12. $X \in \mathbb{R}^{n\times p}$ $Rank(X)=p$ $P_x = X(X^TX)^{-1}X^T$

(1) $y \in \mathbb{R}^n$ 在 $\mu(X)$ 上投影为 $= P_x y$

(2) $P_x = (h_{ij})_{n\times n}$, $0 \le h_{ii} \le 1$, $i=1,\cdots n$

(3) 若 X 第一列元素全是1, 则 $h_{ii} \geq \frac{1}{n}$ $(i=1,\cdots n)$

(1) $(P_x y,\ y-P_x y)$

$= y^T P_x y - \|P_x y\|^2$

$= y^T X(X^TX)^{-1}X^Ty - y^TX(X^TX)^{-1}X^TX(X^TX)^{-1}X^Ty$

$= y^TX(X^TX)^{-1}X^Ty - y^TX(X^TX)^{-1}X^Ty = 0$

又设 $X=[X_1,\cdots X_p]$, $X_i \in \mathbb{R}^n$

$P_x y = X[(X^TX)^{-1}X^Ty] = [X_1\cdots X_p]\begin{pmatrix} c_1\\ \vdots \\ c_p\end{pmatrix} = c_1X_1+\cdots+c_pX_p \in \mu(X)$

(2) 设 $e_i = \begin{pmatrix} 0\\ \vdots \\ 1 \\ \vdots \\ 0\end{pmatrix}$, 其中1在第 i 位

$h_{ii} = e_i^T P_x e_i = e_i^T P_x P_x e_i = (P_x e_i)^T (P_x e_i) \geq 0$

$1-h_{ii} = (I_n - P_x)_{ii} = e_i^T(I_n-P_x)e_i = e_i^T(I_n-P_x)(I_n-P_x)e_i = \|(I_n-P_x)e_i\|_2^2 \geq 0$

(~~易知~~ $P_x = P_x^2$, $(I-P_x)(I-P_x) = I - 2P_x + P_x^2 = I-P_x$).

(3) 考虑 $C = 1_n(1_n^T1_n)^{-1}1_n^T = \frac{1}{n}\begin{pmatrix} 1 & \cdots & 1\\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1\end{pmatrix}$

记 $R_x = P_x - C$ $R_x^T = R_x$

$R_x^2 = P_x^2 - P_xC - CP_x + C^2 = P_x - P_xC - CP_x + C$

由于 $C \in Col(X)$, 有 $P_x C = C$

$C = C(P_x + (I-P_x)) = CP_x$ $(C(I-P_x)y = 0,\ \forall y \in \mathbb{R}^n)$

$\Rightarrow R_x^2 = P_x - C = R_x$ R_x 特征值 $\in \{0,1\}$.

故 $R_{ii} = \frac{e_i^TR_ie_i}{e_i^Te_i} \in [0,1]$

$h_{ii} = R_{ii} + \frac{1}{n} \Rightarrow h_{ii} \geq \frac{1}{n}$ #

14. $\begin{cases} Y = \theta + e \\ A\theta = 0 \end{cases}$

$A \in \mathbb{R}^{r\times n}$, $n > r$ $\quad \mathrm{rank}(A) = r$

试证 θ 的最小二乘估计为 $\hat{\theta} = [I - A^T(AA^T)^{-1}A]Y$

$\hat{\theta} = \arg\min \|Y-\theta\|_2^2 \quad s.t.\ A\theta = 0$

Lagrangian $\mathcal{L}(\theta, \lambda) = \|Y-\theta\|_2^2 + \lambda^T A\theta$, $\lambda \in \mathbb{R}^{r\times 1}$

KKT $\begin{cases} \nabla_\theta \mathcal{L}(\theta,\lambda) = -2Y + 2\theta + A^T\lambda = 0 \quad ① \\ A\theta = 0 \quad ② \end{cases}$

