FYS-STK 4155 AUG 26

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

$\hat{\beta} \in \mathbb{R}^d \qquad y \in \mathbb{R}^n$

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

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

$d \ll n$

$n = 10000 \qquad d = 10$

Derivatives/gradients

$f \Rightarrow f(\vec{x})$

$$\vec{x} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} \in \mathbb{R}^n$$

$$\text{grad} f = \vec{\nabla} f = \begin{bmatrix} \frac{\partial f(\vec{x})}{\partial x_0} \\ \vdots \end{bmatrix}$$

$$\begin{bmatrix} \vdots \\ \frac{\partial f(\vec{x})}{\partial x_{n-1}} \end{bmatrix}$$

multivalued function

$$\vec{f}(\vec{x}) = \begin{bmatrix} f_0(\vec{x}) \\ f_1(\vec{x}) \\ \vdots \\ f_{m-1}(\vec{x}) \end{bmatrix} \in \mathbb{R}^m$$

$$\text{grad } \vec{f} = \begin{bmatrix} \frac{\partial f_0(\vec{x})}{\partial x_0} & \cdots & \frac{\partial f_0}{\partial x_{n-1}} \\ \vdots & & \vdots \\ \frac{\partial f_{m-1}}{\partial x_0} & & \frac{\partial f_{m-1}}{\partial x_{n-1}} \end{bmatrix}$$

$$= J \in \mathbb{R}^{m \times n}$$

$$J_{ij} = \frac{\partial f_i}{\partial x_j}$$

$$f(x) = Ax$$

$$A \in \mathbb{R}^{m \times n} \qquad x \in \mathbb{R}^n$$

$n-1$

$$f_i(x) = \sum_{j=0} a_{ij} x_j = f_i$$

$$A = \begin{bmatrix} a_{00} & a_{01} & \cdots & a_{0\,n-1} \\ \vdots & & & \\ a_{n-1\,0} & & \cdots & a_{n-1\,n-1} \end{bmatrix}$$

$$\frac{\partial f_i}{\partial x_j} = a_{ij}$$

List of useful expression

$$\frac{\partial x^T a}{\partial x} = a^T$$

$$\frac{\partial a^T x b}{\partial x} = a b^T$$

$$\frac{\partial x^T B x}{\partial x} = x^T (B + B^T)$$

$$\frac{\partial}{\partial s} (x - As)^T W (x - As) =$$

$\uparrow$ symmetric

$$W = I = [0, \ldots, 1]$$

$$(x - As)^T (x - As)$$

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

$$= -2(x - As)^T W A$$

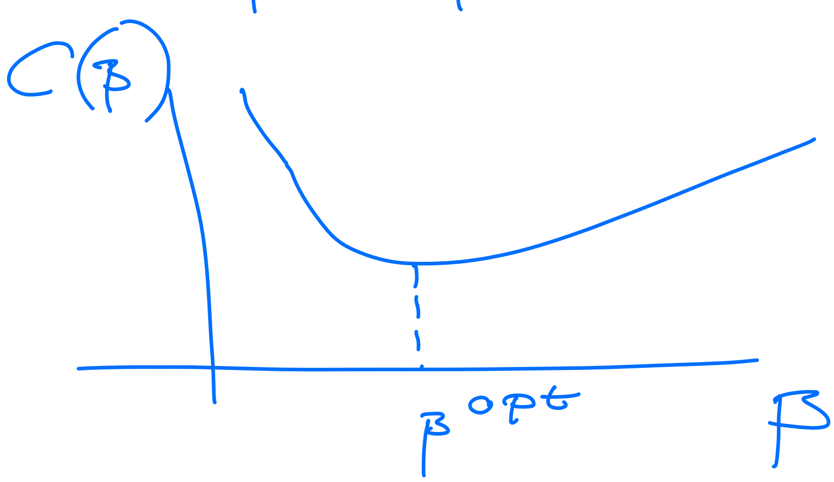
<u>MSE</u> : $\frac{1}{n}(y - X\beta)^T (y - X\beta)$

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

$$= C(\beta)$$

$$\frac{\partial^2 C(\beta)}{\partial \beta^T \partial \beta} = \frac{2}{n} X^T X$$

Exercise



convex

$$\frac{\partial^2 C}{\partial \beta^2} > 0$$

Example

$$X^T X = I$$

$$C(\beta) = ?$$

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

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

$$C(\beta) = 0$$