GD : repeat {

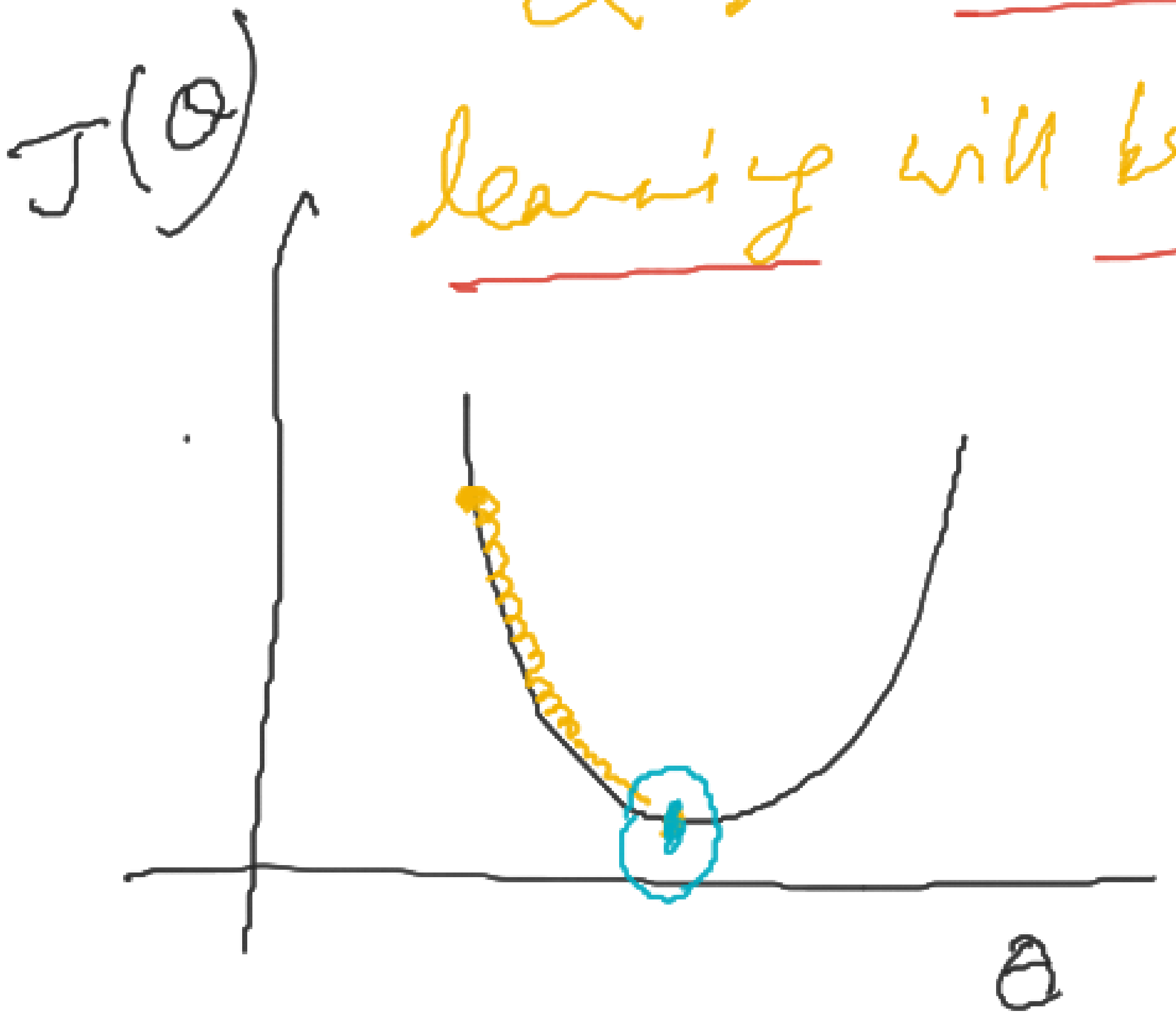
$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$$

}

$j = \{0, 1\}$

α is very small

learning will be done slowly




$\alpha = 0.1$

α is very high

may overshoot the minima

may fail to converge, even diverge

The purpose of GD is to optimize the cost func associated with linear reg.

