Lecture 9: 11 February, 2025

Madhavan Mukund

https://www.cmi.ac.in/~madhavan

Data Mining and Machine Learning
January–April 2025

Linear Regression

Inputs $(x_1, .., x_k)$

$$h(x) = \theta_0 + \theta_1 x_1 + \cdots + \theta_k x_k$$

Which attributes are significant?

If x_i is not "relevant" – expect $\theta_i \approx 0$

Coefficient size indicates significance

Scales must be uniform

Advisable to normalize all attributes

Gradient descent.

$$-\alpha \frac{\partial J}{\partial \theta_i} = -\alpha \cdot x_i$$

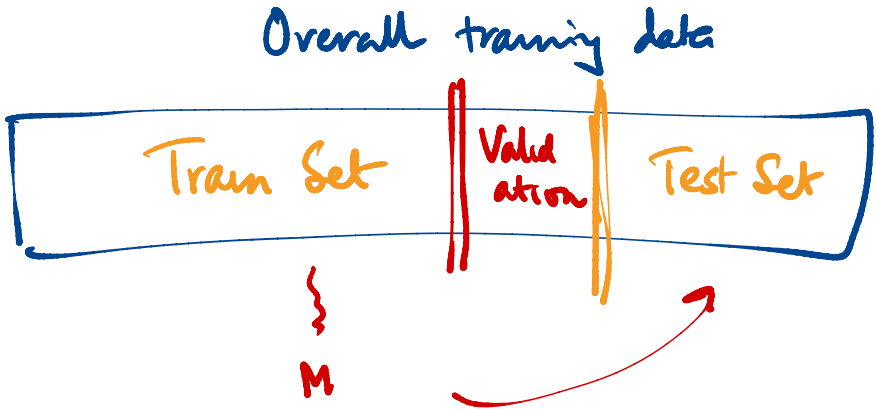
$x_0 = 1$

$h(x)$ $\theta_0 + \theta_1 x_1 \cdots + \theta_k x_k$

θ_0 ↑ 1

$$J = (h(x) - y)^2$$

Uniform α works if x_i is normalized


Overall training data
Train Set
Valid ation
Test Set
M