$\vec{y} = \mathbf{A}\cdot\vec{x} + b = A : \vec{y} = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \end{pmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix}$

$\hat{x}$ - estimat x - true

$\tilde{x}$ - linear estimator

$$E[\Theta] = (F^TF)^{-1}F^T E(y) = (F^TF)^{-1}F^T \cdot E(F\Theta + \epsilon)$$

① a) $y = F\Theta + \epsilon$ lsq: $\min_{\Theta} \epsilon^T\epsilon$ $\epsilon \sim \mathcal{N}(0, I)$

sol.: $\Theta = (F^TF)^{-1}F^Ty$

assuming Θ is deterministic, y is deterministic

$$E[\hat{\Theta}] = E\left[\Theta + (F^TF)^{-1}F^T\epsilon\right] = \Theta + (F^TF)^{-1}F^T \cdot E(\epsilon) = \underline{\underline{\Theta}}$$

$$\mathrm{cov}(\hat{\Theta}) = E\left[(\hat{\Theta} - \Theta)(\hat{\Theta} - \Theta)^T\right] = E\left[(\Theta + (F^TF)^{-1}F^T\epsilon - \Theta)(\Theta + (F^TF)^{-1}F^T\epsilon - \Theta)^T\right]$$

$$= E\left[(F^TF)^{-1}F^T\epsilon\epsilon^TF(F^TF)^{-1}\right] = (F^TF)^{-1}F^TE(\epsilon\epsilon^T)F(F^TF)^{-1}$$

$E[\epsilon\epsilon^T] \rightarrow$ since $\epsilon \sim \mathcal{N}(0, I)$, the off-diagonal elements will be 0, and the diagonal values will be 1.

$\Downarrow$

$$E[\epsilon\epsilon^T] = I$$

$$\mathrm{cov}(\hat{\Theta}) = (F^TF)^{-1}F^TIF(F^TF)^{-1} = \underline{\underline{(F^TF)^{-1}}}$$

b/ $(F^TF) = \begin{bmatrix} 6 & 3 \\ 3 & 2 \end{bmatrix}$ $(F^TF)^{-1} = \begin{bmatrix} \frac{2}{3} & -1 \\ -1 & 2 \end{bmatrix}$

$(F^TF)^{-1}F^T = \begin{bmatrix} -\frac{1}{3} & \frac{1}{3} & \frac{2}{3} \\ 1 & 0 & -1 \end{bmatrix}$ $\hat{\Theta} = (F^TF)^{-1}F^T y = \begin{bmatrix} \frac{1}{3} \\ 1 \end{bmatrix}$

$y - F\hat{\Theta} = \begin{bmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix}$ $\|y - F\Theta\|_2^2 = \frac{1}{\sqrt{3}}$

c/ $\hat{\Theta} = [\hat{\Theta}_a \ \hat{\Theta}_b \ \hat{\Theta}_M]^T$ $\hat{\Theta}_a \in \mathbb{R}$ $\hat{\Theta}_b \in \mathbb{R}$ $\hat{\Theta}_M \in \mathbb{R}^{n-2}$

split up $\hat{\Theta}$ just as F

$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1^T \\ 2 & 1 & 3 1^T \end{bmatrix} \cdot \hat{\Theta}$ $\hat{\Theta}_a + 2\hat{\Theta}_b = 3$ $\hat{\Theta}_a = 3 - 2\hat{\Theta}_b$

$\epsilon = y - F\hat{\Theta} = y - f_a\hat{\Theta}_a - f_b\hat{\Theta}_b - M\hat{\Theta}_M = y - M\hat{\Theta}_M - f_a(3 - 2\hat{\Theta}_b) -$

$- f_b\hat{\Theta}_b = y - 3f_a - (f_b - 2f_a)\hat{\Theta}_b - M\hat{\Theta}_M$

$2 = 3 - 2\hat{\Theta}_b + \hat{\Theta}_b + 1^T\hat{\Theta}_M$

$\hat{\Theta}_b = 1^T\hat{\Theta}_M + 1$

$\epsilon = (y - 3f_a) - (f_b - 2f_a) \cdot (1^T\hat{\Theta}_M + 1) - M\hat{\Theta}_M =$

$= \underbrace{(y - 3f_a - f_b + 2f_a)}_{y_{aug}} - \underbrace{(f_b 1^T - 2f_a 1^T + M)}_{F_{aug}}\hat{\Theta}_M$ → final least-squares problem

$\hat{\Theta}_M = (F_{aug}^T F_{aug})^{-1} F_{aug}^T y_{aug}$

(2) a)

$$p(\Theta) = \frac{1}{\sqrt{(2\pi)^n |P_\Theta|}} \cdot \exp\left(-\frac{1}{2}(\Theta-\mu_\Theta)^T P_\Theta^{-1}(\Theta-\mu_\Theta)\right) \qquad n = \dim(\Theta)$$

$$p(y|\Theta) = \frac{1}{\sqrt{(2\pi)^n |P_{y|\Theta}|}} \cdot \exp\left(-\frac{1}{2}(y-\mu_{y|\Theta})^T P_{y|\Theta}^{-1}(y-\mu_{y|\Theta})\right)$$

