ROB 501 HW8

Yulun Zhuang
yulunz@umich.edu

November 8, 2022

1

Recall the formulas of Best Linear Unbiased Estimate (BLUE).

Given $y \in \mathbb{R}^m$, $x \in \mathbb{R}^n$, $y = Cx + \epsilon$, $E(\epsilon) = 0$, $Var(\epsilon) = Q > 0$, and $rank(C) = n$.

$$\begin{aligned}
\hat{K} &= (C^T Q^{-1} C)^{-1} C^T Q^{-1} \\
\hat{x} &= \hat{K} y \\
Var(\hat{x}) = E\{(\hat{x} - x)(\hat{x} - x)^T\} &= (C^T Q^{-1} C)^{-1}
\end{aligned}$$

1.1

$$\begin{aligned}
\hat{x} &= \begin{bmatrix} 0.6194 \\ 0.4591 \end{bmatrix} \\
Var(\hat{x}) &= \begin{bmatrix} 4.0000 & -2.7500 \\ -2.7500 & 2.0000 \end{bmatrix}
\end{aligned}$$

1.2

$$\begin{aligned}
\hat{x} &= \begin{bmatrix} -1.4303 \\ 1.8791 \end{bmatrix} \\
Var(\hat{x}) &= \begin{bmatrix} 0.0679 & -0.0260 \\ -0.0260 & 0.1129 \end{bmatrix}
\end{aligned}$$

1.3

$$\begin{aligned}
\hat{x} &= \begin{bmatrix} -1.2201 \\ 1.5368 \end{bmatrix} \\
Var(\hat{x}) &= \begin{bmatrix} 0.0487 & 0.0054 \\ 0.0054 & 0.0618 \end{bmatrix}
\end{aligned}$$

2

Given jointly Guassian Random Variables (X, Y, Z) with mean and covariance

$$\mu = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \ \Sigma = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

2.1

Conditional distribution of $\begin{bmatrix} x \\ y \end{bmatrix} \mid \{Z=z\}$.

Let $P = \begin{bmatrix} X \\ Y \end{bmatrix}, Q = z$.

$$\Sigma_{PQ} = \mathrm{Cov}(P,Q) = \mathrm{Cov}\left(\begin{bmatrix} X \\ Y \end{bmatrix}, Z\right) = \begin{bmatrix} \mathrm{Cov}(X,Z) \\ \mathrm{Cov}(Y,Z) \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \Sigma_{QP}^T$$

$$\begin{aligned}
\mu_{\{P|Q=z\}} &= \mu_P + \Sigma_{PQ}\Sigma_Q^{-1}(z-\mu_Q) \\
&= \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix}\frac{1}{2}(z-1) \\
&= \begin{bmatrix} \frac{z-3}{2} \\ z-1 \end{bmatrix} \\
\Sigma_{\{P|Q=z\}} &= \Sigma_P - \Sigma_{PQ}\Sigma_Q^{-1}\Sigma_{QP} \\
&= \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \end{bmatrix}\frac{1}{2}\begin{bmatrix} 1 & 2 \end{bmatrix} \\
&= \begin{bmatrix} \frac{3}{2} & 1 \\ 1 & 2 \end{bmatrix}
\end{aligned}$$

2.2

Distribution of $X|_{\{Z=z\}}$ conditioned on $Y|_{\{Z=z\}} = y$.

Let $P = X|_{\{Z=z\}}, Q = Y|_{\{Z=z\}}$. Compute the conditional distribution of P given $Q = y$.

From 2.1, $\mu_p = \frac{z-3}{2},\ \mu_Q = z-1$ and $\Sigma_{PQ} = 1,\ \Sigma_P = \frac{3}{2},\ \Sigma_Q = 2$.

$$\begin{aligned}
\mu_{\{P|Q=y\}} &= \mu_P + \Sigma_{PQ}\Sigma_Q^{-1}(y-\mu_Q) \\
&= \frac{z-3}{2} + \frac{1}{2}(y-z+1) \\
&= \frac{1}{2}y - 1 \\
\Sigma_{\{P|Q=y\}} &= \Sigma_P - \Sigma_{PQ}\Sigma_Q^{-1}\Sigma_{QP} \\
&= \frac{3}{2} - 1 \times \frac{1}{2} \times 1 \\
&= 1
\end{aligned}$$

2.3

Distribution of $X \mid_{\{Y=y, Z=z\}}$.

Let $P = X$, $Q = \begin{bmatrix} Y \\ Z \end{bmatrix}$ and $q = \begin{bmatrix} y \\ z \end{bmatrix}$.

$$
\Sigma_{PQ} = \mathrm{Cov}\left(X, \begin{bmatrix} Y \\ Z \end{bmatrix}\right) = [\mathrm{Cov}(X,Y), \mathrm{Cov}(X,Z)] = \begin{bmatrix} 2 & 1 \end{bmatrix}
$$

$$
\begin{aligned}
\mu_{\{P|Q=q\}} &= \mu_P + \Sigma_{PQ}\Sigma_Q^{-1}\left(\begin{bmatrix} y \\ z \end{bmatrix} - \mu_Q\right) \\
&= -1 + \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}^{-1} \left(\begin{bmatrix} y \\ z \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \\
&= \frac{1}{2}y - 1 \\
\Sigma_{\{P|Q=q\}} &= \Sigma_P - \Sigma_{PQ}\Sigma_Q^{-1}\Sigma_{QP} \\
&= 2 - \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\
&= 1
\end{aligned}
$$

2.4

They are the same. If we have jointly distributed normal random vectors, when we condition one block of vectors on another, we always obtain either a jointly distributed normal random vector or, if only a scalar quantity is left, a normally distributed random variable.

