Problem 4.1. Since $\hat{\theta}^*$ is the BLUE of $\theta$, we know that $var(\hat{\theta}_i^* - \theta)$ is minimum among all linear unbiased estimator $\hat{\theta}_i$ of $\theta$, for every $i$. (this is how $\hat{\theta}^*$ was constructed).

So, if $\hat{\theta}$ is any linear unbiased estimator of $\theta$

$$var(\hat{\theta}_i^* - \theta_i) \leq var(\hat{\theta}_i - \theta_i) \quad \forall i$$

$$\Rightarrow \sum_{i=1}^{n} var(\hat{\theta}_i^* - \theta_i) \leq \sum_{i=1}^{n} var(\hat{\theta}_i - \theta_i)$$

$$\Leftrightarrow \sum_{i=1}^{n} E[(\hat{\theta}_i^* - \theta_i)^2] \leq \sum_{i=1}^{n} E[(\hat{\theta}_i - \theta_i)^2] \leftarrow \because \text{unbiased, } var(\tilde{\theta}_i) = E[\tilde{\theta}_i^2] \text{ where } \tilde{\theta}_i = \hat{\theta}_i - \theta$$

$$\Leftrightarrow E\left[\sum_i (\hat{\theta}_i^* - \theta_i)^2\right] \leq E\left[\sum_i (\hat{\theta}_i - \theta_i)^2\right] \leftarrow \text{linearity of Expectation}$$

$$\Leftrightarrow E[(\hat{\theta}^* - \theta)^T(\hat{\theta}^* - \theta)] \leq E[(\hat{\theta} - \theta)^T(\hat{\theta} - \theta)]$$

— Q.E.D.

Problem 2.1

$$Z = H\theta + \epsilon, \quad E[\epsilon] = 0, \quad E[\epsilon\epsilon^T] = R$$

in class we constructed the best linear unbiased estimate of $\theta$ in terms of $z$.

Now the question is, what is the best linear unbiased estimate of $L\theta$, where $L$ is some matrix

Since we want a <u>linear</u> estimate of $L\theta$,

call $\gamma \triangleq L\theta$,

then we want to find the 'gains' $g_i^*$

such that $\hat{\gamma}_i^* = g_i^{*T} z$ is unbiased and min var estimate of $\gamma_i$

$$\gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_p \end{bmatrix} = L\theta = \begin{bmatrix} \text{---} L_1^T \text{---} \\ \text{---} L_2^T \text{---} \\ \vdots \\ \text{---} L_p^T \text{---} \end{bmatrix} \theta = \begin{bmatrix} L_1^T\theta \\ L_2^T\theta \\ \vdots \\ L_p^T\theta \end{bmatrix}$$

(lets say $\theta \in \mathbb{R}^n$, $L \in \mathbb{R}^{p\times n}$, so $\gamma \in \mathbb{R}^p$)

An arbitrary linear estimate of $\gamma_i$:

$$\hat{\gamma}_i = g_i^T z$$

<u>unbiasedness</u> need $E[\hat{\gamma}_i] = \gamma_i$

$$E[\hat{\gamma}_i] = E[g_i^T(H\theta + \epsilon)] = g_i^T H\theta \cancel{+ g_i^T E(\epsilon)}^{0}$$

$$= g_i^T H . \theta$$

$\therefore$ we must have $g_i^T H\theta = \gamma_i = L_i^T\theta \quad \forall i$ (for every $i$)

$\therefore$ must have $g_i^T H = L_i^T$

$$\Leftrightarrow H^T g_i = L_i \quad \text{—} 1$$

variance: of $\hat{\gamma}_i = g_i^T z$ ?

