Derivation of Linear Regression

Friday, June 7, 2024 11:20 AM

for analytic solution of the linear regression problem,

$$L(w,b) = \frac{1}{n}\sum_{i=1}^{n}\left(w^T x^{(i)} + b - y^{(i)}\right)^2.$$

$\longrightarrow$ find w^*, b^*, to minimize $L(w,b)$

firstly, minimize $\frac{1}{n}\sum_{i=1}^{n}\left(w^T x^{(i)} - y^{(i)} + b\right)^2$.

we can use the following matrix:

$$\min_{w,b} [w_1, w_2, \cdots, w_n, b] \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1m}, & 1 \\ X_{21} & \cdots & \cdots & X_{2m}, & 1 \\ \vdots & & & & \\ X_{n1} & \cdots & \cdots & X_{nm}, & 1 \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\downarrow$$

$$\min \frac{1}{2} \|Y - w^T X\|^2 \xrightarrow{\text{derivate}} (Y - wX) = 0$$

but note that $(Y - w^T X = 0)$ is impossible, thus we use

$$X^T(Y - Xw) = 0 \Rightarrow X^T Y = X^T X w$$, also X is full-rank matrix.

Then we have : $\boxed{w = [X^T X]^{-1} X^T Y}$ analytic solution.

or $[w, b] = [X^T X]^{-1} X^T Y$

Note : the matrix $X^T X$ should be invertible. (columns are linearly independent.)