1. As we know, LSE: $\hat{\beta} = (X^T X)^{-1} X^T y$.

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

$$E[\hat{\theta}] = E[a^T (X^T X)^{-1} X^T y] = a^T (X^T X^{-1}) X^T E[y]$$

$$\because\ y = X\beta + \varepsilon,\quad E[\varepsilon] = 0$$

$$\Rightarrow E[\hat{\theta}] = a^T (X^T X)^{-1} (X^T X)\beta + 0 = a^T \beta$$

$$\tilde{\theta} = a^T (X^T X)^{-1} X^T y + d^T y = c^T y$$

$$c^T = a^T (X^T X)^{-1} X^T + d^T,\quad y = X\beta + \varepsilon$$

$$\Rightarrow E[\tilde{\theta}] = E[a^T (X^T X)^{-1} X^T (X\beta) + a^T (X^T X)^{-1} X^T \varepsilon) + d^T (X\beta) + d^T \varepsilon].$$

$$= E[a^T \beta] + E[d^T X \beta]$$

$$\Rightarrow E[\tilde{\theta}] = a^T \beta + d^T X \beta$$

$\Rightarrow$ iff unbiased, $(d^T X) = 0$

$$E[\tilde{\theta}] = a^T \beta + d^T X \beta = a^T \beta$$

$$var[\tilde{\theta}] = var(c^T y) = c^T var(y) = \sigma^2 (c^T c)$$

$$= \sigma^2 (a^T (X^T X)^{-1} X^T + d^T)(a^T (X^T X)^{-1} X^T + d^T)^T$$

$$= \sigma^2 [a^T (X^T X)^{-1} X^T + d^T](X (X^T X)^{-1} a + d)$$

$$= \sigma^2 (a^T (X^T X)^{-1} a + d^T d)$$

$$var(\hat{\theta}] = var(a^T (X^T X)^{-1} X y)$$

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

$$= \sigma^2 [a^T (X^T X)^{-1} X][X^T (X^T X)^{-1} a]$$

$$= \sigma^2 (a^T (X^T X)^{-1} a)$$

$$\Rightarrow var[\tilde{\theta}] = var(\hat{\theta}] + \sigma^2 d^T d$$

$$\therefore var(c^T y) = var(a^T \beta) + \sigma^2 d^T d$$

$$\therefore var(c^T y) \geq var(a^T y)$$

proved

2. Given that the columns of the design matrix x are orthogonal. Let the columns of X be $x_1, x_2, \cdots x_p$. we have $x_i^T x_j = 0, i \neq j$

As we know

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad \text{where } X = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

$$\Rightarrow X^T \cdot X = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_p \end{bmatrix} \cdot [x_0, x_1 \cdots x_p] = \begin{bmatrix} x_0^2 & 0 & 0 & \cdots & 0 \\ 0 & x_1^2 & 0 & \cdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & & \cdots & x_p^2 \end{bmatrix}$$

$$\Rightarrow (X^T X)^{-1} = \begin{bmatrix} \frac{1}{x_0^2} & 0 & \cdots & 0 \\ \vdots & \frac{1}{x_1^2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & \frac{1}{x_p^2} \end{bmatrix}$$

$$\Rightarrow (X^T X)^{-1} \cdot X^T = \begin{bmatrix} \frac{1}{x_0^2} & 0 & \cdots & 0 \\ \vdots & \frac{1}{x_1^2} & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & & \frac{1}{x_p^2} \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

$$= \left[\frac{x_1}{\|x_0^2\|} \quad \frac{x_2}{\|x_1\|} \quad \cdots \quad \frac{x_p}{\|x_p\|^2} \right]$$

$$\Rightarrow \hat{\beta}_j = \left[\frac{x_1}{\|x_0\|} \quad \frac{x_2}{\|x_1\|} \quad \cdots \quad \frac{x_p}{\|x_p\|^2} \right] \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_p \end{bmatrix} = \frac{x_0}{\|x_0\|^2} y_0 + \frac{x_1}{\|x_1\|^2} y_1 + \cdots \frac{x_p}{\|x_p\|^2} y_p$$

$$\Rightarrow \beta_j = \sum_{i=0}^{p} \frac{x_i}{\|x_i\|^2} \cdot y_i$$

3.

(a) As we know $X^T X\beta = X^T y$.