$$var(\hat{\gamma}_i) = E[(\hat{\gamma}_i - \bar{\hat{\gamma}}_i)^2] = E[(\hat{\gamma}_i - \gamma_i)^2] \quad \because \text{unbiased}$$

$$= E[(g_i^T \epsilon)^2] \qquad (g_i^T\epsilon)(g_i\epsilon) = (g_i^T\epsilon)(g_i^T\epsilon)^T$$

$$= E[g_i^T \epsilon \epsilon^T g_i] \qquad = g_i^T \epsilon\epsilon^T g_i$$

$$= g_i^T R g_i$$

From this point on, you can repeat the steps carried out in class (to derive the BLUE for $\theta$), to show that

$$g_i^* = R^{-1} H (H^T R^{-1} H)^{-1} L_i.$$

$$\Rightarrow \hat{\gamma}_i^* = g_i^T \cdot z = L_i^T (H^T R^{-1} H)^{-1} H^T R^{-1} z$$

$$\Rightarrow \hat{\gamma}^* = \begin{bmatrix} \hat{\gamma}_1^* \\ \hat{\gamma}_2^* \\ \vdots \\ \hat{\gamma}_p^* \end{bmatrix} = \begin{bmatrix} L_1^T \\ L_2^T \\ \vdots \\ L_p^T \end{bmatrix} (H^T R^{-1} H)^{-1} H^T R^{-1} z = L (H^T R^{-1} H)^{-1} H^T R^{-1} z$$

$$= L \hat{\theta}^*$$

$$\therefore \widehat{(L\theta)}^* = L \hat{\theta}^*$$

— Q.E.D.

4.3 Decompose $P$ as $P = \underbrace{\frac{P+P^T}{2}}_{\text{symmetric}} + \underbrace{\frac{P-P^T}{2}}_{\text{skew-symmetric}}$

$:= P_s + P_k$

where $P_s = P_s^T > 0$, $P_k = -P_k^T$

For every skew symmetric matrix $A$, $x^T A x = 0$ (prove it :)

$\Rightarrow x^T P x = x^T P_s x + x^T P_k x = x^T P_s x$

$\Rightarrow$ We have to prove the result only for symmetric positive semi definite matrices, though the problem statement does not mention symmetric.

Since $P$ is p.s.d, $\exists L$ such that $P = L^T L$

$L\hat{\theta}^*$ is the BLUE of $L\theta$ (from part 4.2)

$$\Rightarrow E\left[(L\theta - L\hat{\theta}^*)^T (L\theta - L\hat{\theta}^*)\right] \leq E\left[(L\theta - L\hat{\theta})^T (L\theta - L\hat{\theta})\right]$$

for every unbiased linear estimate $\hat{\theta}$ of $\theta$ (from part 1)

$$\Leftrightarrow E\left[(\theta - \hat{\theta}^*)^T L^T L (\theta - \hat{\theta}^*)\right] \leq E\left[(\theta - \hat{\theta})^T L^T L (\theta - \hat{\theta})\right]$$

$$\Leftrightarrow E\left[(\theta - \hat{\theta}^*)^T P (\theta - \hat{\theta}^*)\right] \leq E\left[(\theta - \hat{\theta})^T P (\theta - \hat{\theta})\right]$$

— Q.E.D.

4.4. For an arbitrary real vector $x$ of appropriate dimension,

$$x^T E\left[(\theta-\hat{\theta}^*)(\theta-\hat{\theta}^*)^T\right]x = E\left[\underline{x^T\tilde{\theta}^*}\ \underline{\tilde{\theta}^{*T}x}\right]$$

where $\tilde{\theta}^* \triangleq \theta-\hat{\theta}^*$, and $x^T\tilde{\theta}^*$ is a scalar

$$= E\left[(x^T\tilde{\theta}^*)\ (x^T\tilde{\theta}^*)^T\right]$$

$$= E\left[(\tilde{\theta}^{*T}x)\ (x^T\tilde{\theta}^*)\right]$$

transpose of a scalar is itself, so $\tilde{x}^T\tilde{\theta}^* = (\tilde{x}^T\tilde{\theta}^*)^T = \tilde{\theta}^{*T}\tilde{x}$

etc.

