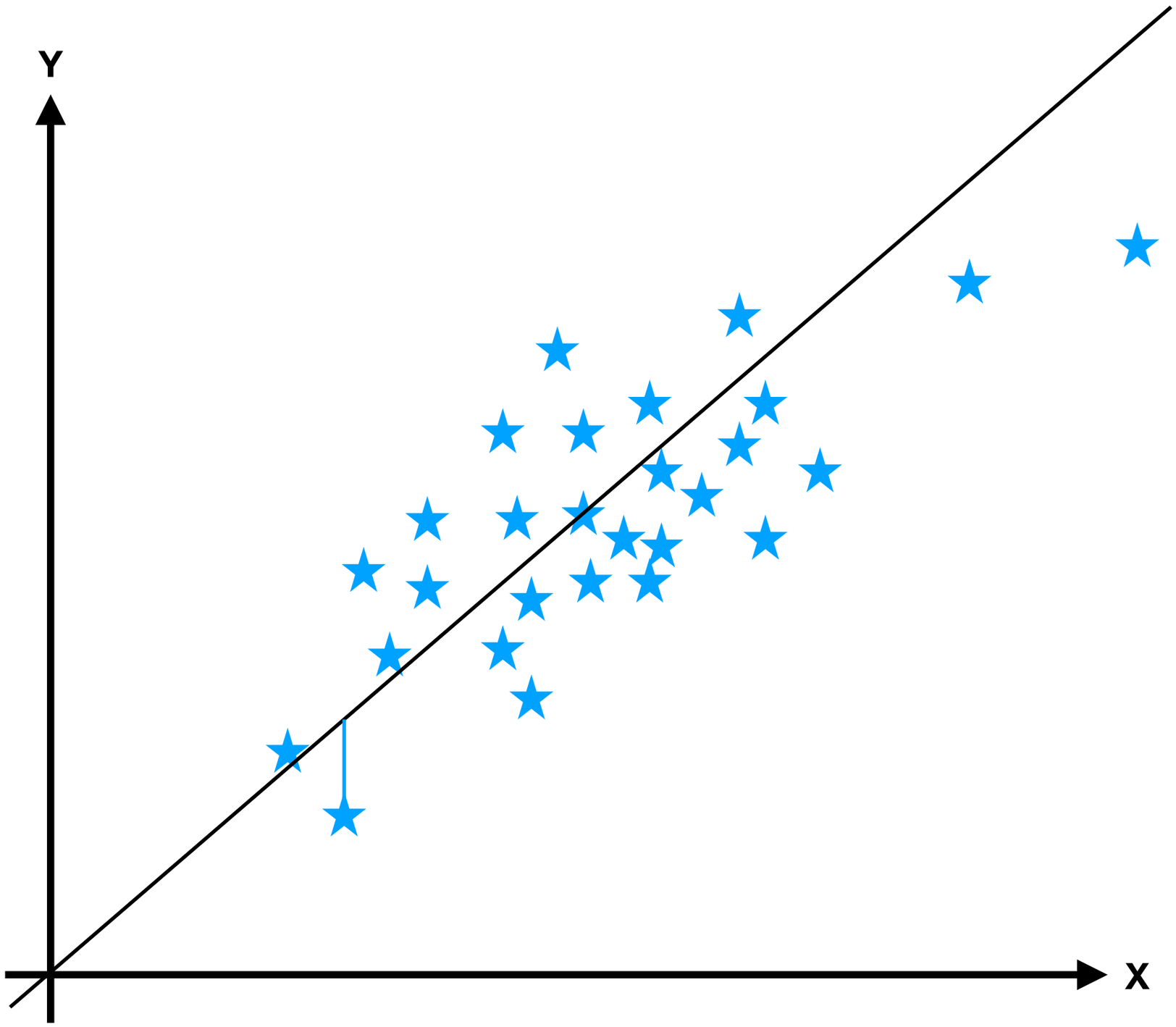
$$L = (\theta_1 x_1 - y_1)^2$$



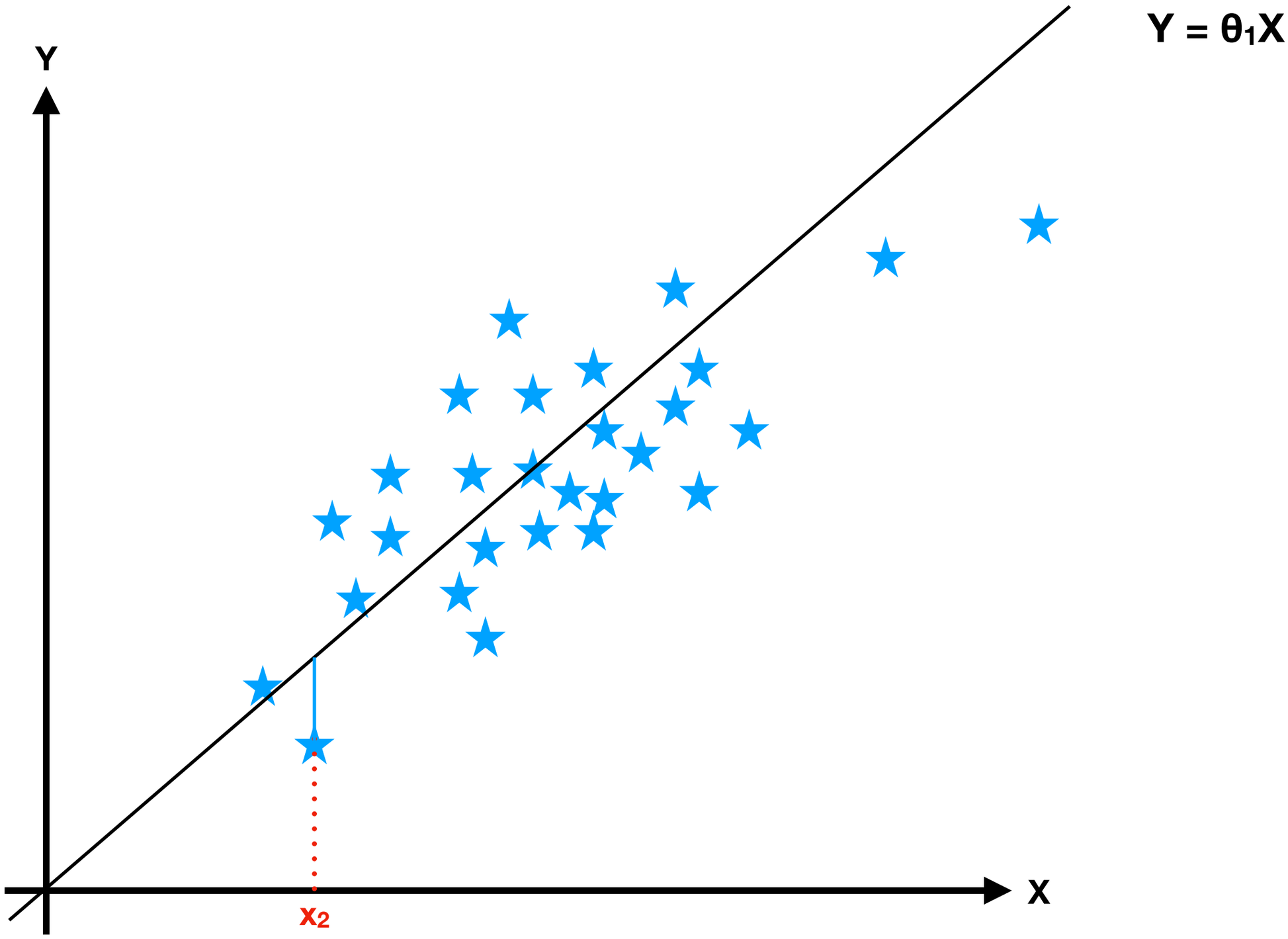
$$\theta = \theta - \alpha\,(2(\theta_1 x_1 - y_1)x_1)$$



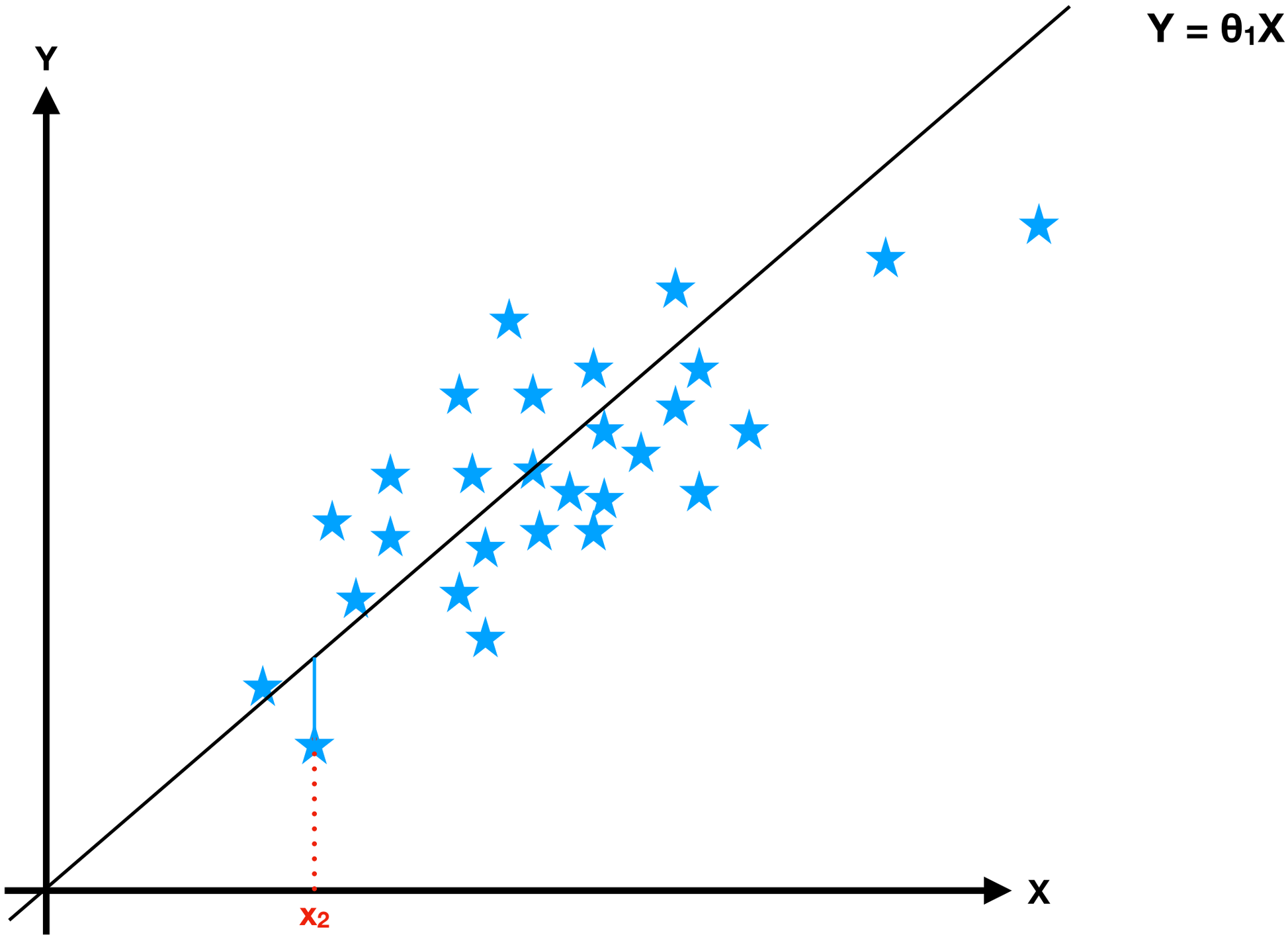
$$\theta = \theta - \alpha \left(2(\theta_1 x_1 - y_1)x_1\right)$$



