Day 1 January 18

Design/feature-matrix

$X \in \mathbb{R}^{n \times p}$

n = # data entries

for example # of patients

p = # features

Regression-linear:

data (x_i, y_i)

ideally:

$$\boxed{y_i = f(\underline{x_i}) + \varepsilon_i}$$

output/target

continuous function

x_i = inputs

normally distributed noise $\sim N(0, \sigma^2)$

mean $\mu = 0$

variance σ^-

$f(x_i)$ is unknown !

$\tilde{y}_i \approx f(x_i)$

Model linear in unknown parameters $\beta = \{\beta_0, \beta_1, \dots \beta_{p-1}\}$

Example : polynomial of degree - p -

$$\tilde{y}_i = \sum_{j=0}^{p-1} \beta_i x_i^j$$

$$= \beta_0 + \beta_1 x_i^1 + \beta_2 x_i^2 + \dots \beta_{p-1} x_i^{p-1}$$

$\underline{y} = \{y_0, y_1, y_2, \dots, y_{n-1}\}$

$\underline{\tilde{y}} = \{\tilde{y}_0, \tilde{y}_1, \tilde{y}_2 \dots \tilde{y}_{n-1}\}$

$$\tilde{y}_0 = \beta_0 + \beta_1 x_0^1 + \beta_2 x_0^2 + \dots + \beta_{p-1} x_0^{p-1}$$

$$\tilde{y}_1 = \beta_0 + \beta_1 x_1^1 + \beta_2 x_1^2 + \dots \beta_{p-1} x_1^{p-1}$$

$\vdots$

$$\tilde{y}_{n-1} = \beta_0 + \beta_1 x_{n-1}^1 + \dots \beta_{p-1} x_{n-1}^{p-1}$$

$$\boxed{\tilde{y} = X\beta}$$

$$\beta \in \mathbb{R}^p \qquad \tilde{y}, y \in \mathbb{R}^n$$

$$X = \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^{p-1} \\ 1 & x_1 & x_1^2 & \dots & x_1^{p-1} \\ 1 & x_2 & x_2^2 & \dots & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \dots & x_{n-1}^{p-1} \end{bmatrix}$$

Define an error:

Mean-Squared Error = MSE

$$MSE(y, X, \beta) = C(y \mid X, \beta)$$

$$= \frac{1}{n} \sum_{i=0}^{n-1} (y_i - \tilde{y}_i)^2$$

$$\beta^{opt} = \hat{\beta} = \underset{\beta}{\operatorname{argmin}}$$

$$\sqrt{\frac{1}{n}\sum_{i=0}^{n-1}\left(y_i-\tilde{y}_i\right)^2}$$

$$\tilde{y} = X\beta$$

$$\tilde{y}_i = X_{i*}\beta$$

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^{p-1} \\ \vdots & & & & \\ 1 & x_i & x_i^2 & \cdots & x_i^{p-1} \\ \vdots & & & & \\ 1 & x_{n-1} & & \cdots & x_{n-1}^{p-1} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix}$$

$$MSE = \frac{1}{n}\sum_{i=0}^{n-1}\left(y_i - X_{i*}\beta\right)^2$$

$$= \frac{1}{n}\left(y - X\beta\right)^T\left(y - X\beta\right)$$

$$= C(y|X,\beta) = C(\beta)$$

Tomorrow:

$$\frac{\partial C(\beta)}{\partial \beta} = 0 \Rightarrow$$

$$\boxed{\beta^{opt} = \hat{\beta} = \underline{(X^T X)}^{-1} \underline{X}^T \underline{y}}$$

known

known

$$X \in \mathbb{R}^{n \times p}$$

$$X^T X \in \mathbb{R}^{p \times p}$$

$$X^T \in \mathbb{R}^{p \times n}$$

$$y \in \mathbb{R}^{n}$$

$$\beta \in \mathbb{R}^{p}$$

$$\tilde{y} = \underline{X \beta^{opt}}$$