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

$\Rightarrow (U\Sigma V^T)^T \cdot U\Sigma V^T \beta = (U\Sigma V^T)^T y$.

$\Rightarrow V\Sigma^T \cdot U^T \cdot U \Sigma V^T \beta = V \cdot \Sigma^T U^T \cdot y$.

$\Rightarrow V^T \cdot V\Sigma^T U^T U \Sigma V^T \beta = V^T \cdot V \Sigma^T \cdot U^T \cdot y$.

$\because V^T V = I_r$, $U^T \cdot U = I_r$

$\Rightarrow \Sigma^T \cdot \Sigma V^T \beta = \Sigma^T \cdot U^T \cdot y$.

$(\Sigma^T)^{-1} \Sigma^T \Sigma V^T \beta = (\Sigma^T)^{-1} \Sigma^T \cdot U^T y$.

$\Rightarrow \Sigma^T V^T \beta = U^T y$.

$\Rightarrow \beta = (V^T)^{-1} (\Sigma^T)^{-1} U^T y$.

$\because V = (V^T)^{-1}$, $\Sigma^T = \Sigma$.

$\Rightarrow \beta = V \cdot \Sigma^{-1} U^T \cdot y$.

$\therefore \beta_{mns} = V \Sigma^{-1} U^T y$, proved.

(B) As we know:

$X^T X\beta = X^T y$, $\beta = V\Sigma^{-1} U^T + b$.

$\Rightarrow (U\Sigma V^T)^T (U\Sigma V^T)(V\Sigma^{-1} U^T y + b) = (U\Sigma V^T)^T y$.

$\Rightarrow V\Sigma^T U^T U\Sigma V^T V\Sigma^{-1} U^T y + V\Sigma^T U^T U\Sigma V^T b = V\Sigma^T U^T y$.

$\because \Sigma\cdot\Sigma^{-1} = I$, $U^T U = I_r$, $V^T V = I_r$.

$\Rightarrow V\Sigma^2 V^T b = V\Sigma^T U^T y - V\Sigma^T V^T y$

if $b=0$, there is only a solution to normal equation.

$\Rightarrow V\Sigma^T U^T y - V\Sigma^T U^T y = 0 \Rightarrow \|\beta\| = \|\beta_{mns}\|$

if $b \neq 0$, then not a solution to normal equation.

$\Rightarrow \|\beta\| > \|\beta_{mns}\|$

$\therefore \|\beta\| \geq \|\beta_{mns}\|$ proved.

C. Prove pseudo properties. for $X^+ = V\Sigma^{-1} U^T$

$X = U\Sigma V^T$

① $XX^+X = X$.

$XX^+X = (U\Sigma V^T)(V\Sigma^{-1} U^T)(U\Sigma V^T) = U\Sigma\Sigma^{-1}\Sigma V^T$

$= U\Sigma V^T$

$= X$ proved

② $X^+XX^+ = X^+$

$$X^+XX^+ = (V\Sigma^{-1}U^T)(U\Sigma V^T)(V\Sigma^{-1}U^T) = V\Sigma^{-1}\Sigma\Sigma^{-1}U^T$$

$$= V\Sigma^{-1}U^T.$$

$$= X^+$$

$\Rightarrow$ proved

③ $(XX^+)^T = X \cdot X^+$

$$(XX^+)^T = (U\Sigma V^T V\Sigma^{-1}U^T)^T$$

$$= (U\Sigma\Sigma^{-1}U^T)^T$$

$$= (U \cdot U^T)^T$$

$$= U \cdot U^T.$$

$$= I$$

$$X \cdot X^+ = (U\Sigma V^T)(V\Sigma^{-1}U^T) = U \cdot U^T = I.$$

$\Rightarrow (XX^+)^T = X \cdot X^+$ $\therefore$ proved.

④ $(X^+X)^T = X^+ \cdot X$.

$$\Rightarrow (X^+X)^T = (V\Sigma^{-1}U^T \cdot U\Sigma V^T)^T = V \cdot V^T = I.$$

$$X^+ \cdot X = V \cdot \Sigma^{-1}U^T \cdot U\Sigma V^T = V \cdot V^T = I$$

$\Rightarrow (X^+X)^T = X \cdot X^+ \Rightarrow$ proved.

$\therefore V\Sigma^{-1}U$ is the psedo inverse of X.