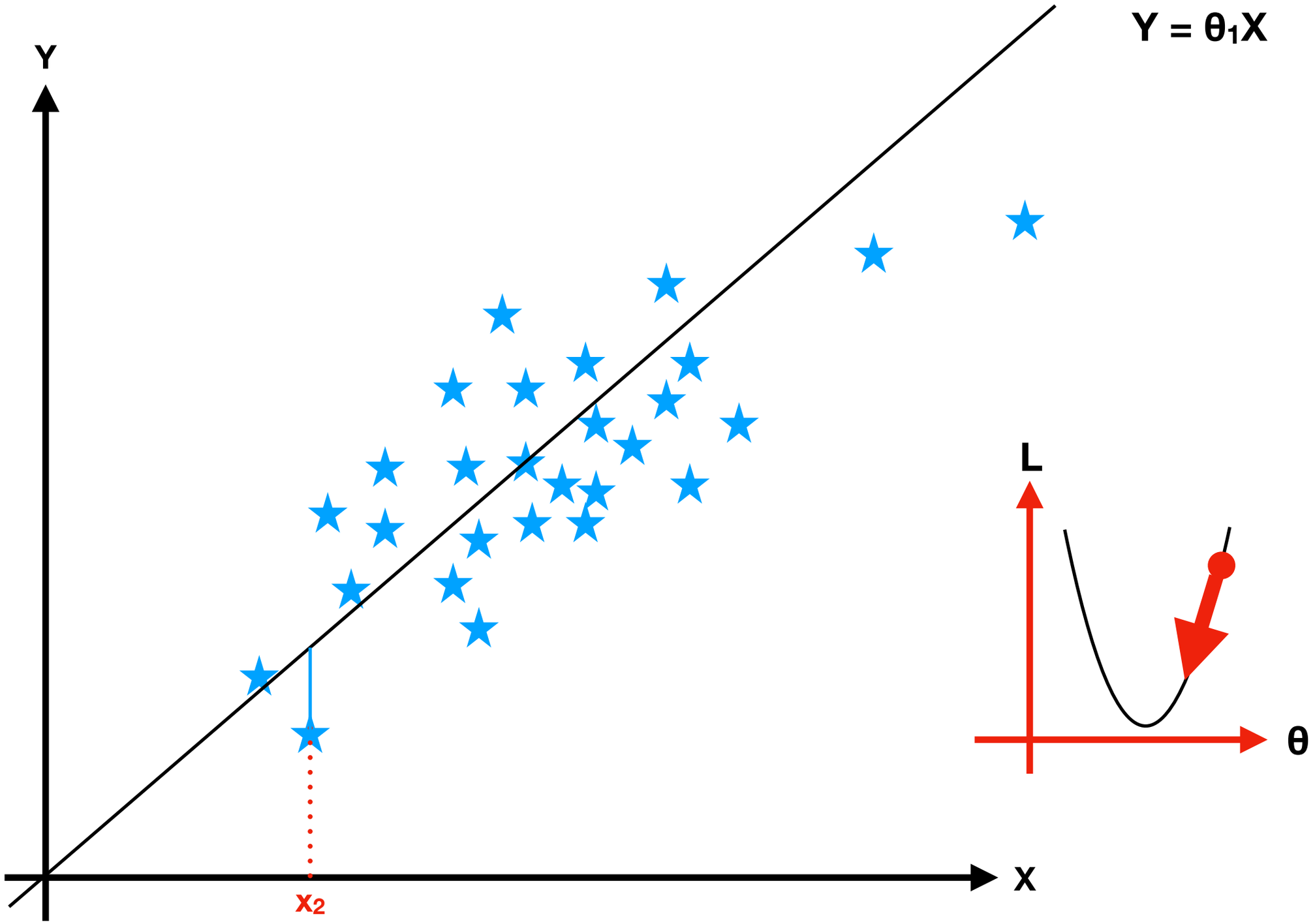
$$L = (y' - y_2)^2$$



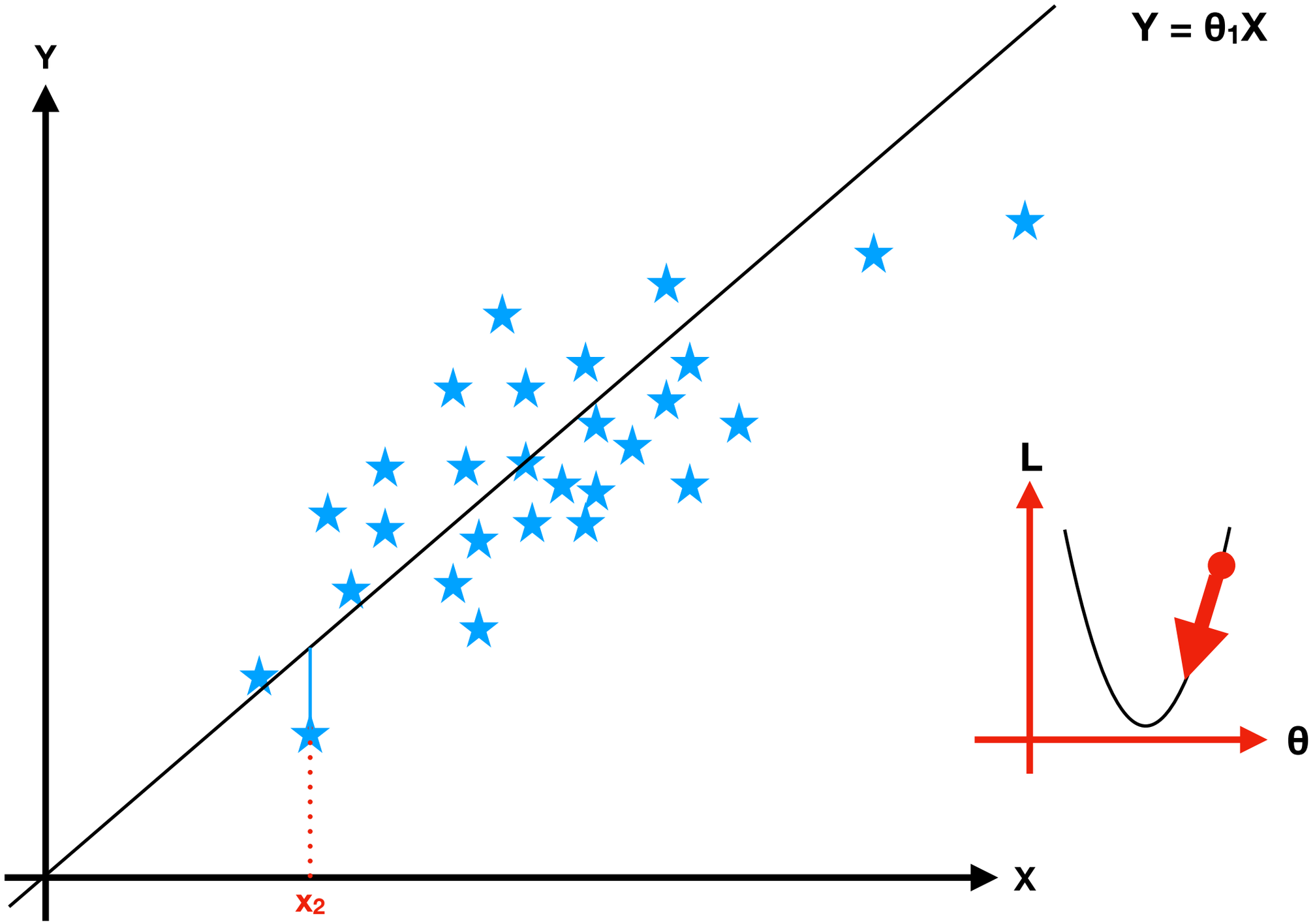
$$L = (y' - y_2)^2$$



$$L = (\theta_1 x_2 - y_2)^2$$

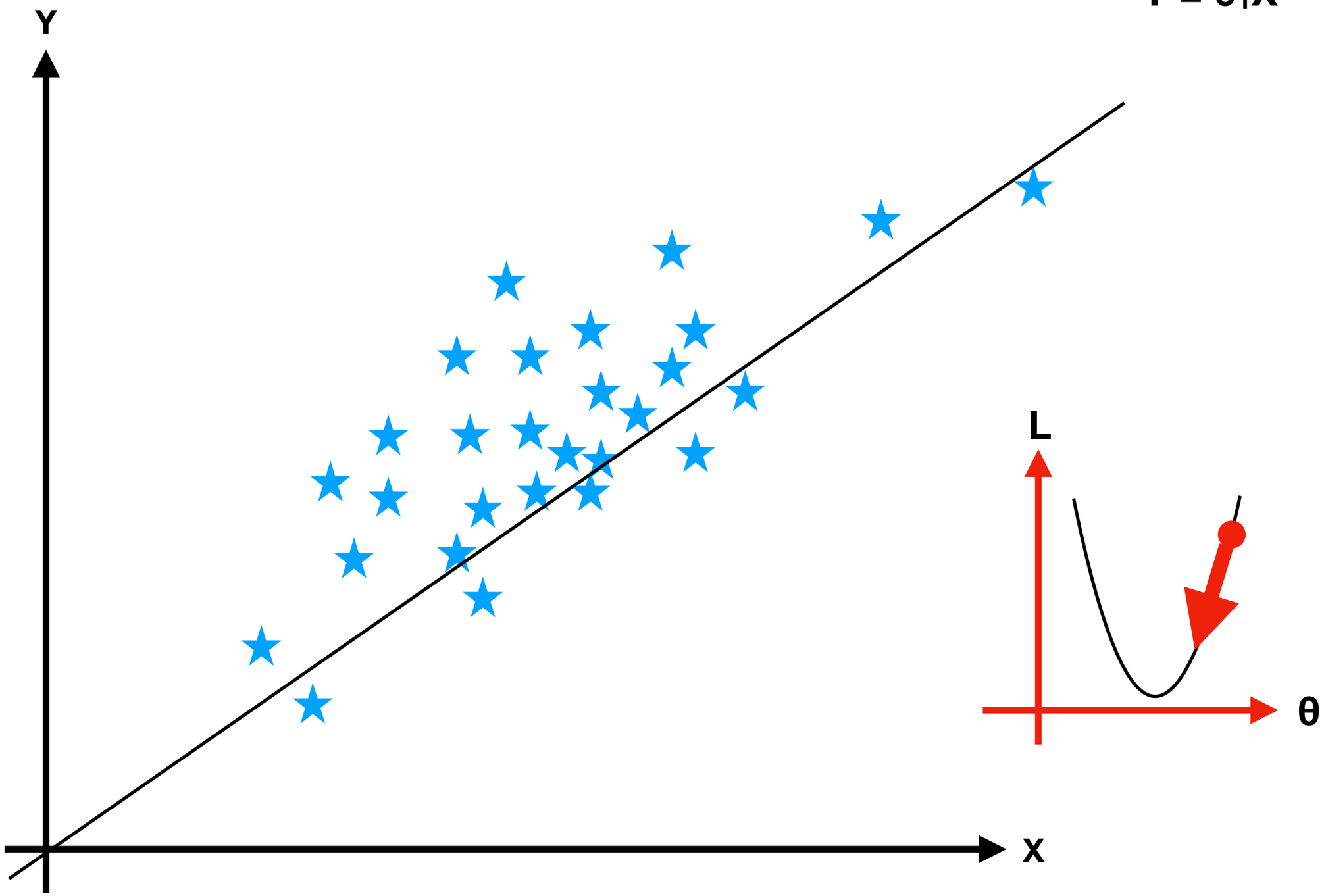


$$L = (\theta_1 x_2 - y_2)^2$$

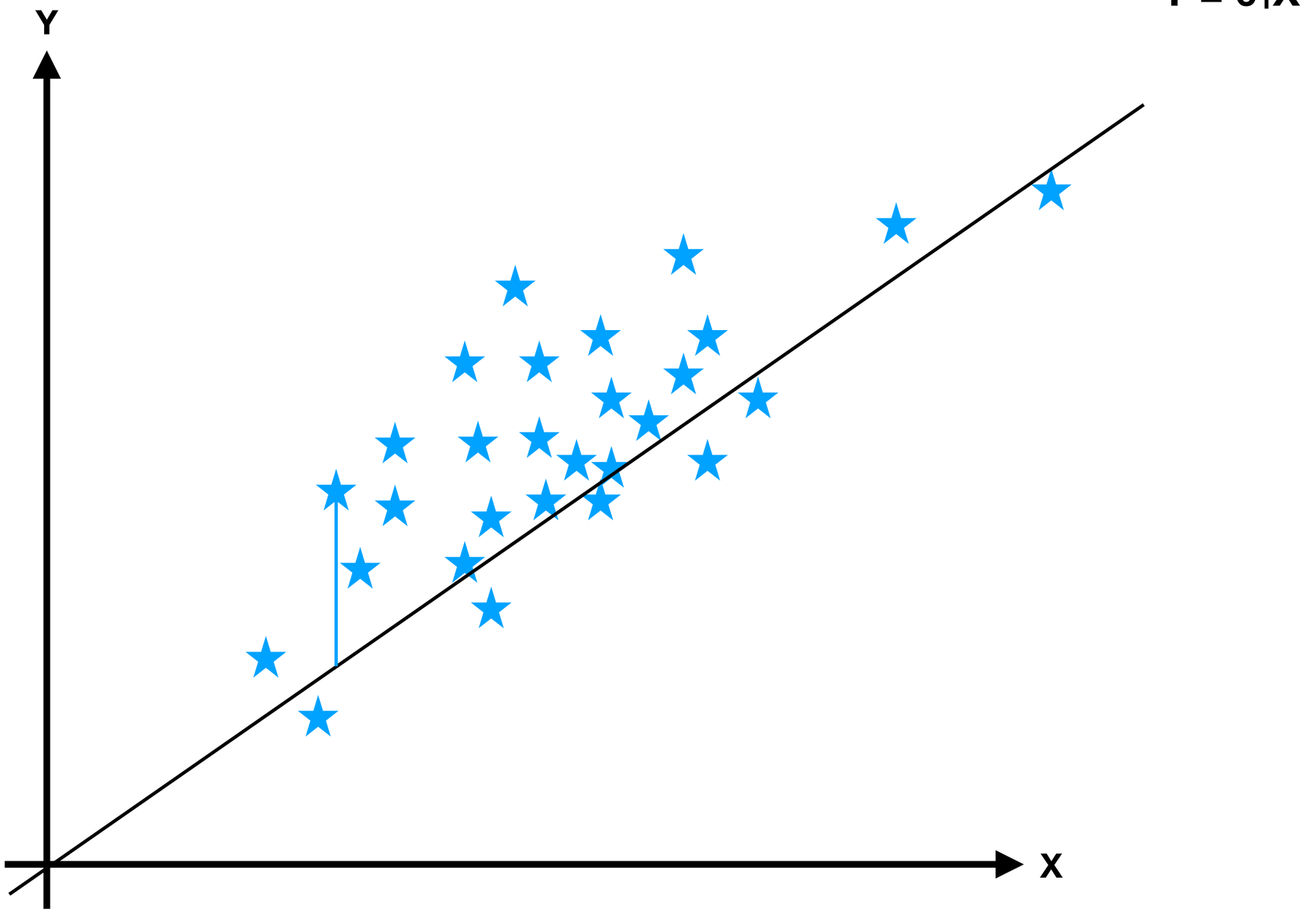


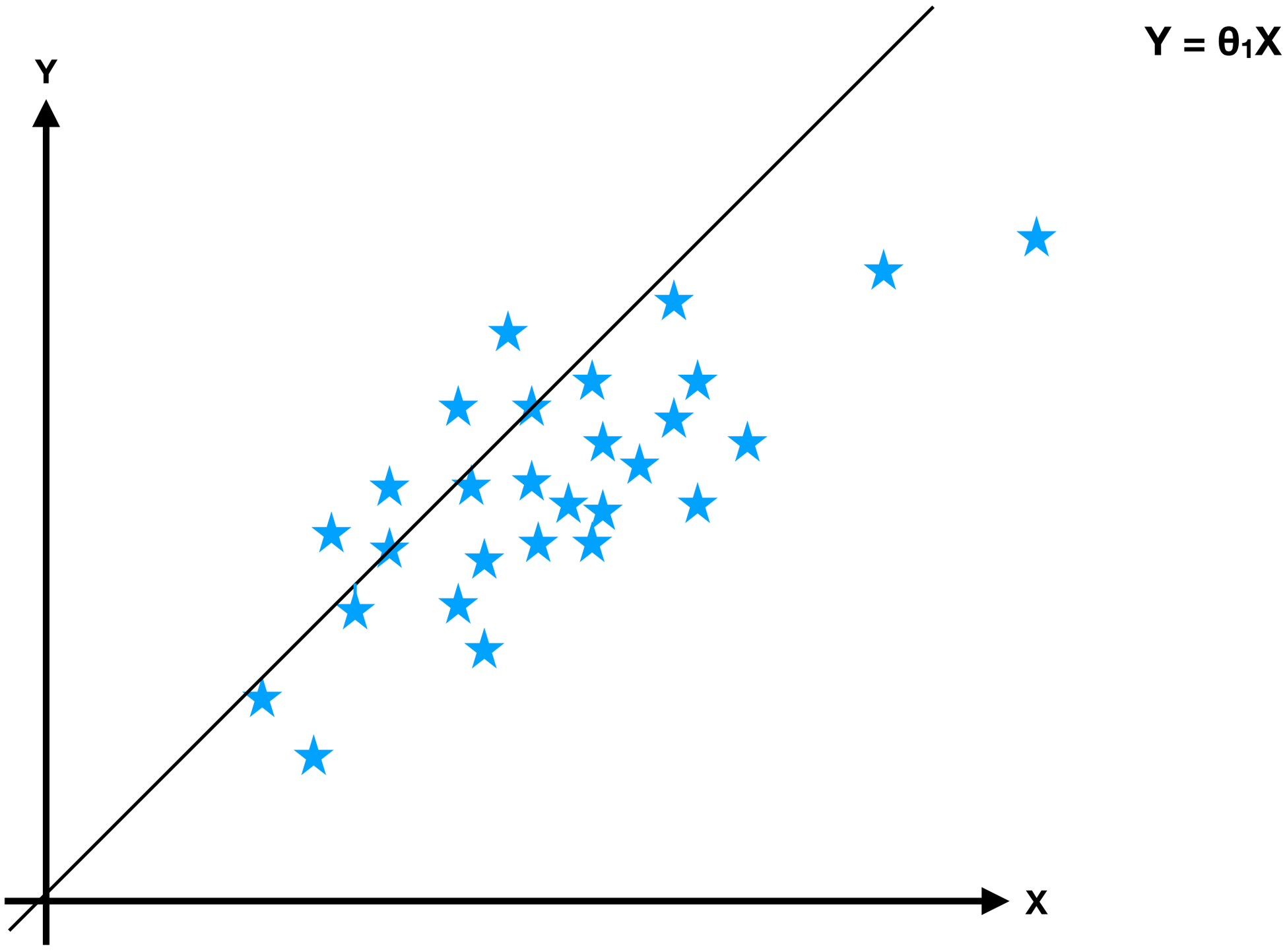
$$\theta = \theta - \alpha\,(2(\theta_1 x_2 - y_2)x_2)$$