$n = \dim(y)$

$$p(\Theta|y) = \frac{1}{\sqrt{(2\pi)^n |P_{y|\Theta}|}} \exp\left(-\frac{1}{2}(\Theta-\mu_{\Theta|y})^T P_{\Theta|y}^{-1}(\Theta-\mu_{\Theta|y})\right)$$

$$P_y^{-1} = (LL^T)^{-1} = W = W^T \quad P_\Theta > 0 \rightarrow \quad P_{\hat{\Theta}} = (P_\Theta^{-1} + F^T W F)$$

$$\mu_{y|\Theta} = F\Theta \qquad \mu_{y|\Theta}^T = \Theta^T F^T$$

$$p(y|\Theta)\cdot p(\Theta) = \alpha \cdot \exp\left(-\frac{1}{2}(y^T - \Theta^T F^T) W (y - F\Theta) - \frac{1}{2}(\Theta - \mu_\Theta)^T P_\Theta^{-1} (\Theta - \mu_\Theta)\right) = \alpha \cdot \exp\left(-\frac{1}{2}y^T W y - \frac{1}{2}(\Theta^T F^T W F \Theta) + \right.$$

$$\frac{1}{2}\Theta^T F^T W y + \frac{1}{2} y^T W F \Theta - \frac{1}{2}\Theta^T P_\Theta^{-1}\Theta - \frac{1}{2}\mu_\Theta^T P_\Theta^{-1}\mu_\Theta + \frac{1}{2}\mu_\Theta^T P_\Theta^{-1}\Theta +$$

$$\left. + \frac{1}{2}\Theta^T P_\Theta^{-1}\mu_\Theta\right) = \underline{\alpha \cdot \exp\left(-\frac{1}{2}\Theta^T(F^T W F + P_\Theta^{-1})\Theta + \frac{1}{2}(\Theta^T F^T W y + \right.}$$

$$\underline{\left. y^T W F \Theta + \mu_\Theta^T P_\Theta^{-1}\Theta + \Theta^T P_\Theta^{-1}\mu_\Theta) - \frac{1}{2}y^T W y - \frac{1}{2}\mu_\Theta^T P_\Theta^{-1}\mu_\Theta\right)}$$

$\Downarrow$

2nd part

$$p(\Theta|y) = \beta \cdot \exp\left(\underline{-\frac{1}{2}\Theta^T P_{\Theta|y}^{-1}\Theta + \frac{1}{2}(\Theta^T P_{\Theta|y}^{-1}\mu_{\Theta|y} + \mu_{\Theta|y}^T P_{\Theta|y}^{-1}\Theta)} - \right.$$

$$\left.\underline{\frac{1}{2}(\mu_{\Theta|y} P_{\Theta|y}^{-1}\mu_{\Theta|y})}\right)$$

1st part

$\Theta^T P_{\Theta|y}^{-1}\mu_{\Theta|y}$ may not be symmetric

b) $E\left[(\Theta-\mu_\Theta)\epsilon^T\right]=0$ $\qquad y=F\Theta+L\epsilon$

unbiased : $E[\mu_{\Theta|y}]=\mu_\Theta$

minimum-variance : $E\left[(\Theta-\mu_{\Theta|y})(\Theta-\mu_{\Theta|y})^T\right]$ is minimized

$$\mu_{\Theta|y}=[M\ N]\begin{bmatrix}y\\ \mu_\Theta\end{bmatrix}$$

μ_Θ is deterministic

$\rightarrow E[\mu_{\Theta|y}]=E[My]+N\mu_\Theta=E[MF\Theta+M\epsilon]+N\mu_\Theta=MFE[\Theta]+ME[\epsilon]+$

$+N\mu_\Theta=MF\mu_\Theta+N\mu_\Theta=\mu_\Theta \Rightarrow MF+N=I \qquad E[\epsilon]=0$

$N=(I-MF)$

$\rightarrow (\Theta-\mu_{\Theta|y})=\Theta-MF\Theta-ML\epsilon-N\mu_\Theta=(I-MF)\Theta-(I-MF)\mu_\Theta-ML\epsilon=$

$=(I-MF)(\Theta-\mu_\Theta)-ML\epsilon$

$\rightarrow E\left[(\Theta-\mu_{\Theta|y})(\Theta-\mu_{\Theta|y})^T\right]=E\left[((I-MF)(\Theta-\mu_\Theta)-ML\epsilon)((I-MF)(\Theta-\mu_\Theta)-ML\epsilon)^T\right]=$

$=E\Big[(I-MF)(\Theta-\mu_\Theta)(\Theta-\mu_\Theta)^T(I-MF)^T-ML\epsilon(\Theta-\mu_\Theta)^T(I-MF)-$

$-(I-MF)(\Theta-\mu_\Theta)\epsilon^TL^TM^T+ML\epsilon\epsilon^TL^TM\Big]=(I-MF)\underbrace{E\left[(\Theta-\mu_\Theta)(\Theta-\mu_\Theta)^T\right]}_{P}$

