Lecture August 31 Lineær regression analyses.

$$x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad x_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$x_0^T x_1 = 0$$

$$\hat{P}_0 = x_0 x_0^T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad \hat{P}_0^2 = \hat{P}_0$$

$$\hat{P}_1 = x_1 x_1^T = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \qquad \hat{P}_1^2 = \hat{P}_1$$

$$\alpha = a x_0 + b x_1 = \begin{bmatrix} a \\ b \end{bmatrix}$$

$$\hat{P}_0 \alpha = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ 0 \end{bmatrix} = a x_0$$

$$\hat{P}_0 + \hat{P}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbb{I}$$

<u>SVD</u>

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

$$X = U \Sigma V^T$$

$$U \in \mathbb{R}^{n \times n} \qquad U^T U = U U^T = \mathbb{1}$$

$$V \in \mathbb{R}^{p \times p} \qquad V V^T = V^T V = \mathbb{1}$$

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

$$\Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad \underset{2\times 2}{\Sigma^T\Sigma} = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_1^2 \end{bmatrix}$$

$$\sigma_0 > \sigma_1 > \cdots > \sigma_{p-1} > 0$$

$$\underline{MSE} \quad \frac{\partial^2 C}{\partial \beta \partial \beta^T} = \frac{2}{n} X^T X$$

$$X = U \Sigma V^T$$

$$X^T X = V \Sigma^T \underbrace{U^T U}_{\mathbb{I}} \Sigma V^T = V \underbrace{\Sigma^T \Sigma}_{\Sigma^2} V^T$$

$$X^T X = V \Sigma^2 V^T$$

$$\Sigma^2 \in \mathbb{R}^{p \times p} \qquad \begin{bmatrix} \sigma_0^2 & & & \\ & \sigma_1^2 & & \\ & & \ddots & \\ & & & \sigma_{p-1}^2 \end{bmatrix}$$

$$V = \begin{bmatrix} v_0 & v_1 & v_2 & \cdots & v_{p-1} \end{bmatrix}$$

$$v_i^T v_j = \delta_{ij}$$

$$X^T X V = V \Sigma^2 \underbrace{V^T V}_{I} = V \Sigma^2$$

$$X^T X v_i = v_i \sigma_i^2 \Rightarrow X^T X u -$$

presents a convex optimization problem since $\sigma_0^2 > \sigma_1^2 > \sigma_2^2 \cdots \sigma_{p-1}^2 \Rightarrow X^T X$ is positive definite

$$\tilde{y}_{OLS} = \tilde{y}_{opt} = \tilde{y} = X\hat{\beta}$$

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

$$\tilde{y} = \underbrace{X (X^T X)^{-1} X^T}_{A} y \qquad X = U\Sigma V^T$$

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

$$= U \Sigma V^T \left(\frac{1}{V \Sigma^2 V^T} \right) V \Sigma^T U^T y$$

$$\Sigma^2 = \begin{bmatrix} \sigma_0^2 & & \\ & \ddots & \\ & & \sigma_{p-1}^2 \end{bmatrix} \in \mathbb{R}^{p \times p}$$

$$V, V^T \in \mathbb{R}^{p \times p} \qquad \Sigma \in \mathbb{R}^{n \times p}$$

$$U \in \mathbb{R}^{n \times n}$$

$$V I = I V$$

$$V \Sigma^2 V^T = \underbrace{V V^T}_{I} \Sigma^2 = \Sigma^2$$

$$\Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{3\times 2} \quad \sigma_2 = 0$$

$$U = \begin{bmatrix} u_0 & u_1 & u_2 \end{bmatrix} = \begin{bmatrix} u_{00} & u_{01} & u_{02} \\ u_{10} & u_{11} & u_{12} \\ u_{20} & u_{21} & u_{22} \end{bmatrix}$$

$$\Sigma^T U^T = \begin{bmatrix} \sigma_0 u_{00} & \sigma_0 u_{10} & \sigma_0 u_{11} \\ \sigma_1 u_{01} & \sigma_1 u_{11} & \sigma_1 u_{21} \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_0 u_0 \\ \sigma_1 u_1 \end{bmatrix}$$

$$\tilde{y} = \sum_{i=0}^{p-1} u_i u_i^T y$$

$X^T X \propto$ covariance matrix.

Basic statistical quantities

$$\mathbb{E}[x^n] = \langle x^n \rangle = \int_{x \in D} p(x) x^n dx$$

$$\int_{x \in D} p(x) dx = 1$$

Discrete case

$$\mathbb{E}[x^n] = \sum_{x_i \in D} p(x_i) x_i^n$$

sample average

$$\mathbb{E}[x^n] = \frac{1}{n} \sum_{x_i \in D} x_i^n$$

$i = 0, 1, 2, \ldots, n-1$

$$\mu_x = \int p(x)\, x\, dx \neq \bar{\mu}_x = \frac{1}{n} \sum_i x_i$$

variance

$$\sigma_x^2 = \int dx\, p(x) (x - \mu_x)^2$$

sample variance

$$\bar{\sigma}_x^2 = \frac{1}{n} \sum_i (x_i' - \bar{\mu}_x)^2$$

sample average

covariance

$$\mathrm{cov}(x_i x_j) = \int dx_i' \int dx_j'$$

$$\times\, p(x_i x_j)(x_i' - \mu_{x_i})(x_j' - \mu_{x_j})$$

i.i.d = independent and identically distributed

$$P(x_i, x_j) = \underline{P(x_i)}\,\underline{P(x_j)}$$

same function

$$\operatorname{cov}(x_i, x_j) = 0 \qquad \mu_{x_i} = \mu_{x_j}$$

sample covariance

$$\operatorname{cov}(\underset{x_i}{x_i}, \underset{x_j}{y_j}) = \frac{1}{n}\sum_{k=0}^{n-1}(x_{ki} - \mu_{x_i})(y_{kj} - \mu_{y_j})$$

$$x_0 = \begin{bmatrix} x_{00} \\ x_{10} \end{bmatrix} \qquad x_1 = \begin{bmatrix} x_{01} \\ x_{11} \end{bmatrix}$$

$$X = \begin{bmatrix} x_0 & x_1 \end{bmatrix} = \begin{bmatrix} x_{00} & x_{01} \\ x_{10} & x_{11} \end{bmatrix}$$

$$\mathrm{cov}(x_0, x_0) = \frac{1}{2} \sum_{k=0}^{1} (x_{k0} - \mu_0)(x_{k0} - \mu_0)$$

$$= ? \qquad \sigma_{x_0}^2$$

$$\mathrm{cov}(x_0\, x_1) = \frac{1}{2} \sum_{k=0}^{1} (x_{k0})\,(x_{k1})$$

$$= x_{00} x_{01} + x_{10} x_{11} = \mathrm{cov}(x_1\, x_0)$$

cov-matrix for $X = [x_0 \; x_1]$

$$= C[X] = \begin{bmatrix} var[x_0] & cov[x_0 x_1] \\ cov[x_0 x_1] & var[x_1] \end{bmatrix}$$

$$= \frac{1}{n} X^T X \in \mathbb{R}^{p \times p}$$

$\rightarrow$ singular values of $X = U \Sigma V^T$ as eigenvalues

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