Cost $f^n$ : $J_\theta = \frac{1}{2n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2$

where $\hat{y} = h_\theta(x_i) = \theta_0 + \theta_1 x_i$ (for single var. Lin Reg)

Aim

$$\frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1) = \frac{\partial}{\partial \theta_0} \left[ \frac{1}{2n} \sum_{i=1}^{n} (\theta_0 + \theta_1 x_i - y_i)^2 \right]$$

$$= \frac{1}{n} \sum_{i=1}^{n} (\theta_0 + \theta_1 x_i - y_i) \leftarrow$$

$$\frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1) = \frac{1}{n} \sum_{i=1}^{n} (\theta_0 + \theta_1 x_i - y_i) x_i$$

GP

for single variable

Repeat{

$$\theta_0 := \theta_0 - \alpha \frac{1}{n} \sum_{i=1}^{n} (\theta_0 + \theta_1 x_i - y_i)$$

$$\theta_1 = \theta_1 - \alpha \frac{1}{n} \sum_{i=1}^{n} (\theta_0 + \theta_1 x_i - y_i) x_i$$

}

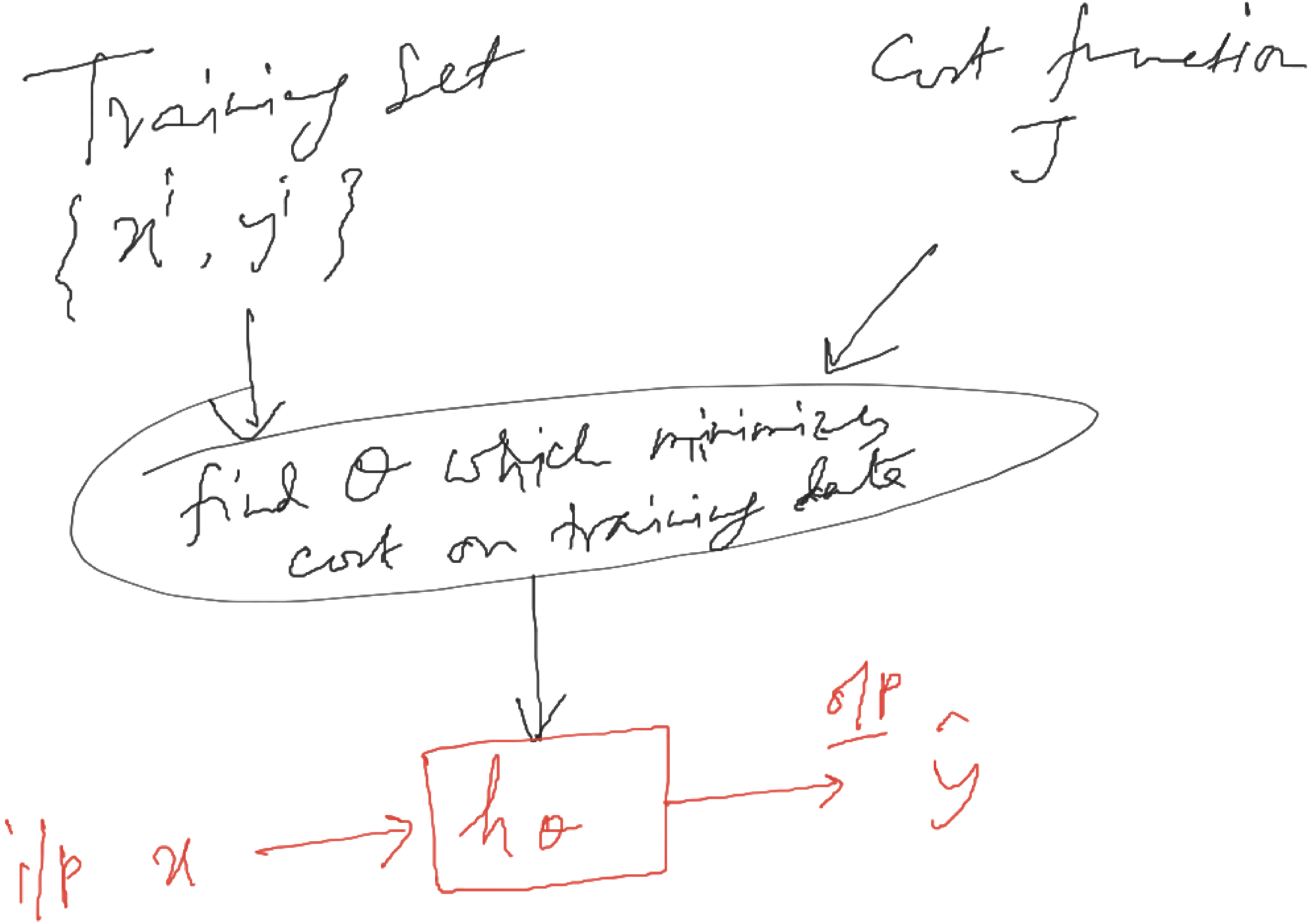
for multiple variable

Repeat ( ){

$$\theta_j := \theta_j - \alpha \frac{1}{n} \sum_{i=1}^{n} (\theta_0 + \theta_1 x_j^i - y^i) x_j^i$$

}

$x_0^i = 1$ ; $\forall j \in \{0, 1, 2, \ldots, k\}$


Training Set
{x^i, y^i}
Cost function
J
find θ which minimizes cost on training data
i/p x
h_θ
o/p
ŷ

Termination Criteria in GD

1) fixed # of iteration

2) min Gradient is reached

⋮

Polynomial Regression

Data Mining reference book: han, kamber, Pei