$$= E\left[\tilde{\theta}^{*T}\ xx^T\ \tilde{\theta}^*\right]$$

$$= E\left[(\theta-\hat{\theta}^*)^T\ xx^T\ (\theta-\hat{\theta}^*)\right]$$

$$\le E\left[(\theta-\hat{\theta})^T\ xx^T\ (\theta-\hat{\theta})\right]$$ (by Problem 2.3,

since $xx^T$ is symmetric positive semi definite $\forall x$)

for every linear unbiased estimate $\hat{\theta}$ of $\theta$

$$= E\left[(\tilde{\theta}^T x)(x^T\tilde{\theta})\right] \qquad \tilde{\theta} \triangleq \theta-\hat{\theta}$$

$$= E\left[x^T\tilde{\theta}\cdot\tilde{\theta}^T x\right]$$

$$= x^T E\left[\tilde{\theta}\tilde{\theta}^T\right]x$$

$$= x^T E\left[(\theta-\hat{\theta})(\theta-\hat{\theta})^T\right]x$$

$$\Rightarrow x^T E\left[(\theta-\hat{\theta})^*(\theta-\hat{\theta}^*)\right]x \le x^T E\left[(\theta-\hat{\theta})(\theta-\hat{\theta})^T\right]x \quad \forall x$$

which proves the statement.

3. $$y_i - \bar{y} = (x_i - \bar{x})\, a + b + \epsilon$$
$$= \begin{bmatrix} (x_i - \bar{x}) & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + \epsilon$$

$$\underbrace{\begin{bmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ \vdots \\ y_m - y \end{bmatrix}}_{Z} = \underbrace{\begin{bmatrix} (x_1 - \bar{x}) & 1 \\ (x_2 - \bar{x}) & 1 \\ \vdots & \\ (x_m - \bar{x}) & 1 \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\theta} + \underbrace{\begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_m \end{bmatrix}}_{\epsilon}$$

$\hat{\theta}_{LS}$ is the solution of $(H^T H)\,\hat{\theta} = H^T Z$

$$H^T H = \begin{bmatrix} x_1 - \bar{x} & x_2 - \bar{x} & \cdots & x_m - \bar{x} \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 - \bar{x} & 1 \\ x_2 - \bar{x}_2 & 1 \\ \vdots & \vdots \\ x_m - \bar{x} & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{m} (x_i - \bar{x})^2 & \sum_{i=1}^{m} (x_i - \bar{x}) \\ \sum_{i=1}^{m} (x_i - \bar{x}) & m \end{bmatrix}$$

$$= \begin{bmatrix} \sum_i (x_i - \bar{x})^2 & 0 \\ 0 & m \end{bmatrix}$$

since $\Sigma(x_i - \bar{x}) = \Sigma x_i - m\bar{x}$
$= \Sigma x_i - \Sigma x_i$
$= 0$

$$\Rightarrow (H^T H)^{-1} = \begin{bmatrix} \frac{1}{\sum_{i=1}^{m} (x_i - \bar{x})^2} & 0 \\ 0 & \frac{1}{m} \end{bmatrix}$$

$$H^T Z = \begin{bmatrix} x_1 - \bar{x} & x_2 - \bar{x}_2 & \cdots & x_m - \bar{x}_m \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ \vdots \\ y_m - \bar{y} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{m} (x_i - \bar{x})(y_i - \bar{y}) \\ \sum_{i=1}^{m} (y_i - \bar{y}) \end{bmatrix}$$

$$= \begin{bmatrix} \sum_i (x_i - \bar{x})(y_i - y) \\ 0 \end{bmatrix}$$

$$\Rightarrow \hat{\theta}_{LS} = (H^T H)^{-1} H Z$$ (okay, I inverted the matrix !)

$$= \begin{bmatrix} \frac{1}{\sum_1^m (x_i - \bar{x})^2} & 0 \\ 0 & \frac{1}{m} \end{bmatrix} \begin{bmatrix} \sum_1^m (x_i - \bar{x})(y_i - \bar{y}) \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \hat{a}_{LS} \\ \hat{b}_{LS} \end{bmatrix} = \begin{bmatrix} \frac{\sum_1^m (x_i - \bar{x})(y_i - \bar{y})}{\sum_1^m (x_i - \bar{x})^2} \\ 0 \end{bmatrix}$$

— Q.E.D.