



Machine Learning

Dr. Mehran Safayani

safayani@iut.ac.ir

safayani.iut.ac.ir



<https://www.aparat.com/mehran.safayani>



https://github.com/safayani/machine_learning_course



Department of Electrical and computer engineering, Isfahan university of technology, Isfahan, Iran

ارتباط روش LS , MLE

Least square (LS)

یک تابع هزینه تعریف کردیم و با توجه به داده ها مدلی را پیدا کردیم که تابع هزینه را کمینه می کرد.

در این قسمت یک نگاه جدید داریم و می خواهیم از منظر مدل های احتمالاتی به این مسئله نگاه کنیم و به عبارتی یک تعبیر احتمالاتی از مسئله **LS** داشته باشیم.

یک مدل احتمالاتی برای LS

فرض کنید داده های ما توسط مدل زیر تولید می شوند:

$$y_n = x_n^T w + \varepsilon_n$$

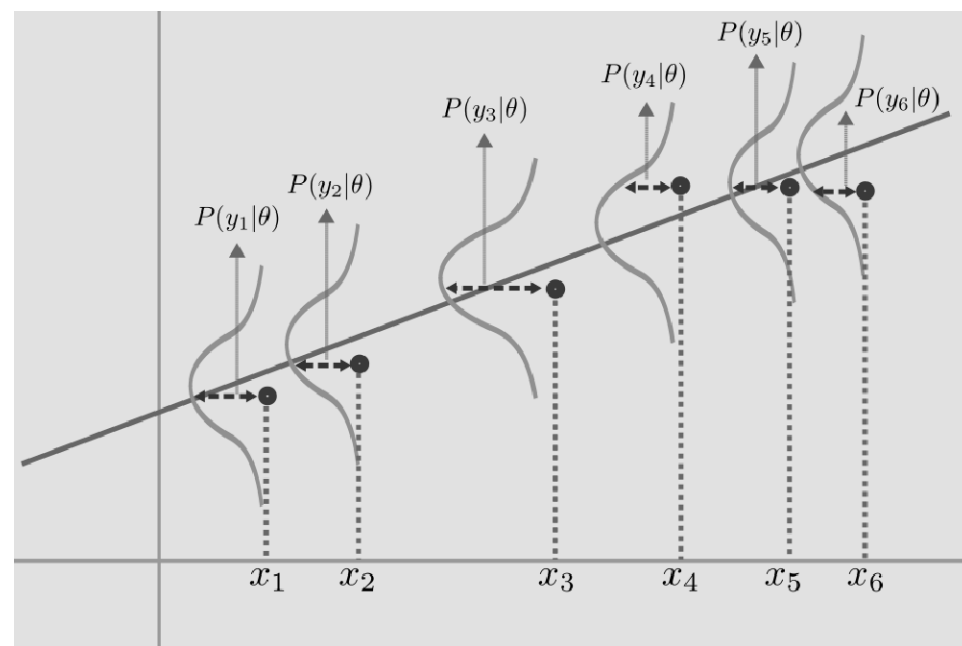
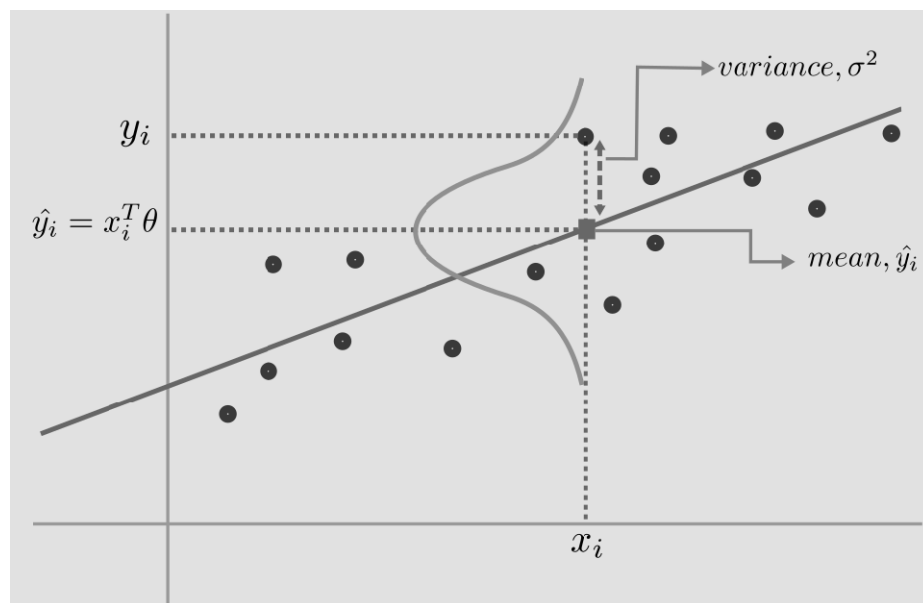
$$\varepsilon_n \sim N(\mu, \sigma^2):$$

ε_n یک نویز گاوسی با میانگین صفر و واریانس σ^2 است.
نویز با نمونه های تبدیل یافته جمع می شود و مستقل از نمونه هاست.

w : پارامترهای مدل است

$$P(y_n | x_n, w) = N(x_n^T w, \sigma^2)$$

$$P(y_n | x_n, w) = N(x_n^T w, \sigma^2)$$



ادامه یک مدل احتمالاتی برای LS

به شرط N نمونه درست نمایی (Likelihood) برای داده $Y = (y_1, y_2, \dots, y_n)$ با داشتن ورودی های X (هر سطر یک داده) و پارامترهای مدل w به صورت زیر است:

$$P(Y | X, w) = \prod_{n=1}^N P(y_n | x_n, w) = \prod_{n=1}^N N(y_n | x_n^T w, \sigma^2)$$

ما بایستی این Likelihood را نسبت به پارامترهای مدل w بیشینه کنیم. یعنی بهترین مدل مدلی است که این درست نمایی را بیشینه کند.

رابطه LS , log-likelihood

Log Likelihood:

$$L_{LL}(w) = \log P(y | X, w) = \frac{-1}{2\sigma^2} \sum_{n=1}^N (y_n - x_n^T w)^2 + \text{con}$$

LS:

$$L_{MSE}(w) = \frac{1}{2N} \sum_{n=1}^N (y_n - x_n^T w)^2$$

$$\underset{w}{\operatorname{argmin}} L_{MSE}(w) = \underset{w}{\operatorname{argmax}} L_{LL}(w)$$