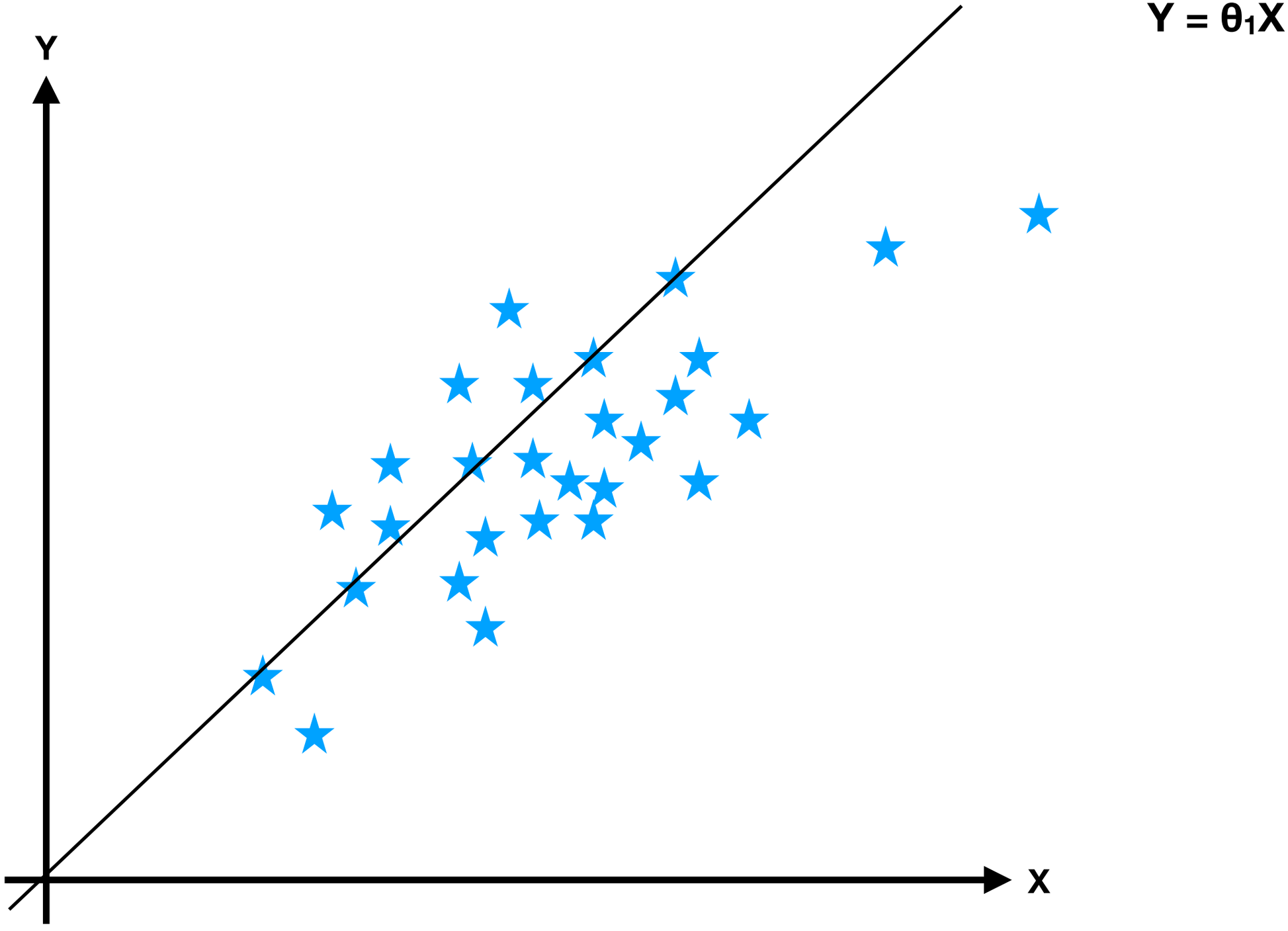
$$Y = \theta_1 X$$

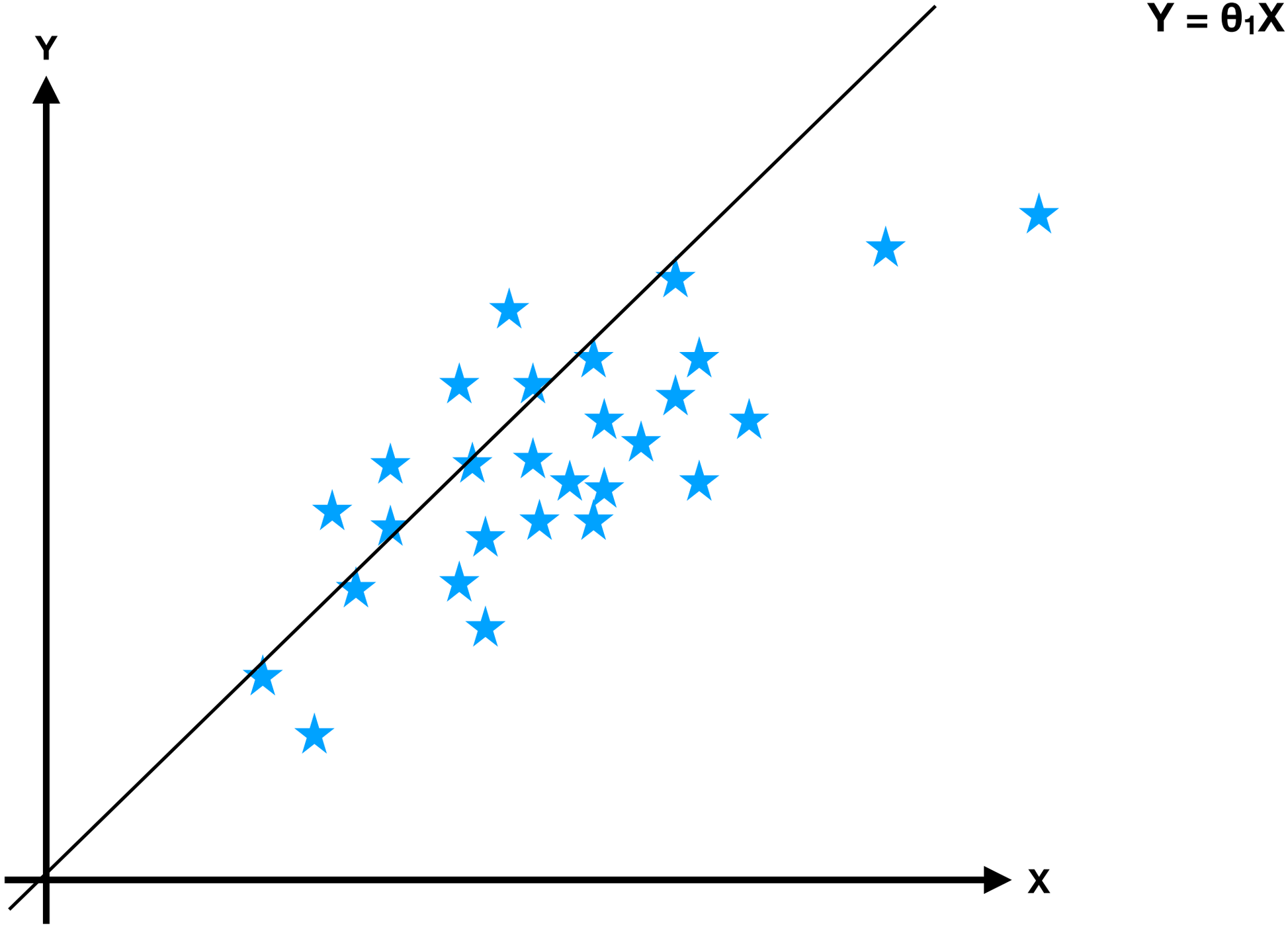


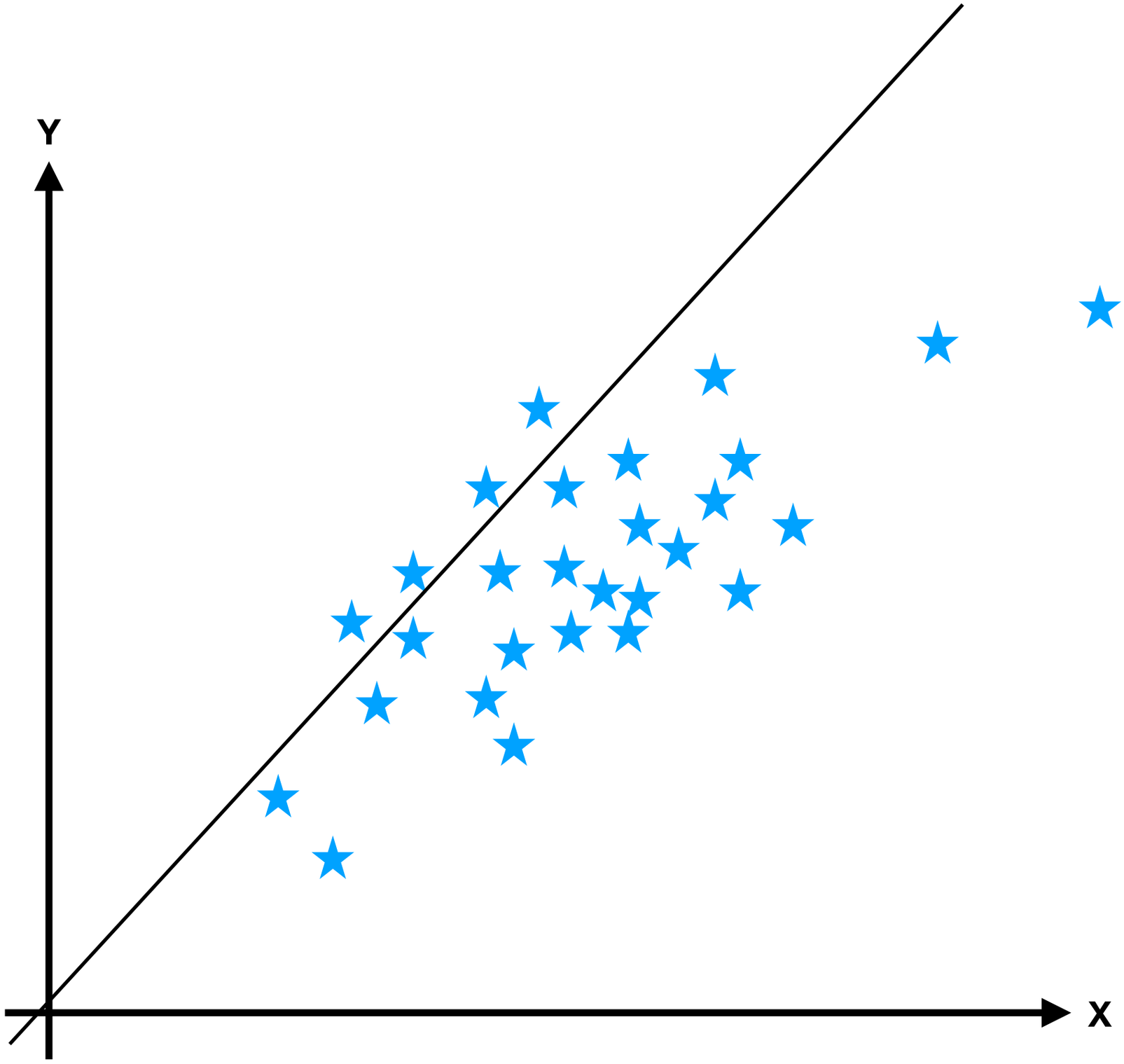
$$\theta = \theta - \alpha \left(2(\theta_1 x_2 - y_2)x_2\right)$$

