• Bùi Phương Thảo • 11215341

# Homework Week 3

## Bài 1. Linear regression

Các biến độc lập <independent variable> $x = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$

Các biến phụ thuộc <dependent variable> $t = \begin{bmatrix} t_1 \\ t_2 \\ \dots \\ t_n \end{bmatrix}$

n: Số quan sát

Vì đây là bài toán dự đoán giá trị của t dựa trên giá trị đã biết của x, giả sử các điểm dữ liệu độc lập tuyến tính và có cùng phân phối → $t = y(x, w) + \text{noise}$

<noise là sai số của model với giá trị thực>

Vì noise ngẫu nhiên nên assume noise tuân theo normal dis

$\to t = y(x, w) + N(\mu, \sigma^2)$

$\to t = N(y(x, w) + \mu, \sigma^2)$

Simplify: $t = N(y(x, w), \sigma^2)$ / $t = N(y(x, w), \beta^{-1})$

$\beta = \frac{1}{\sigma^2}$

→ Hàm likelihood của bộ dữ liệu

$$p(t \mid x, w, \beta) = \prod_{n=1}^{N} N(t_i \mid y(x_i, w_i), \beta^{-1}) \quad (1)$$

Cần tìm maximum likelihood, dùng log

$$(1) \to p(t \mid x, w, \beta) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(t_i - y(x_i, w))^2}{2\sigma^2}\right)$$

$$L = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(t_i - y(x_i, w))^2}{2\beta}\right)$$

$$\log L = \sum_{n=1}^{N} \log \frac{1}{\sqrt{2\pi\delta^2}} + \left(-\frac{(t_i - y(x_i,w))^2}{2\beta}\right)$$

$$= \underbrace{N.\log \frac{1}{\beta^{1/2}\sqrt{2\pi}}}_{\text{const}} - \underbrace{\frac{1}{2\beta}}_{\text{const}} \sum_{i=1}^{N} (y(x_i,w) - t_i)^2$$

$\Rightarrow$ Để xét maximum $\log L$ cần tìm minimum $\sum_{i=1}^{N} (y(x_i,w) - t_i)^2$

Đặt $P = \sum_{i=1}^{N} (y(x_i,w) - t_i)^2$

Ta có

$$t = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix} ;\quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} w_1x_1 + w_0 \\ w_2x_2 + w_0 \\ \vdots \\ w_nx_n + w_0 \end{bmatrix} ;\quad w = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$$

$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$, để dễ nhân x với w, thêm cột $\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$, $x \to \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$

$$\to y = x \cdot w$$

$$t - y = \begin{bmatrix} t_1 - y_1 \\ t_2 - y_2 \\ \vdots \\ t_n - y_n \end{bmatrix} \Rightarrow \|t-y\|_2^2 = (t_1-y_1)^2 + \dots + (t_n-y_n)^2 = \sum^{n} (t_i - y_i)^2 = P$$

$$\to P = \|t-y\|_2^2 = \|t - xw\|_2^2 = (xw-t)^T(xw-t)$$

$P = ((xw)^T - t^T)(xw - t)$

$= (w^Tx^T - t^T)(xw - t)$

$= w^Tx^Txw - w^Tx^Tt - t^Txw + tt^T$

$= w^Tx^Txw - 2w^Tx^Tt + t^Tt$

$t^T_{(1,m)} x_{(m,n+1)} w_{(n+1,1)} = w^T_{(1,n+1)} x^T_{(n+1,m)} t_{(m,1)}$

Unit vector

$\frac{\partial P}{\partial w} = \frac{\partial}{\partial w}(w^Tx^Txw - 2w^Tx^Tt_j + tt^T)$

$= x^T x (w^Tw)' - 2x^Tt(w^T)' + 0$

$= x^Tx\,2w - 2x^Tt = 0$

$\Rightarrow x^Txw = x^Tt$

$\Rightarrow w = \frac{x^Tt}{x^Tx} = (x^Tx)^{-1}(x^Tt)$

Ex 4.

Rank of a matrix : # of cols /rows ≠ null

Full rank  Rank =

X full rank →

Chứng minh $X^T X$ invertible khi X full rank

Giả sử $X_{(m \times n)} \rightarrow X^T_{(n \times m)} \rightarrow X^T X$ là ma trận vuông ①

Lấy $\vec{v} \in N(X^T X) \rightarrow \vec{v} . X^T X = \vec{0}$

$$\vec{v} . X^T . X . \vec{v}^T = \vec{0} . \vec{v}^T = \vec{0}$$

$$(X\vec{v})^T X . \vec{v} = \vec{0}$$

$$X\vec{v} . X\vec{v} = 0$$

$$\rightarrow |Xv|^2 = 0.$$

Vì $\vec{v} \in N(X^T X) \rightarrow \vec{v} \in N(X) \rightarrow \vec{v} = 0$

→ X độc lập tuyến tính , $X^T X$ cũng vậy ②

①② ⇒ $X^T X$ khả nghịch.