$① \Rightarrow \quad AA^T\lambda = 2AY - 2A\theta = 2AY$

$\lambda = 2(AA^T)^{-1}AY$

代回①, $\hat{\theta} = Y - \frac{1}{2}A^T\lambda = Y - A^T(AA^T)^{-1}AY$

$= [I - A^T(AA^T)^{-1}A]Y$ #

15. $\begin{cases} Y = X\beta + e \\ H\beta = r \end{cases}$

$X \in \mathbb{R}^{n\times p}$, $n > p$, $\quad H \in \mathbb{R}^{q\times p}$, $q < p$, $\mathrm{rank}(X) = p$, $\quad \mathrm{rank}(H) = q$

β 最小二乘估计 $\hat{\beta} = \tilde{\beta} + (X^TX)^{-1}H^T[H(X^TX)^{-1}H^T]^{-1}(r - H\tilde{\beta})$

其中 $\tilde{\beta} = (X^TX)^{-1}X^TY$ 为无约束最小二乘估计

$\hat{\beta} = \arg\min_\beta \|Y - X\beta\|_2^2 \quad s.t.\ H\beta = r$

Lagrangian $\mathcal{L}(\beta, \lambda) = \|Y - X\beta\|_2^2 + \lambda^T(H\beta - r) \quad \lambda \in \mathbb{R}^{q\times 1}$

KKT $\begin{cases} \nabla_\beta \mathcal{L}(\beta,\lambda) = 2X^TX\beta - 2X^TY + H^T\lambda = 0 \quad ① \\ H\beta = r \quad ② \end{cases}$

由①: $\beta = (X^TX)^{-1}(X^TY - \frac{1}{2}H^T\lambda) = \tilde{\beta} - \frac{1}{2}(X^TX)^{-1}H^T\lambda$

$r = H\beta = H(X^TX)^{-1}(X^TY - \frac{1}{2}H^T\lambda)$

$r = H(X^TX)^{-1}X^TY - \frac{1}{2}H(X^TX)^{-1}H^T\lambda$

$r - H\tilde{\beta} = -\frac{1}{2}H(X^TX)^{-1}H^T\lambda$

$\lambda = -2[H(X^TX)^{-1}H^T]^{-1}(r - H\tilde{\beta})$

$\Rightarrow \beta = \tilde{\beta} + (X^TX)^{-1}H^T[H(X^TX)^{-1}H^T]^{-1}(r - H\tilde{\beta})$ #

(易知 $H(X^TX)^{-1}H^T$ 可逆, 与证 X^TX 可逆方法一样)

16.

x_1	1.9	0.8	1.1	0.1	-0.1	4.4	4.6	1.6	5.5	3.4
x_2	66	62	64	61	63	70	68	62	68	66
y	0.7	-1	-0.2	-1.2	-0.1	3.4	0.0	0.8	3.7	2.0

设 $\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ 求 $\beta_0, \beta_1, \beta_2$. 并估计 σ^2

$$y = \begin{pmatrix} y_{1,1} \\ \vdots \\ y_{1,10} \end{pmatrix} = \begin{pmatrix} 1 & x_{1,1} & x_{2,1} \\ \vdots & \vdots & \vdots \\ 1 & x_{1,10} & x_{2,10} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} + \varepsilon$$

记 $X = \begin{pmatrix} 1 & x_{1,1} & x_{2,1} \\ \vdots & \vdots & \vdots \\ 1 & x_{1,10} & x_{2,10} \end{pmatrix}$ 验证知 $\text{rank}(X) = 3$

$$\hat{\beta} = (X^T X)^{-1} X^T y = \begin{pmatrix} -11.45278793 \\ 0.45077533 \\ 0.17251714 \end{pmatrix}$$

$\Rightarrow \beta_0 = -11.45 \quad \beta_1 = 0.45 \quad \beta_2 = 0.17$

我们有 $\hat{\sigma}^2 = \dfrac{Q(\hat{\beta})}{n-p-1} = \dfrac{1}{7} Q(\hat{\beta})$

$$Q(\hat{\beta}) = \| Y - X\hat{\beta} \|_2^2 = 8.8654$$

$$\hat{\sigma}^2 = 1.2665$$