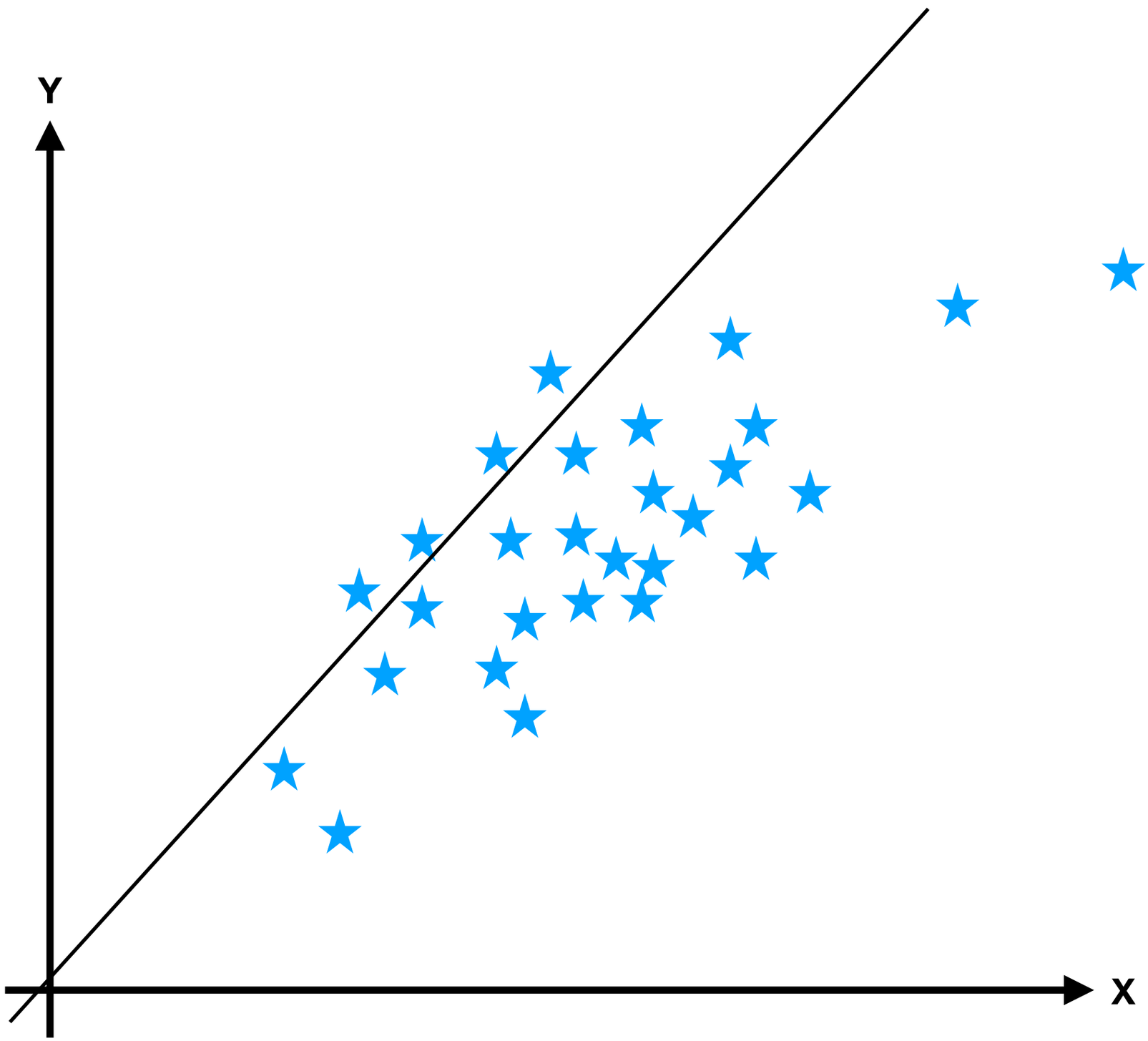

Y = θ1X
Y
X

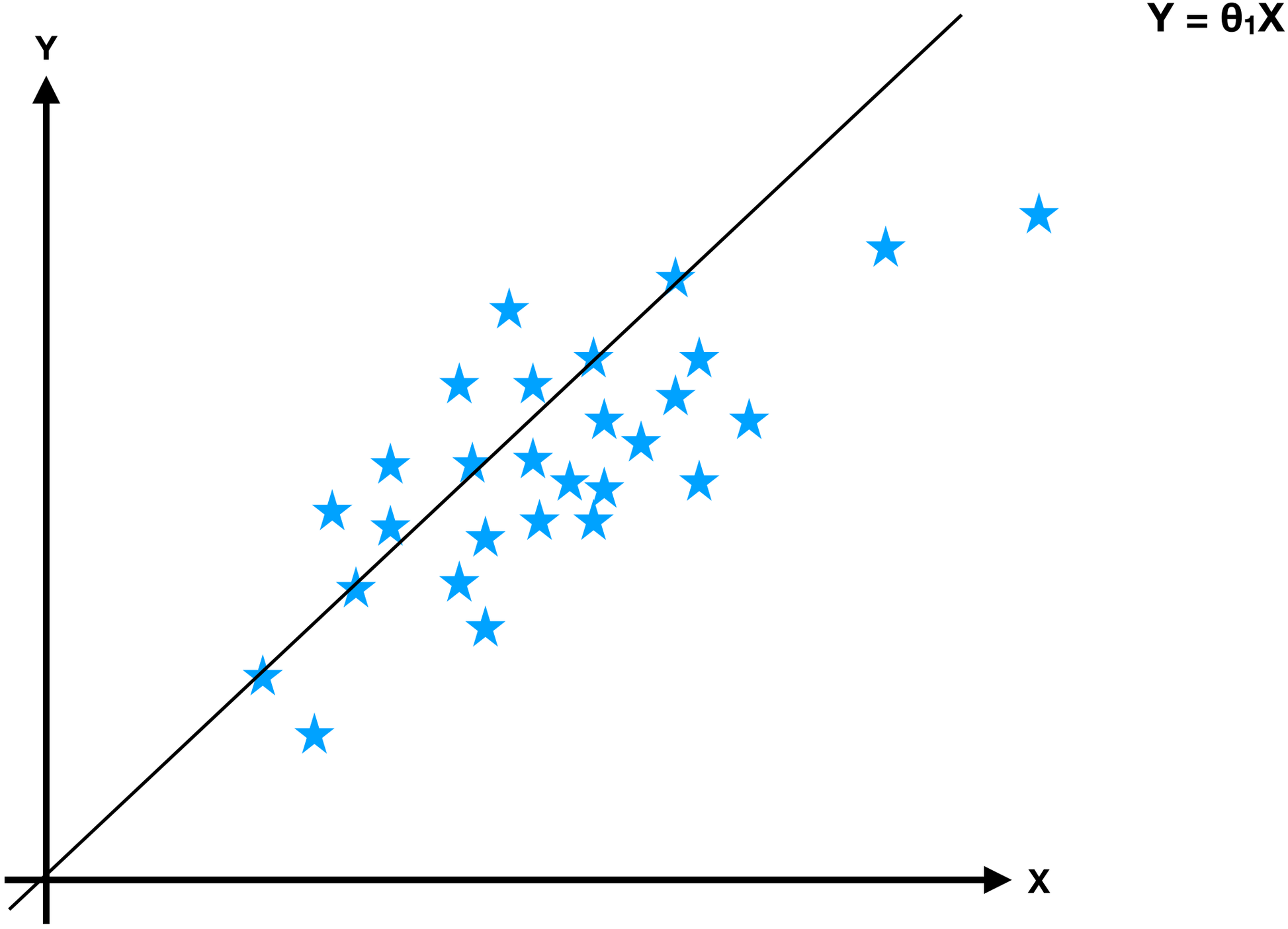

Y = θ1X
Y
X

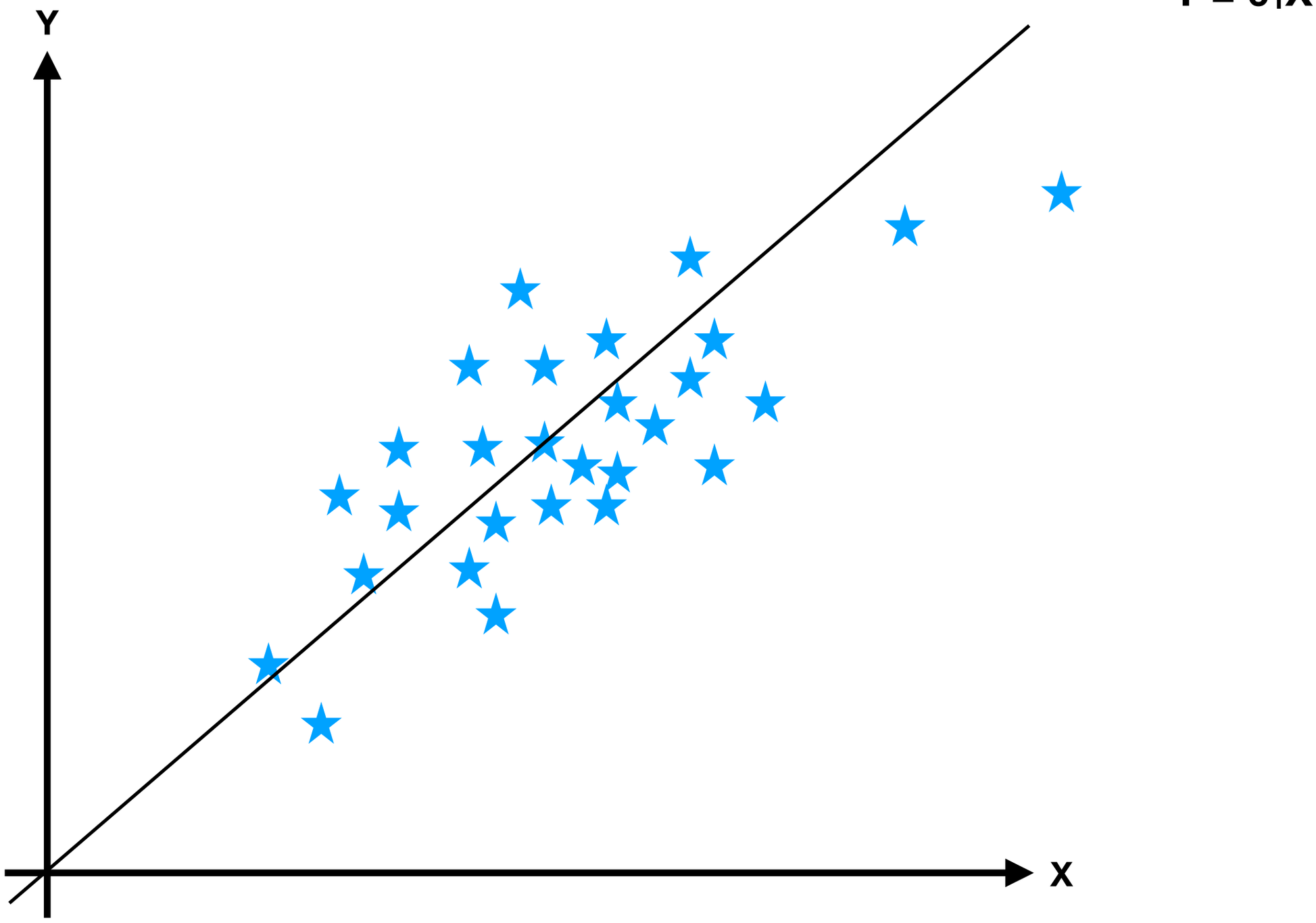

Y = θ1X
Y
X

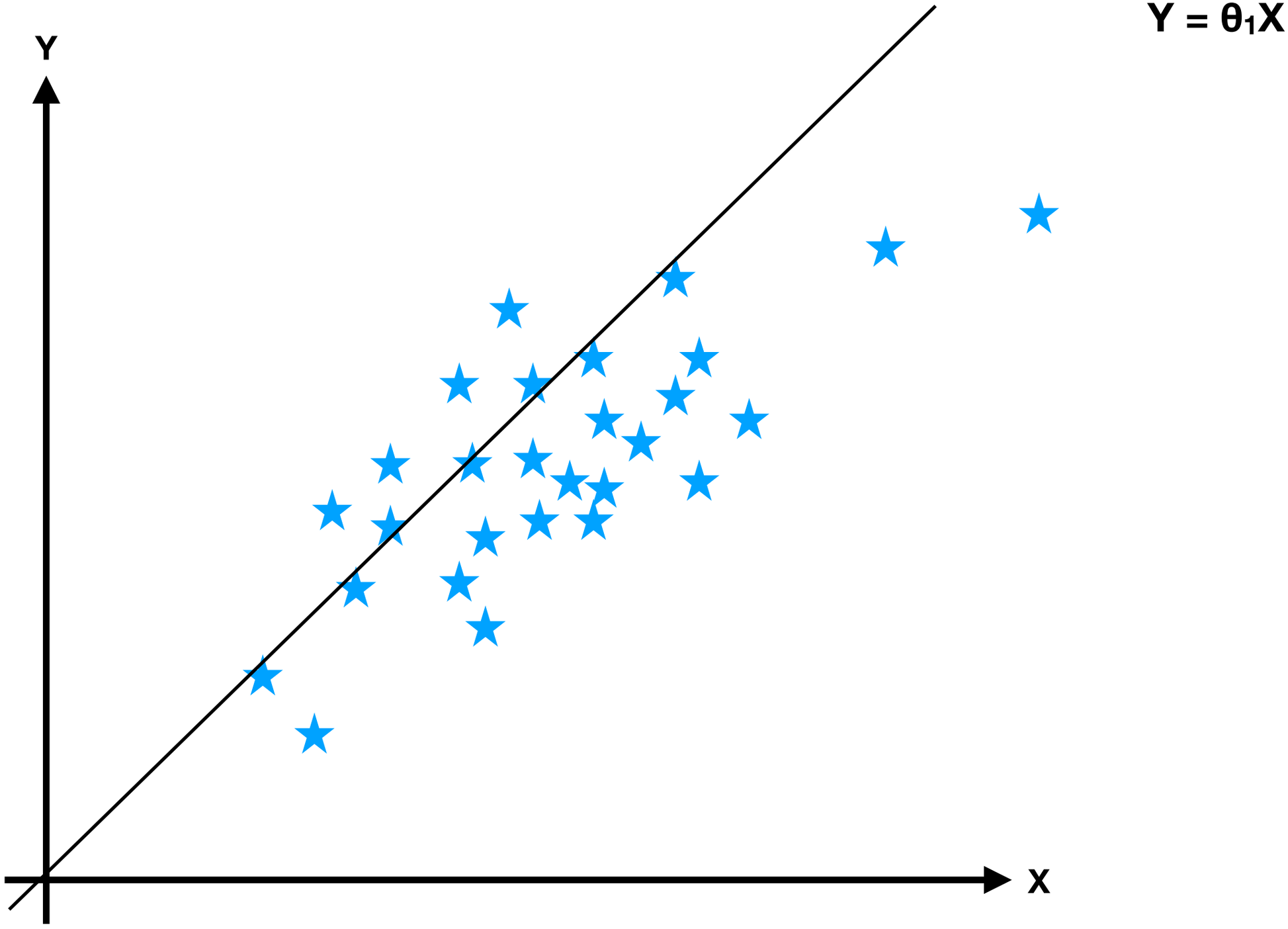

Y = θ1X
Y
X

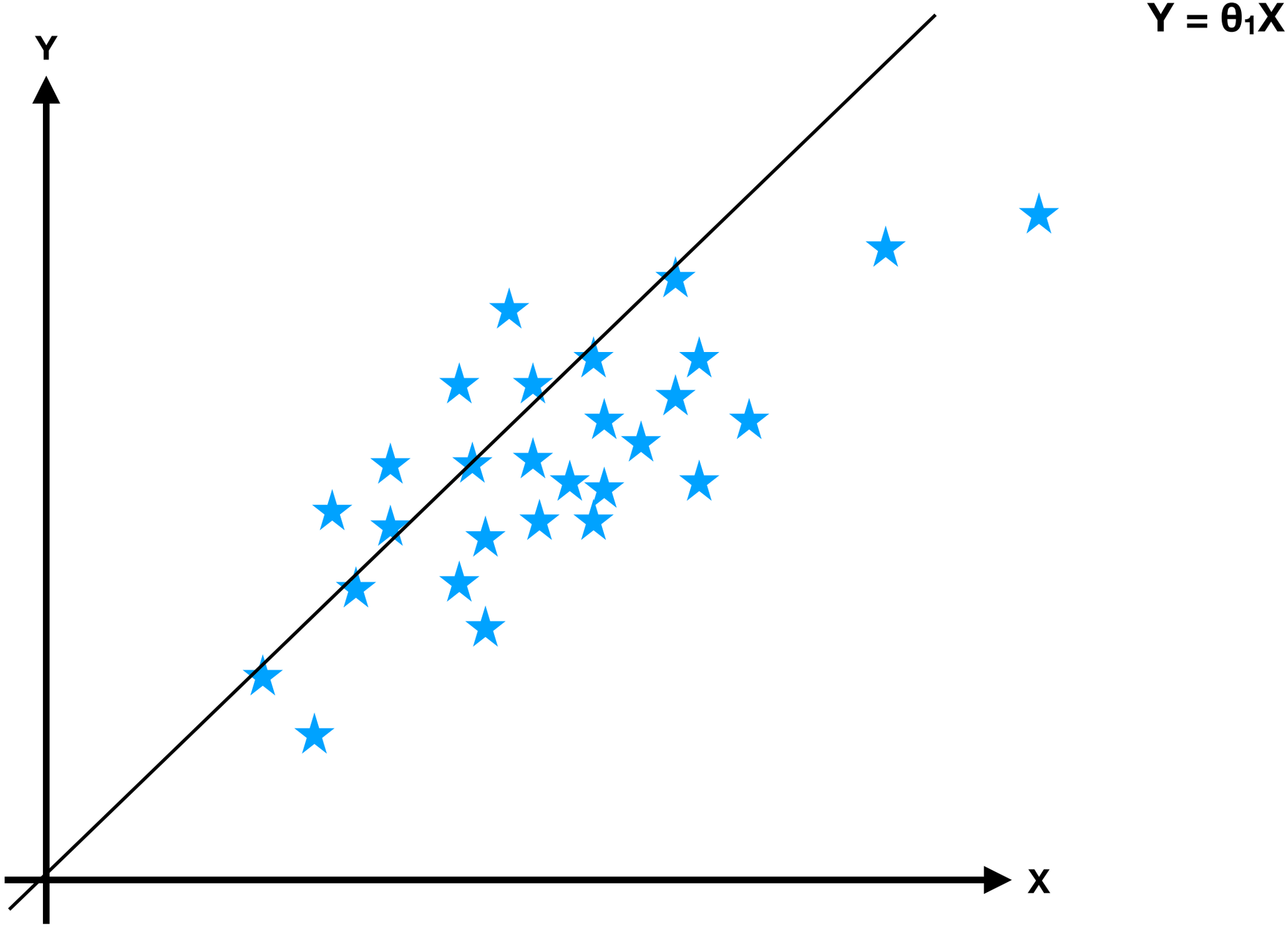

Y = θ1X
Y
X


Y = θ1X
Y
X

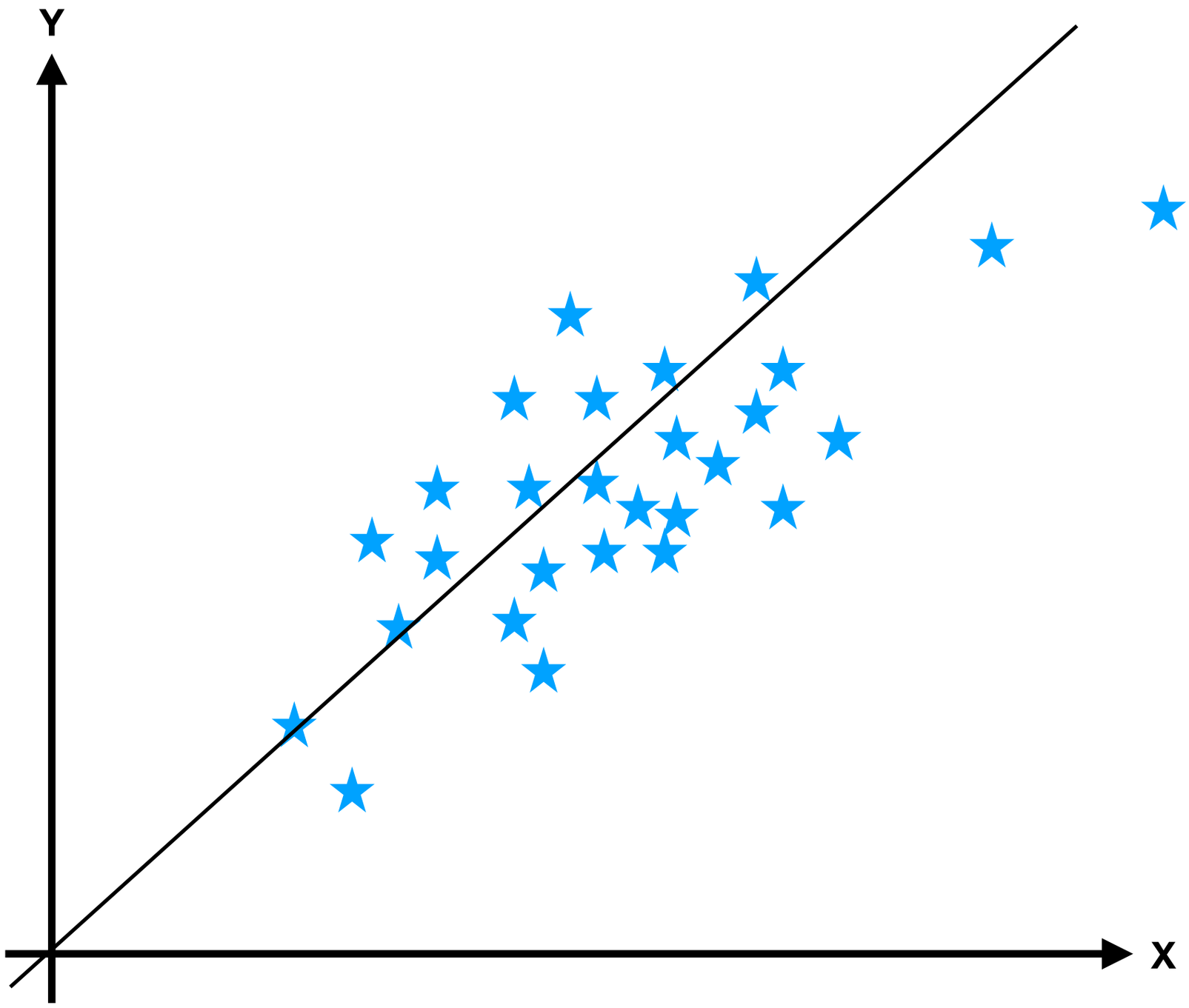

Y = θ1X
Y
X


Y = θ1X
Y
X