$(I-MF)^T-ML\underbrace{E\left[\epsilon(\Theta-\mu_0)^T\right]}_{(E[(\Theta-\mu_0)\epsilon^T])^T=0^T=0}(I-MF)-(I-MF)\underbrace{E\left[(\Theta-\mu_\Theta)\epsilon^T\right]}_{0}+MLE\underbrace{[\epsilon\epsilon^T]}_{I}L^TM^T$

$(LL^T)=W^{-1}$

$\Rightarrow E\left[(\Theta-\mu_{\Theta|y})(\Theta-\mu_{\Theta|y})^T\right]=(I-MF)P(I-MF)^T+MW^{-1}M^T=$

$$=[I\ -M]\begin{bmatrix}P & PF^T\\ FP & FPF^T\end{bmatrix}\begin{bmatrix}I\\ -M^T\end{bmatrix}+MW^{-1}M^T=[I\ -M]\underbrace{\begin{bmatrix}P & PF^T\\ FP & FPF^T+W^{-1}\end{bmatrix}}_{Q}\begin{bmatrix}I\\ -M^T\end{bmatrix}$$

$L>0\Rightarrow W^{-1}>0 \qquad P>0\Rightarrow FPF^T>0\Rightarrow FPF^T+W^{-1}>0$

since the bottom right entry of Q is positive definite the schur-decomposition can be applied.

$$\rightarrow Q = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & BD^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A-BD^{-1}C & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} I & 0 \\ D^{-1}C & I \end{bmatrix}$$

$$\rightarrow Q = \begin{bmatrix} I & PF^T(FPF^T+W^{-1})^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} P - PF^T(FPF^T+W^{-1})^{-1}FP & 0 \\ 0 & FPF^T+W^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ (FPF^T+W^{-1})^{-1}FP & I \end{bmatrix}$$

$\rightarrow$ substituting it back into $E\left((\Theta-\mu_\Theta)(\Theta-\mu_\Theta)^T\right)$

$$\rightarrow E\left[(\Theta-\mu_{\Theta|y})(\Theta-\mu_{\Theta|y})^T\right] = \begin{bmatrix} I & -M \end{bmatrix} \begin{bmatrix} I & PF \\ & \end{bmatrix} =$$

$$= \begin{bmatrix} I & PF^T(FPF^T+W^{-1})^{-1}-M \end{bmatrix} \begin{bmatrix} P-PF^T(FPF^T+W^{-1})^{-1}FP & 0 \\ 0 & (FPF^T+W^{-1}) \end{bmatrix} \begin{bmatrix} I \\ (FPF^T+W^{-1})^{-1}FP-M^T \end{bmatrix} =$$

$$= \begin{bmatrix} P-PF^T(FPF^T+W^{-1})^{-1}FP & PF^T-M(FPF^T+W^{-1}) \end{bmatrix}$$

$$\begin{bmatrix} I \\ (FPF^T+W^{-1})^{-1}FP-M^T \end{bmatrix} = \left(P-PF^T(FPF^T+W^{-1})^{-1}FP\right) +$$

$$+\left[\left(PF^T(FPF^T+W^{-1})^{-1}-M\right)(FPF^T+W^{-1})\right]\cdot\left((FPF^T+W^{-1})^{-1}FP-M^T\right)$$

since the goal is to minimize the above expression one can set either $\left(PF^T(FPF^T+W^{-1})^{-1}-M\right)$ or its transpose to 0

$$\text{So} \Rightarrow M = PF^T(FPF^T + W^{-1})^{-1}$$

$$\rightarrow \mu_{\Theta|y} = PF^T(FPF^T + W^{-1})^{-1} y + \mu_\Theta - PF^T(FPF^T + W^{-1})^{-1} F \mu_\Theta =$$

$$= \underline{\underline{PF^T(FPF^T + W^{-1})^{-1}(y - F\mu_\Theta) + \mu_\Theta}}$$

If $(PF^T(FPF^T + W^{-1})^{-1} - M)$ is set to 0 the remaining term in the expression for $E\left[(\Theta - \mu_{\Theta|y})(\Theta - \mu_{\Theta|y})^T\right] =$

$$= \left(P - PF^T(FPF^T + W^{-1})^{-1} FP\right) = P_{\Theta|y}$$

c) since $P > 0$, and the Matrix inversion lemma

$\exists P^{-1}$

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)DA^{-1}$$

$\rightarrow (FPF^T + W^{-1})^{-1} = W - WF(P^{-1} + F^TWF)^{-1}F^TW$, $\mu_{\Theta|y}$ becomes

$$PF^T(W - WF(P^{-1} + F^TWF)^{-1}F^TW) = P(F^TW - F^TWF(P^{-1} + F^TWF)^{-1}F^TW) =$$

$$= P(I - F^TWF(P^{-1} + F^TWF)^{-1})F^TW = P((P^{-1} + F^TWF)\cdot(P^{-1} + F^TWF)^{-1} - F^TWF(P^{-1} + F^TWF))$$

$$F^TW = \cancel{P}((\cancel{P^{-1}} + \cancel{F^TWF} - \cancel{F^TWF})(P^{-1} + F^TWF))F^TW = \underline{\underline{(P^{-1} + F^TWF)F^TW}}$$

P