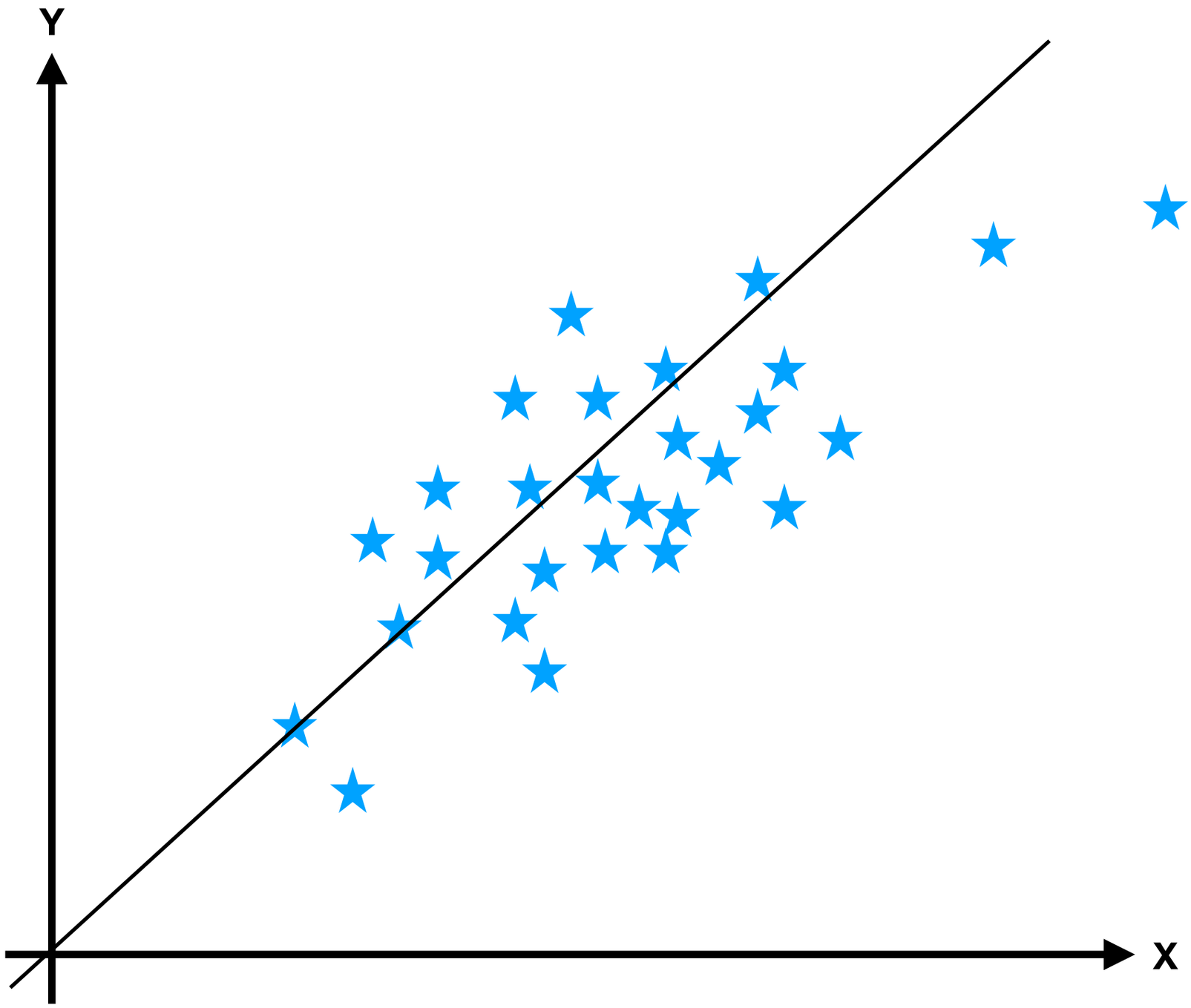

Y = θ1X
Y
X


Y = θ1X
Y
X

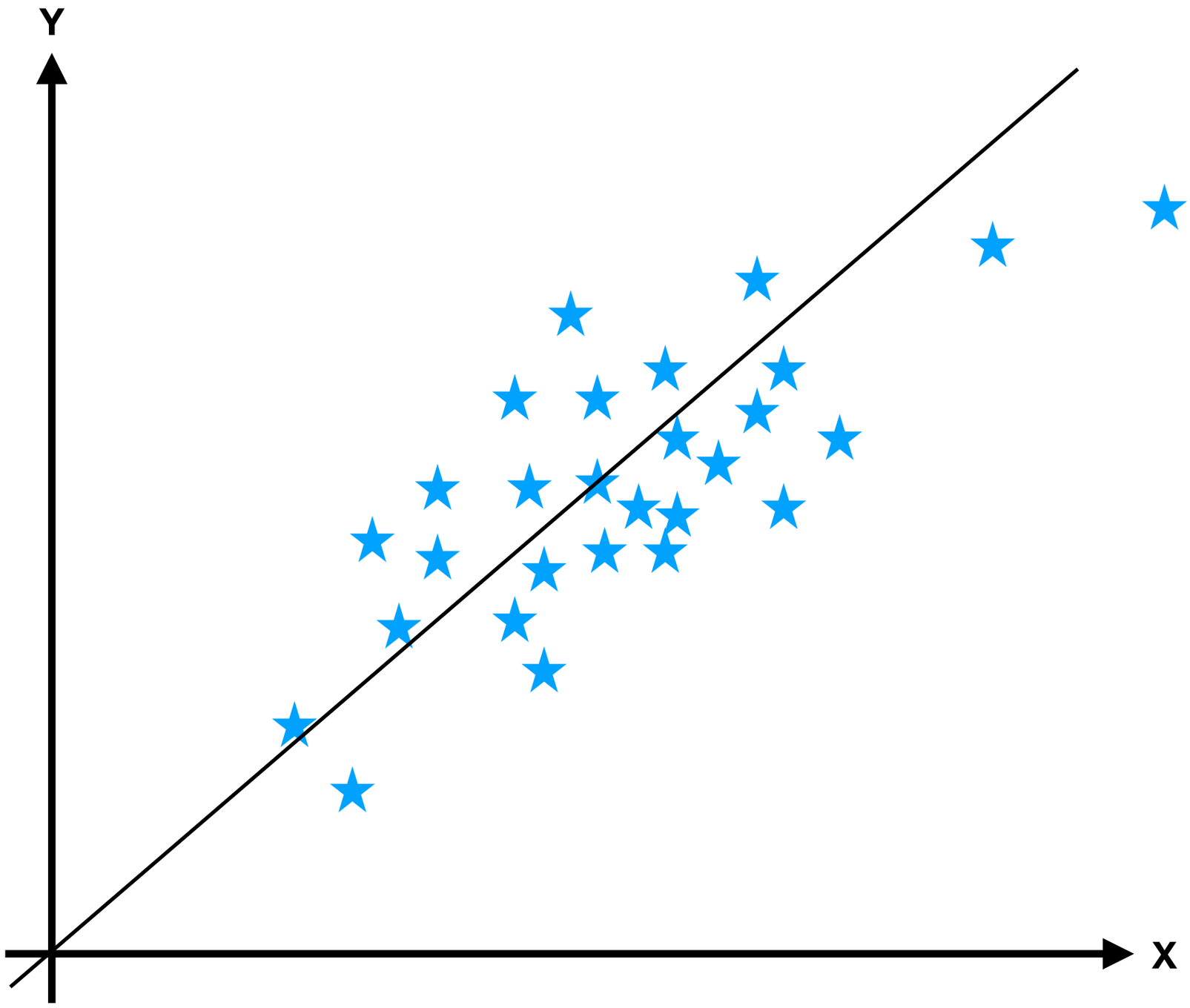
$Y = \theta_1 X$



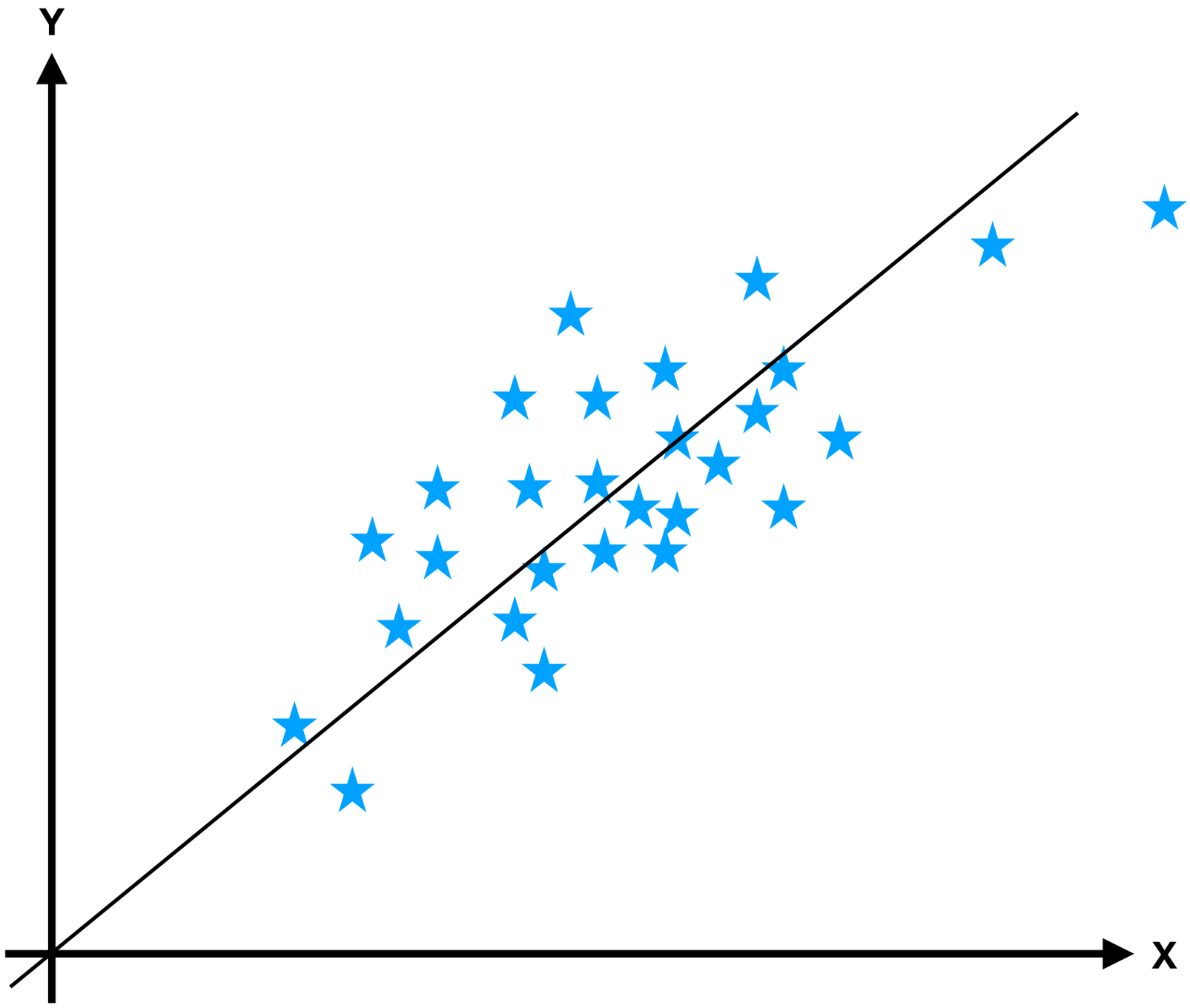
$Y = \theta_1 X$



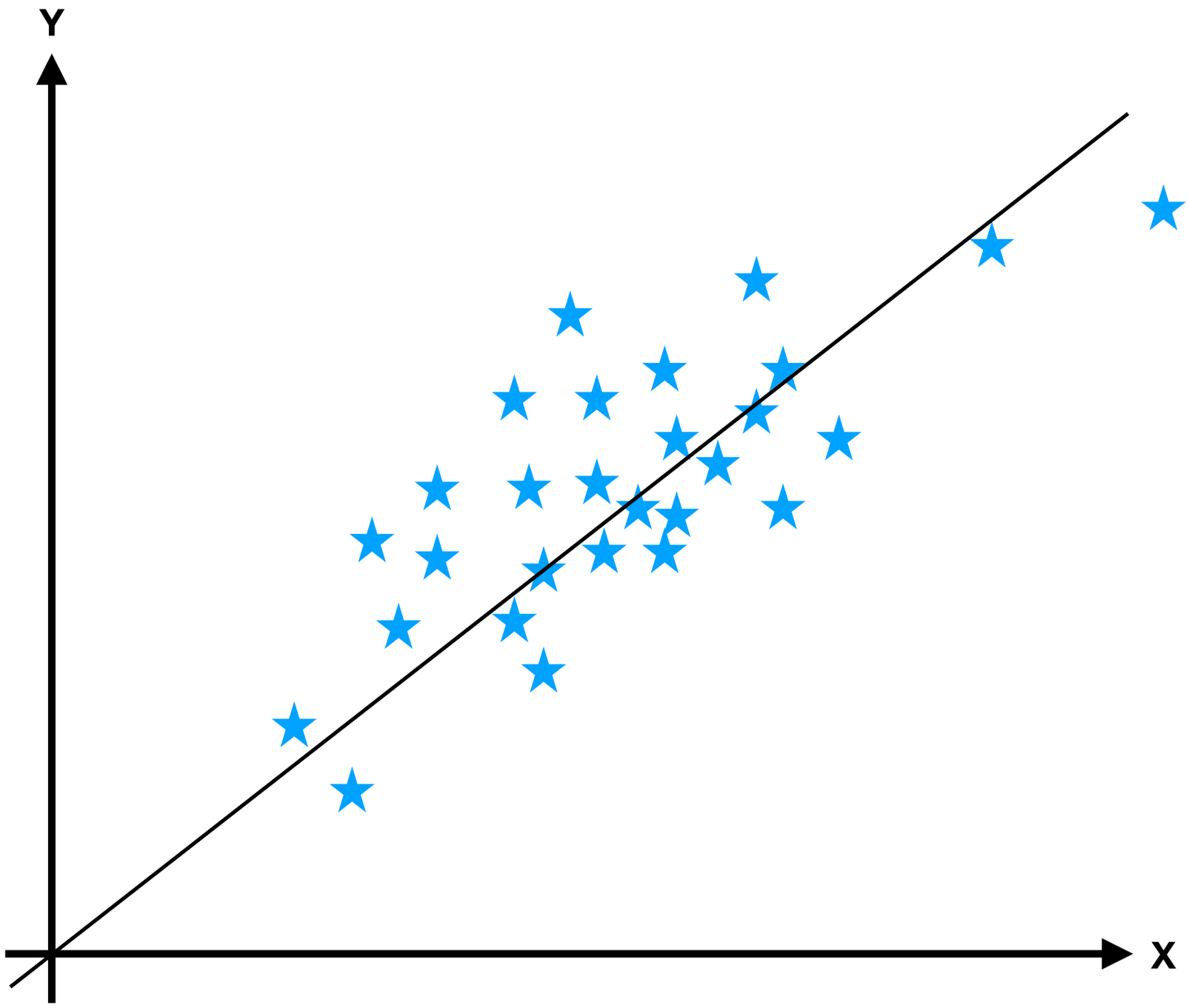
$Y = \theta_1 X$



$Y = \theta_1 X$

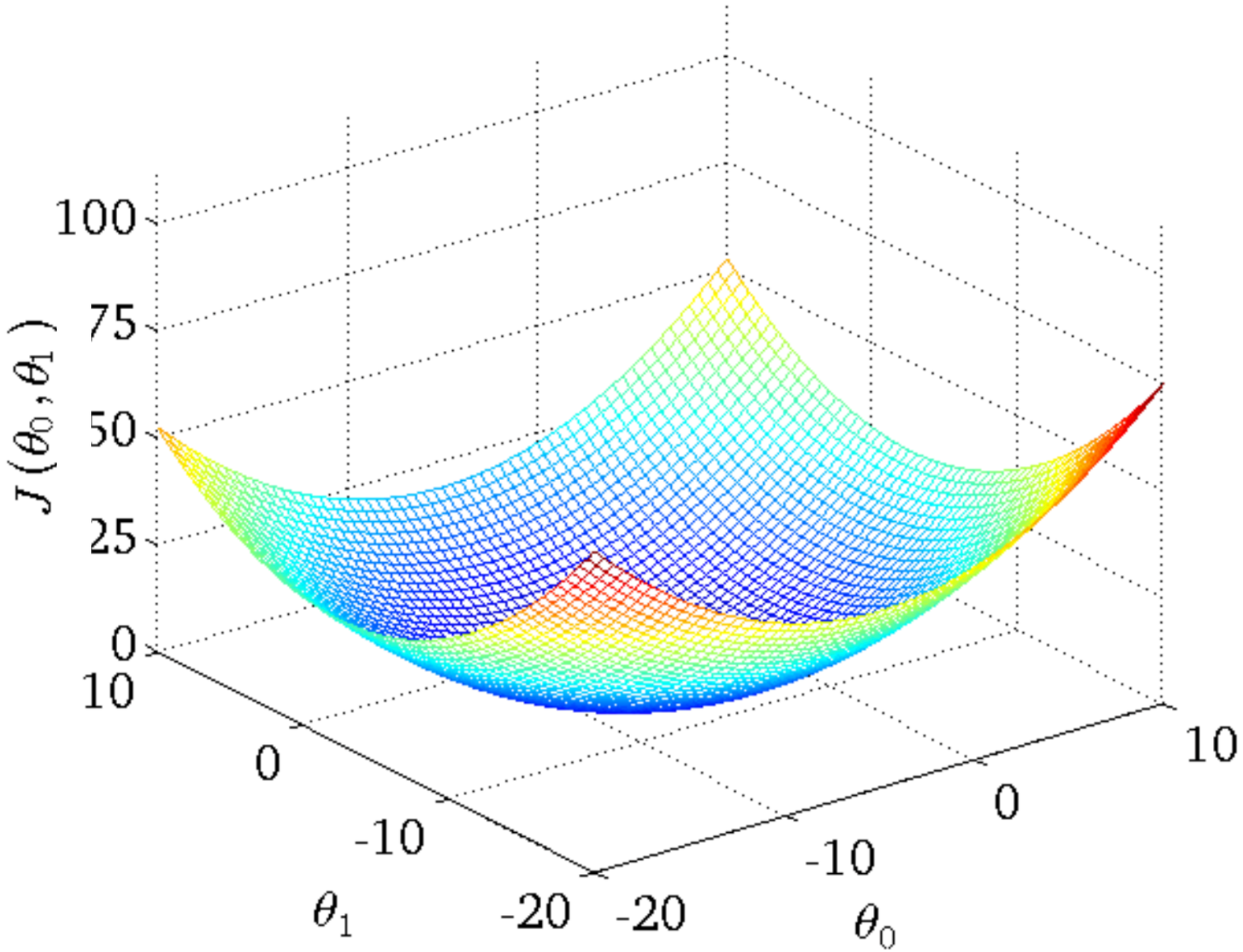


$Y = \theta_1 X$



Stochastic Gradient Descent

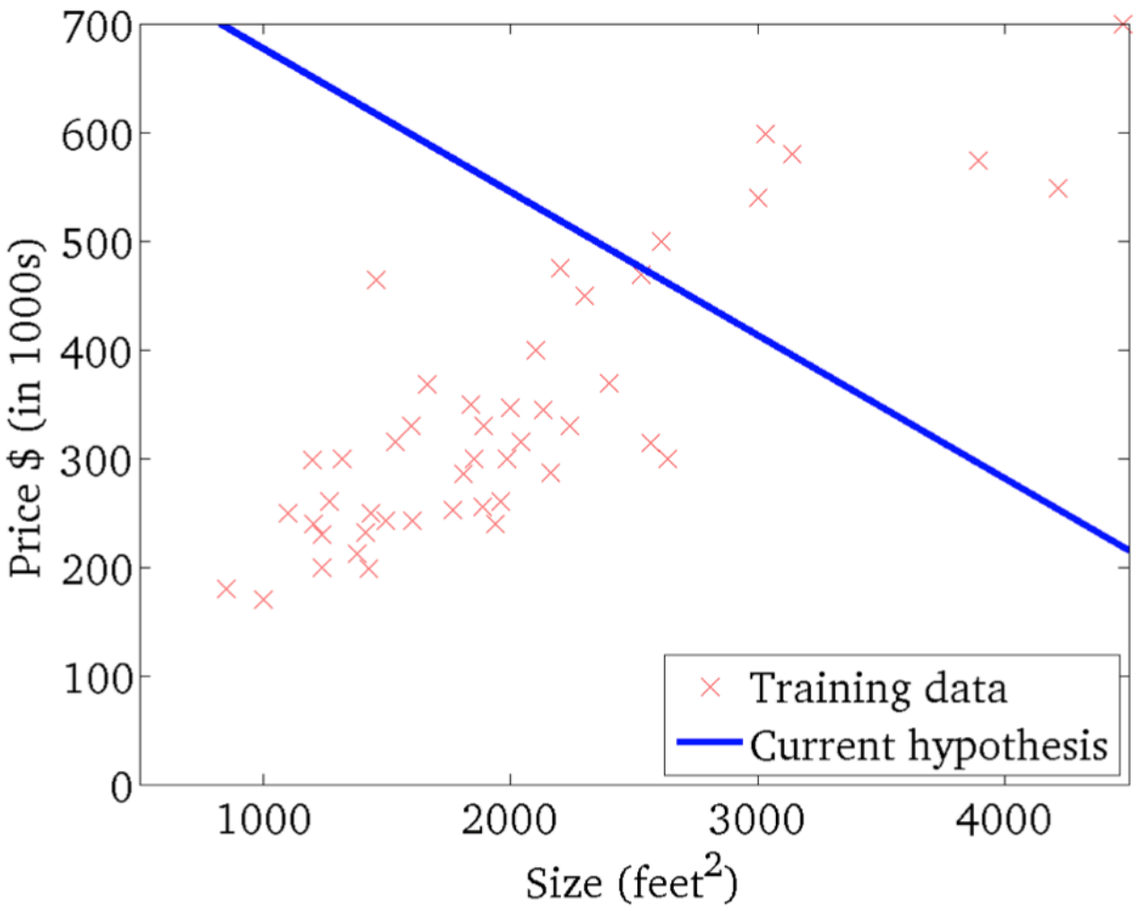
Contour Plot



Courtesy: Andrew Ng

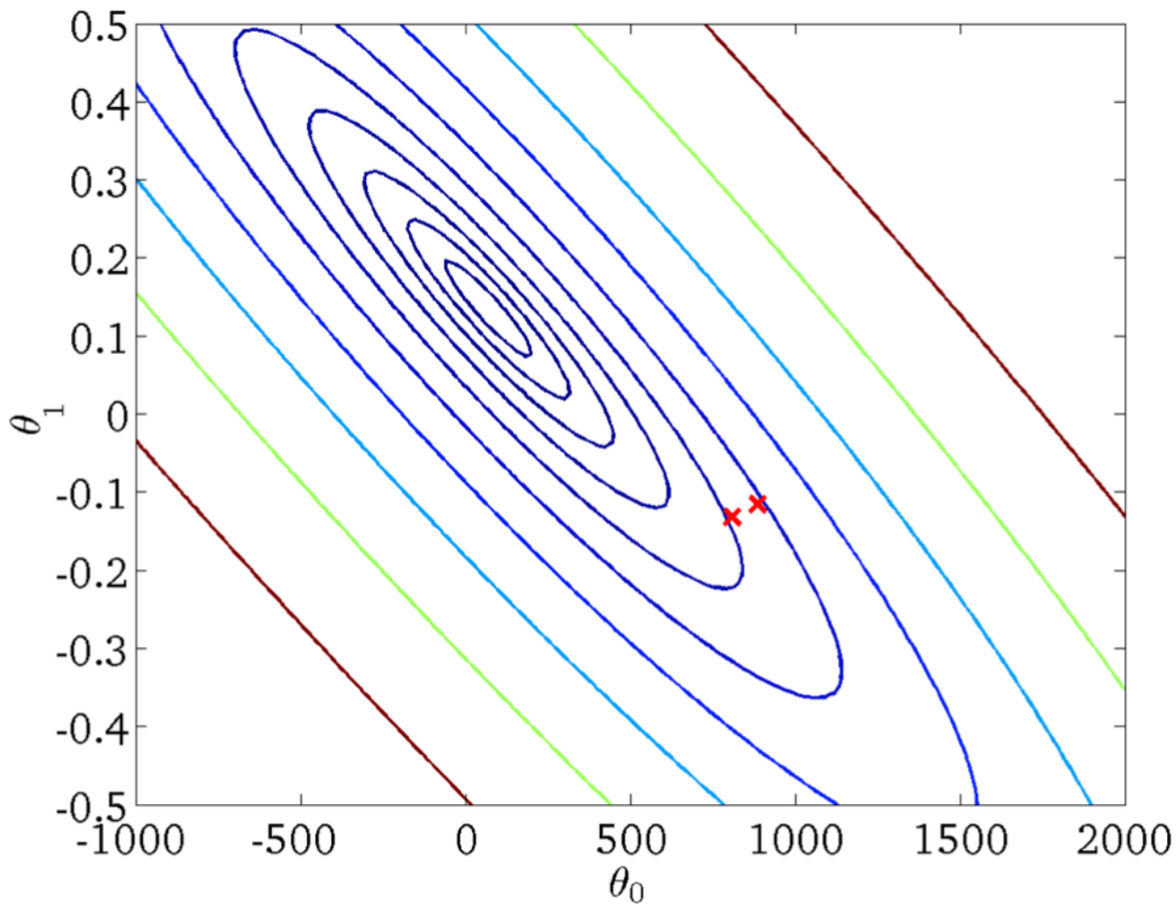
$h_\theta(x)$

(for fixed θ_0, θ_1, this is a function of x)



$J(\theta_0, \theta_1)$

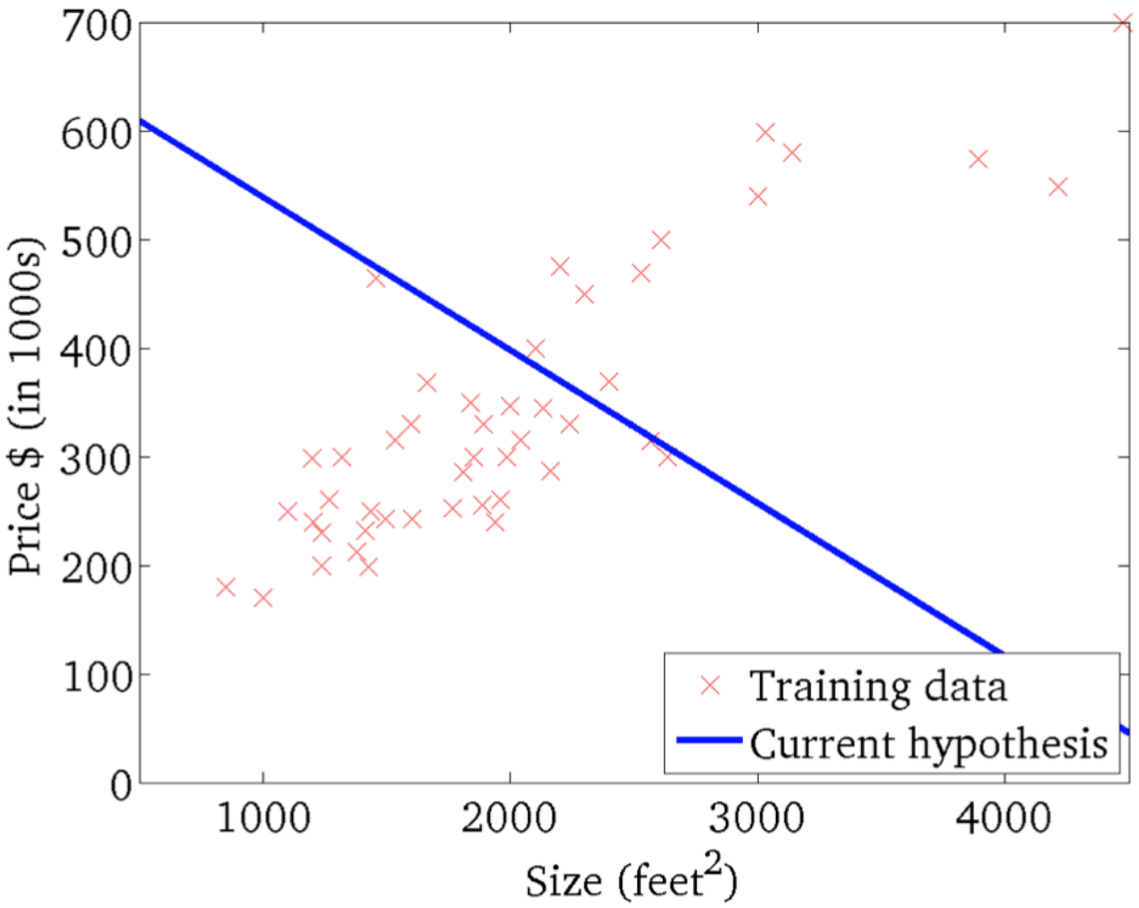
(function of the parameters θ_0, θ_1)



Courtesy: Andrew Ng

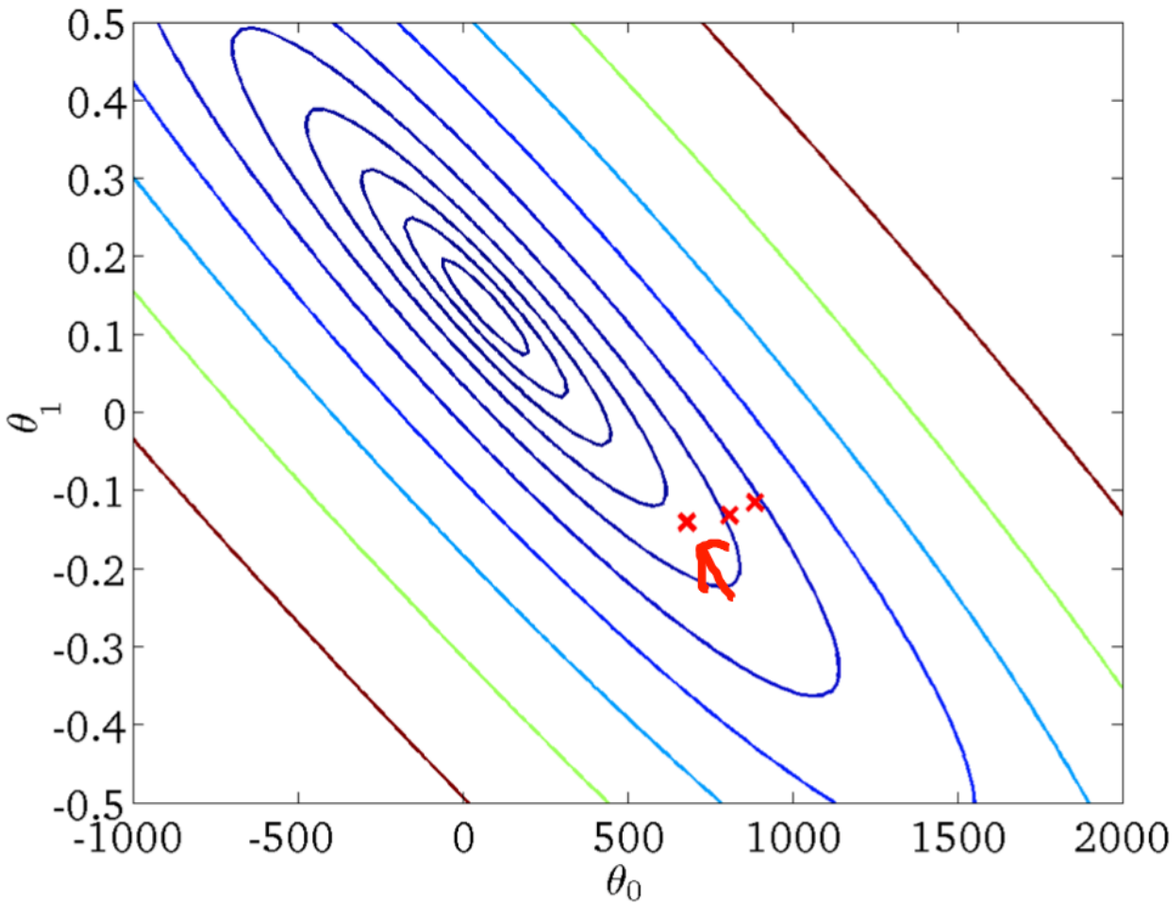
$h_\theta(x)$

(for fixed θ_0, θ_1, this is a function of x)



$J(\theta_0, \theta_1)$

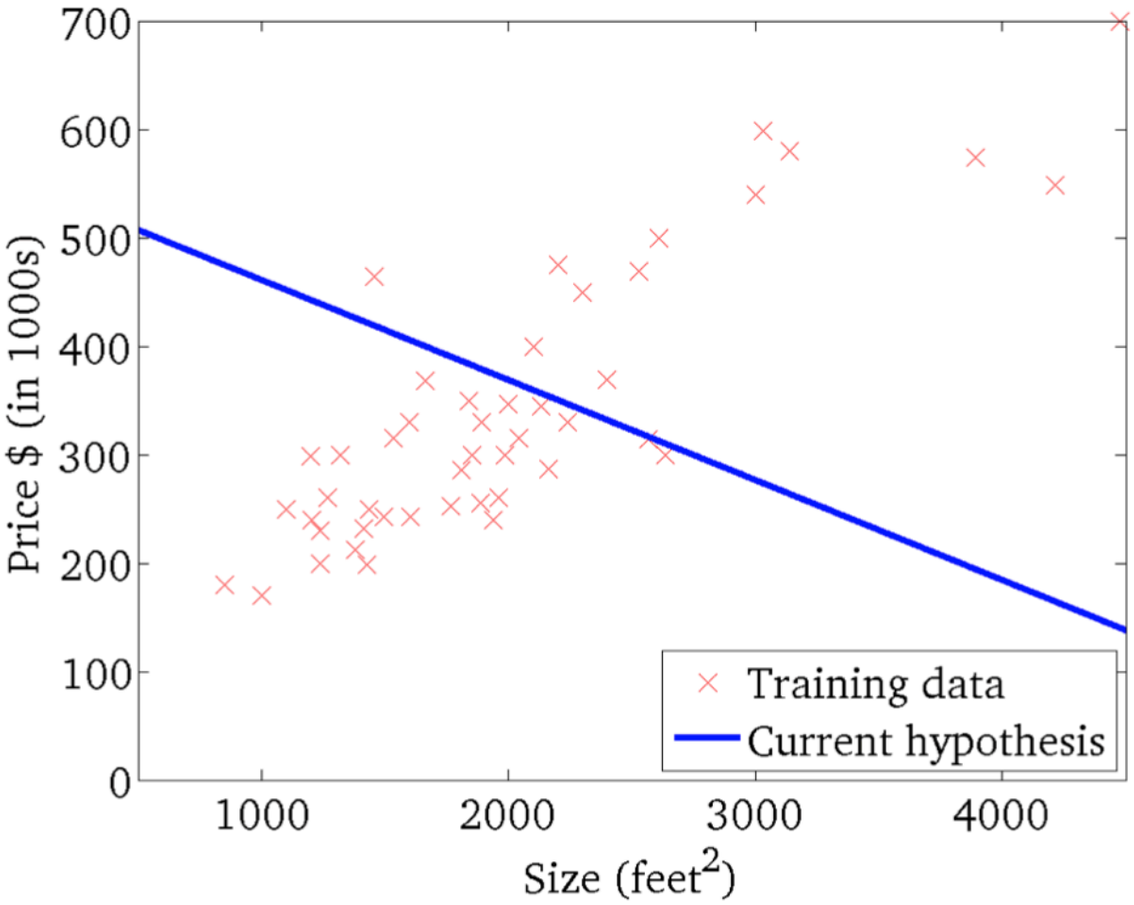
(function of the parameters θ_0, θ_1)



Courtesy: Andrew Ng

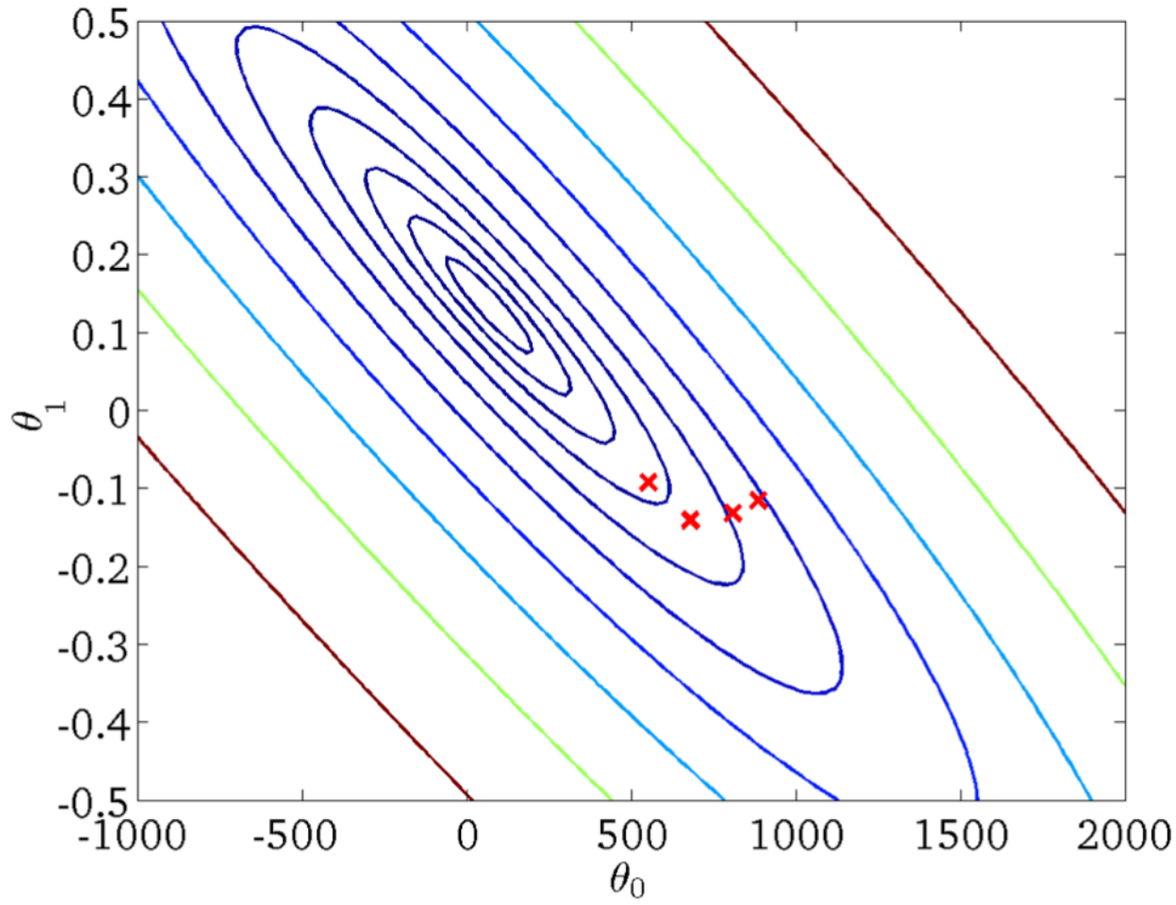
$h_\theta(x)$

(for fixed θ_0, θ_1, this is a function of x)



$J(\theta_0, \theta_1)$

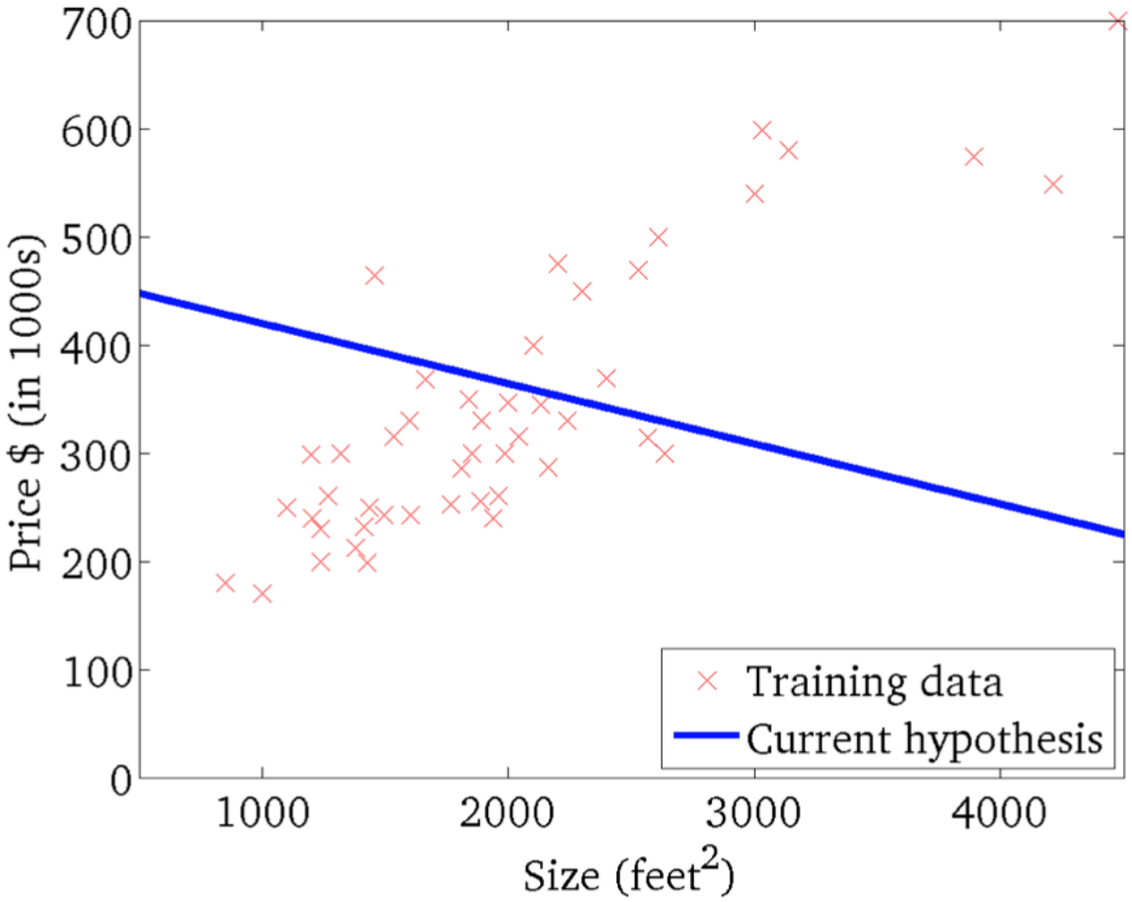
(function of the parameters θ_0, θ_1)



Courtesy: Andrew Ng

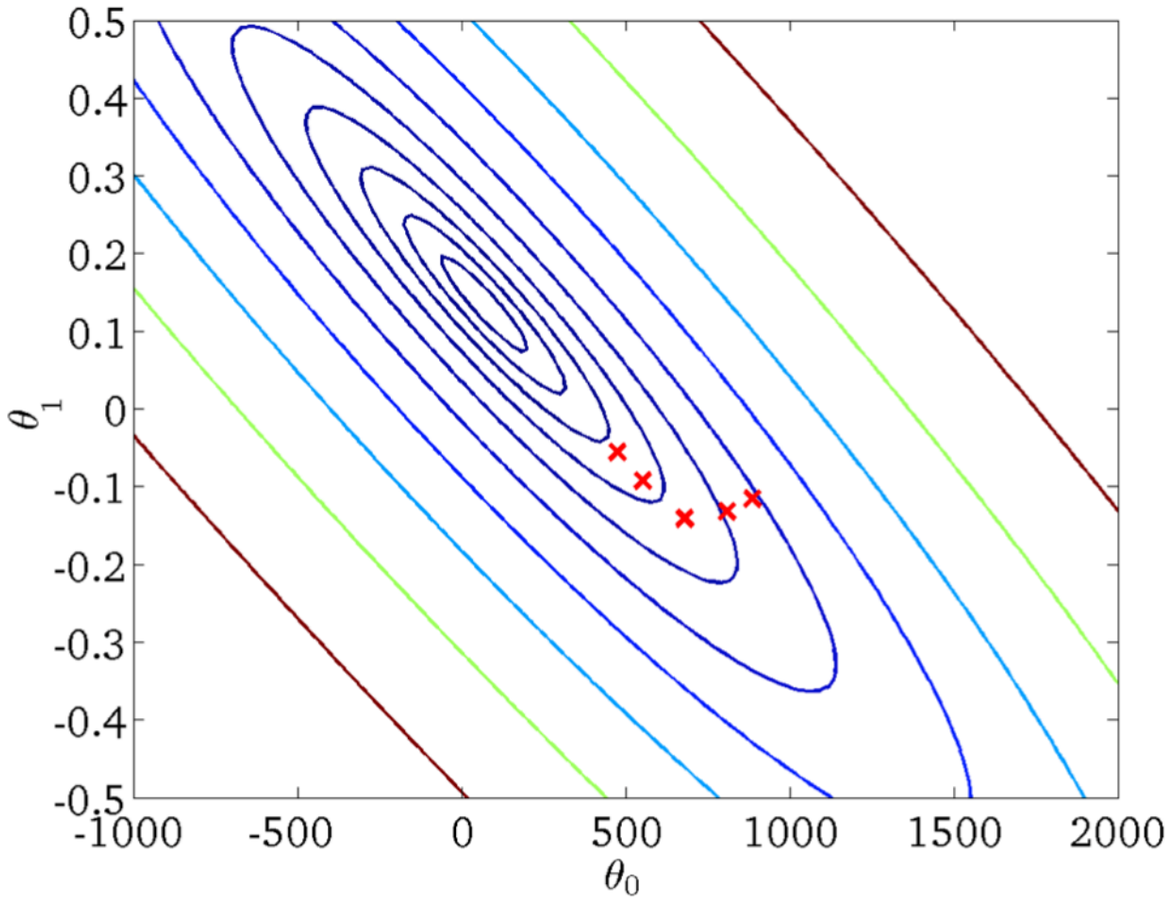
$h_\theta(x)$

(for fixed θ_0, θ_1, this is a function of x)



$J(\theta_0, \theta_1)$

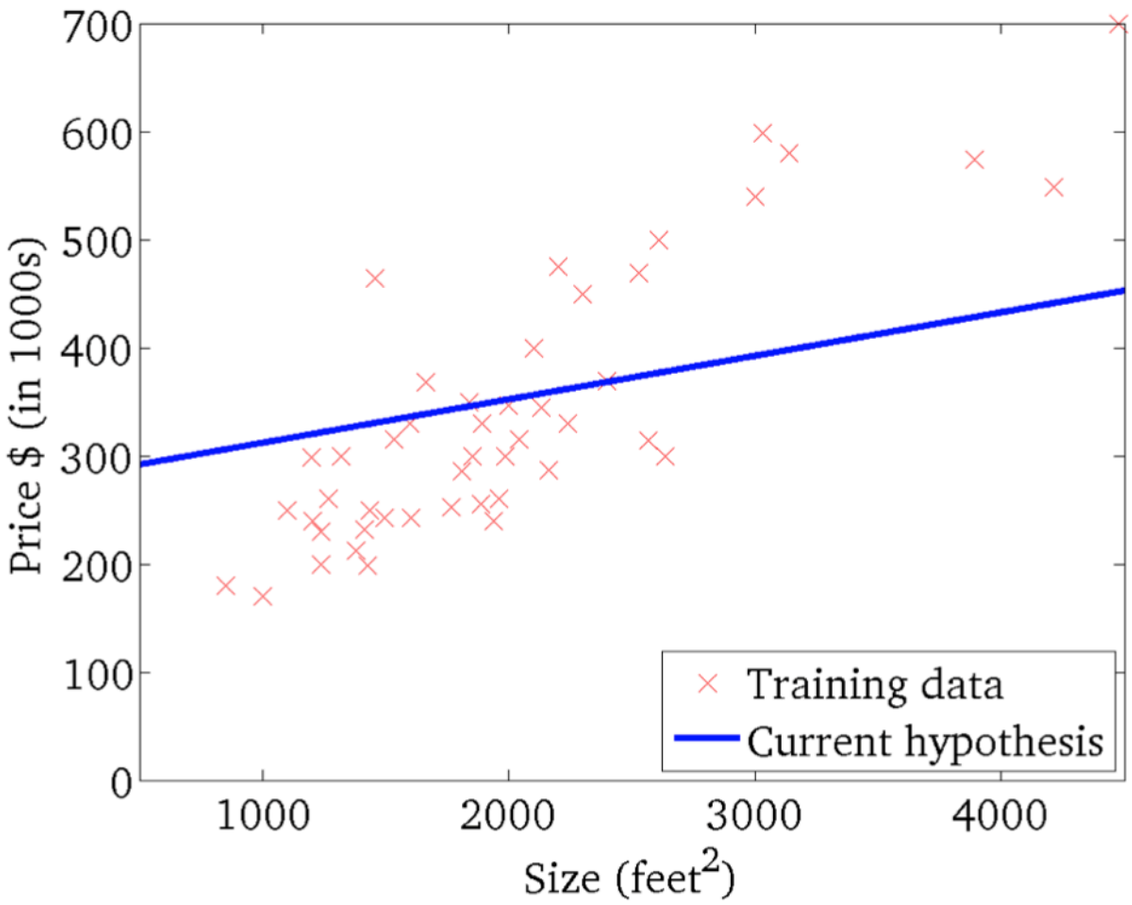
(function of the parameters θ_0, θ_1)



Courtesy: Andrew Ng

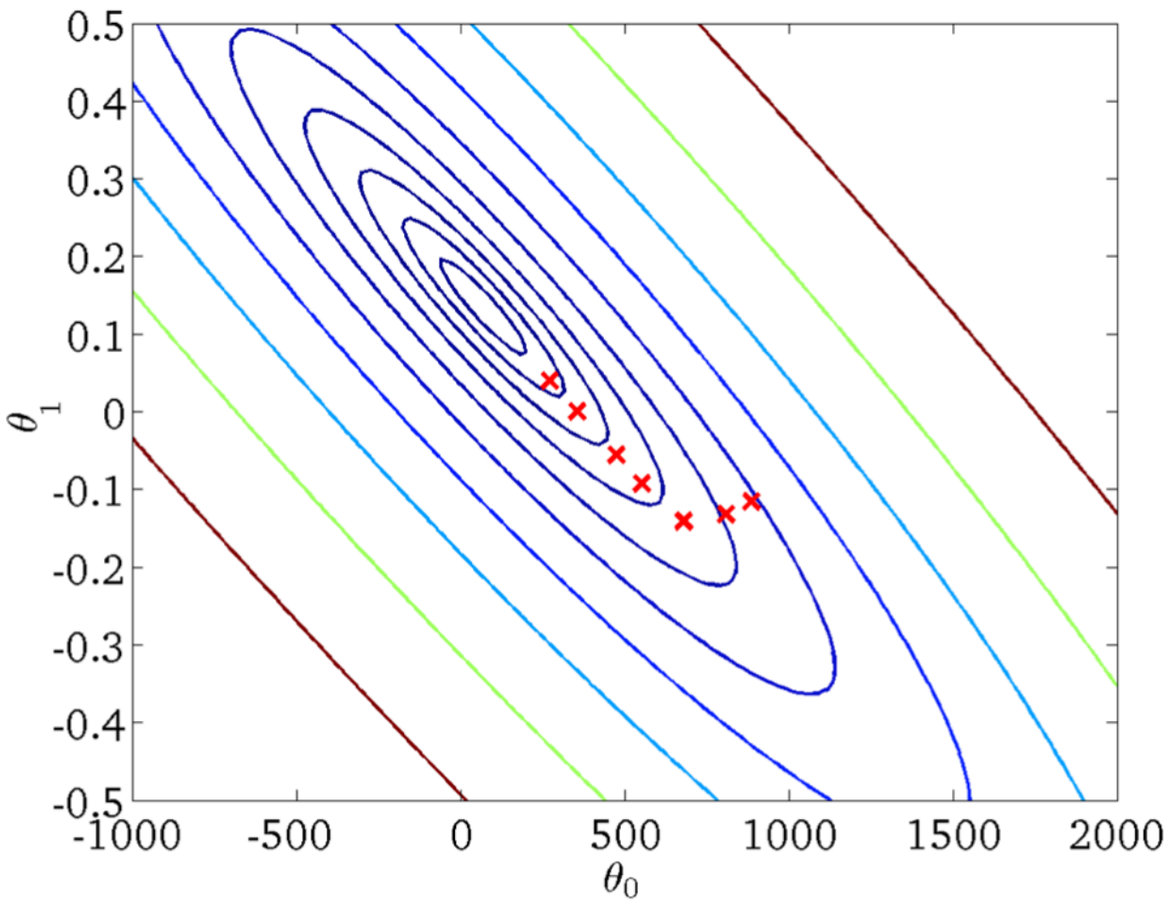
$h_\theta(x)$

(for fixed θ_0, θ_1, this is a function of x)



$J(\theta_0, \theta_1)$

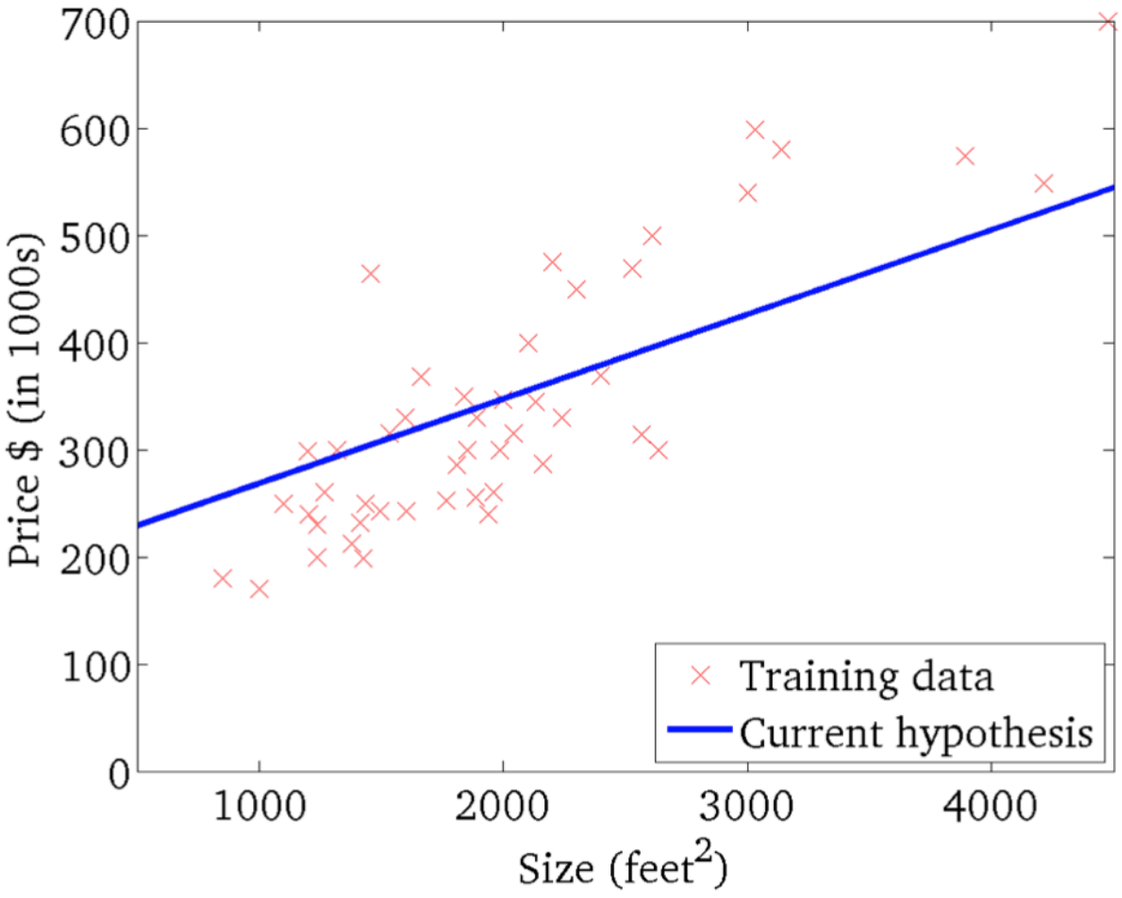
(function of the parameters θ_0, θ_1)



Courtesy: Andrew Ng

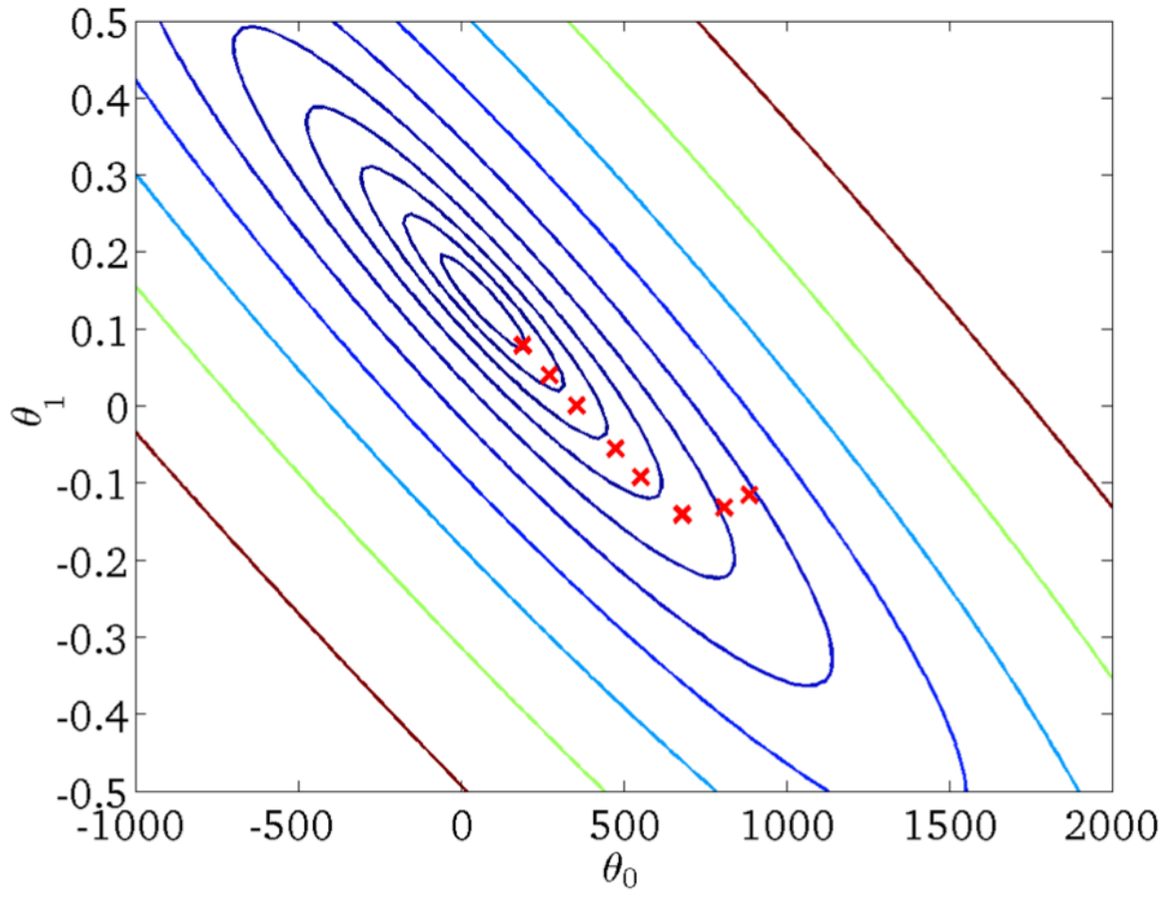
$h_\theta(x)$

(for fixed θ_0, θ_1, this is a function of x)



$J(\theta_0, \theta_1)$

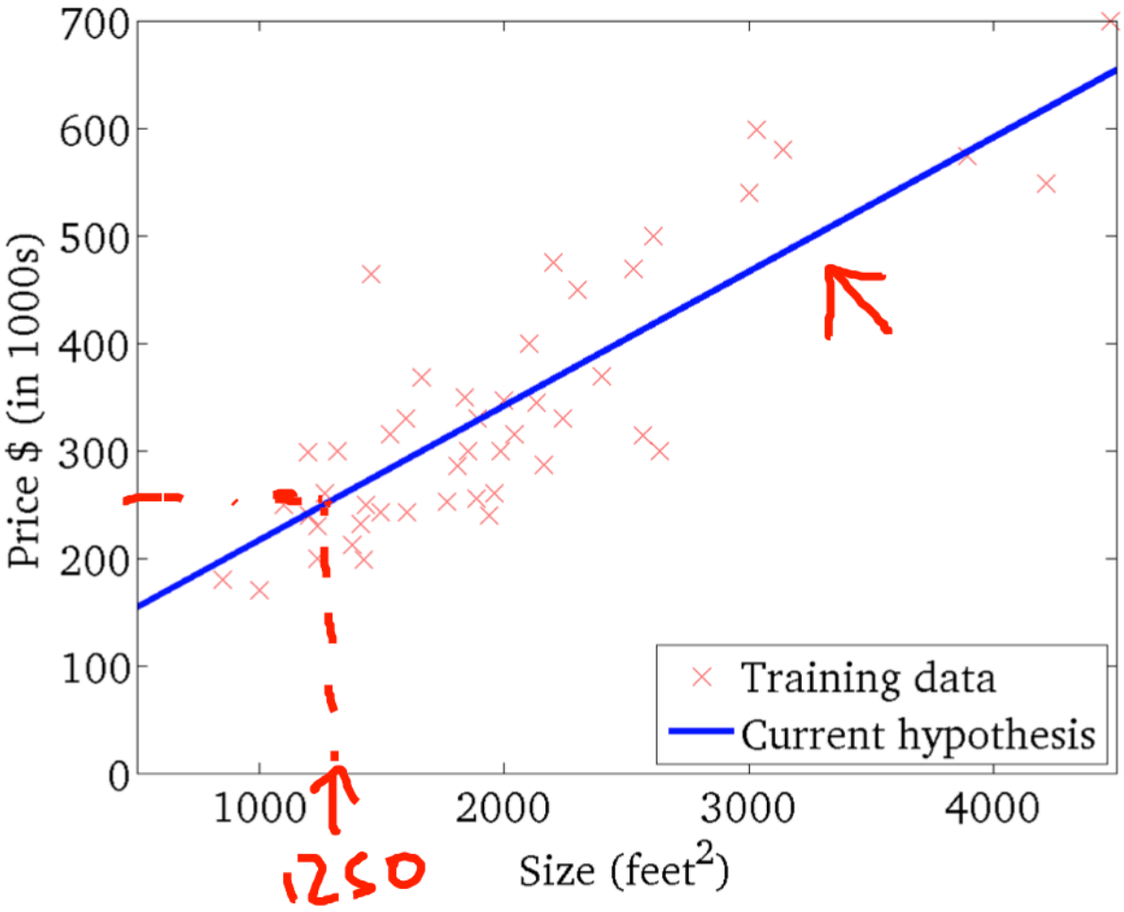
(function of the parameters θ_0, θ_1)



Courtesy: Andrew Ng

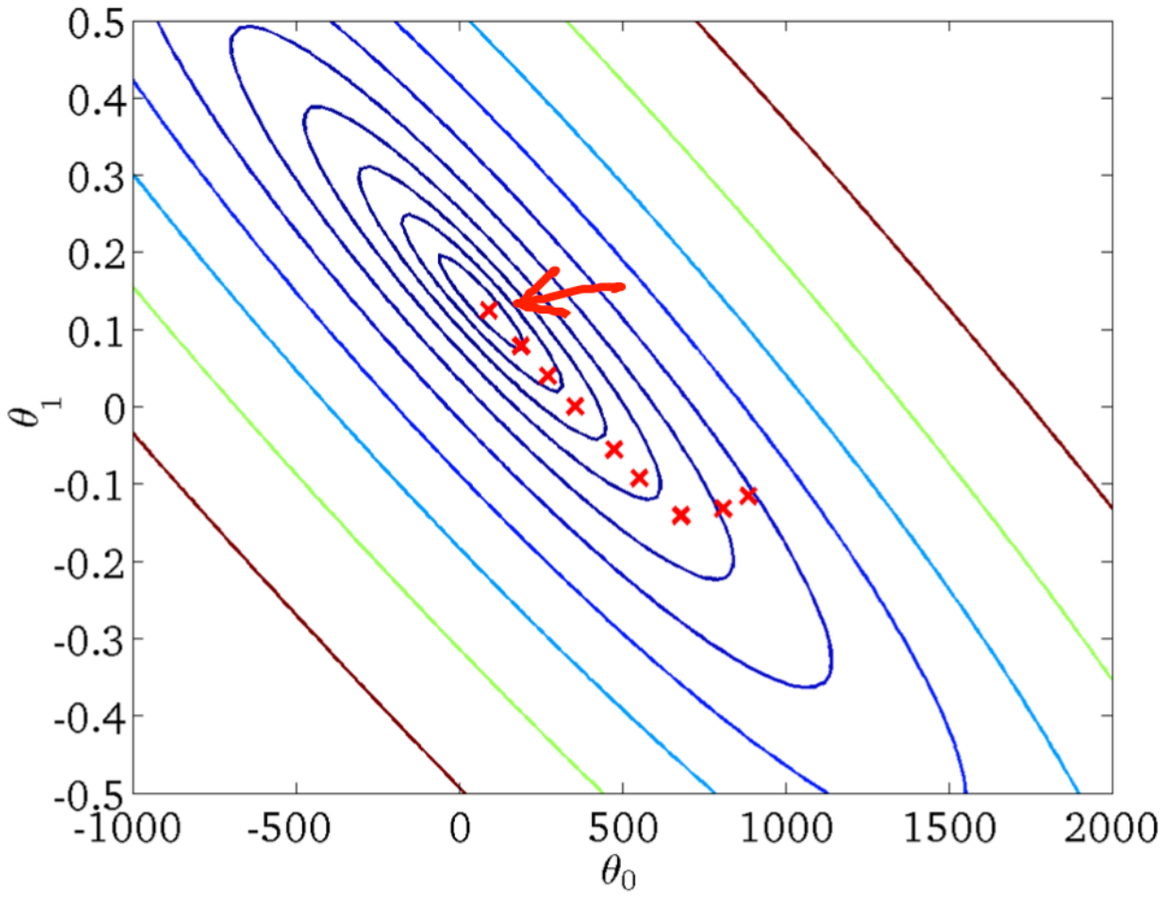
$h_\theta(x)$

(for fixed θ_0, θ_1, this is a function of x)



$J(\theta_0, \theta_1)$

(function of the parameters θ_0, θ_1)



Courtesy: Andrew Ng