3

3.1

Let G_k be the Gram matrix for M_k and G_{k+1} be the Gram matrix for M_{k+1}. Using $y_{k+1} \perp M_k$, we have

$$
G_{k+1} = \begin{bmatrix} G_k & 0_{k\times 1} \\ 0_{1\times k} & < y_{k+1}, y_{k+1} > \end{bmatrix}
$$

Apply the Normal equations $G^T\alpha = \beta$.

$$
\begin{aligned}
\alpha &= G_{k+1}^{-T}\beta \\
&= \begin{bmatrix} G_k^{-1} & 0_{k\times 1} \\ 0_{1\times k} & \langle y_{k+1}, y_{k+1}\rangle \end{bmatrix} \begin{bmatrix} \langle x, y_1\rangle \\ \vdots \\ \langle x, y_k\rangle \\ \langle x, y_{k+1}\rangle \end{bmatrix} \\
&= \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_k \\ \dfrac{\langle x, y_{k+1}\rangle}{\langle y_{k+1}, y_{k+1}\rangle} \end{bmatrix} \\
\hat{x}_{k+1} &= \sum_{i=1}^{k+1} \alpha_i x_i = \hat{x}_k + \frac{\langle x, y_{k+1}\rangle}{\langle y_{k+1}, y_{k+1}\rangle} y_{k+1} \\
\Rightarrow \beta &= \frac{\langle x, y_{k+1}\rangle}{\langle y_{k+1}, y_{k+1}\rangle}
\end{aligned}
$$

3.2

From the Projection Theorem, $y_{k+1} - \hat{y}_{k+1|k}$ is orthogonal to M_k, such that $\forall v \in M_k$

$$M_{k+1} = M_k \oplus span\{y_{k+1}\},\ M_k \perp (y_{k+1} - v)$$

Since $\hat{y}_{k+1|k} \in M_k$, we have

$$M_{k+1} = M_k \oplus span\{y_{k+1} - \hat{y}_{k+1|k}\},\ M_k \perp (y_{k+1} - \hat{y}_{k+1|k})$$

Apply the result in 3.1,

$$\begin{aligned}
\hat{x}_{k+1} &= \hat{x}_k + \frac{\left\langle x, y_{k+1} - \hat{y}_{k+1|k} \right\rangle}{\left\langle y_{k+1} - \hat{y}_{k+1|k}, y_{k+1} - \hat{y}_{k+1|k} \right\rangle} \left(y_{k+1} - \hat{y}_{k+1|k} \right) \\
\Rightarrow \beta &= \frac{\left\langle x, y_{k+1} - \hat{y}_{k+1|k} \right\rangle}{\left\langle y_{k+1} - \hat{y}_{k+1|k}, y_{k+1} - \hat{y}_{k+1|k} \right\rangle}
\end{aligned}$$

4

Recall the formulas of Minimum Variance Estimate (MVE).

Given $y \in \mathbb{R}^m$, $x \in \mathbb{R}^n$, $y = Cx + \epsilon$, $E(X) = 0$, $Var(X) = P > 0$, $E(\epsilon) = 0$, $Var(\epsilon) = Q > 0$, and $rank(C) = n$.

$$\begin{aligned}
\hat{K} &= [C^T Q^{-1} C + P^{-1}]^{-1} C^T Q^{-1} \\
\hat{x} &= \hat{K} y \\
Var(\hat{x}) &= E\{(\hat{x} - x)(\hat{x} - x)^T\} = P - PC^T (CPC^T + Q)^{-1} CP
\end{aligned}$$

4.1

$$\begin{aligned}
\hat{x} &= \begin{bmatrix} 0.3417 \\ 0.4271 \end{bmatrix} \\
Var(\hat{x}) &= \begin{bmatrix} 0.2778 & -0.0278 \\ -0.0278 & 0.1528 \end{bmatrix}
\end{aligned}$$

4.2

$$\begin{aligned}
\hat{x} &= \begin{bmatrix} 0.4504 \\ 0.4963 \end{bmatrix} \\
Var(\hat{x}) &= \begin{bmatrix} 0.1938 & -0.0812 \\ -0.0812 & 0.1188 \end{bmatrix}
\end{aligned}$$

4.3

$$\begin{aligned}
\hat{x} &= \begin{bmatrix} -1.0134 \\ 1.2402 \end{bmatrix} \\
Var(\hat{x}) &= \begin{bmatrix} 0.0545 & -0.0105 \\ -0.0105 & 0.0828 \end{bmatrix}
\end{aligned}$$

4.4

$$\begin{aligned}
\hat{x} &= \begin{bmatrix} -1.0296 \\ 1.2667 \end{bmatrix} \\
Var(\hat{x}) &= \begin{bmatrix} 0.0437 & 0.0072 \\ 0.0072 & 0.0538 \end{bmatrix}
\end{aligned}$$

5

5.1

Standard least squares approximation:

$$\hat{x} = \begin{bmatrix} -1.3169 \\ 1.4368 \end{bmatrix}$$

5.2

BLUE estimation:

$$\hat{x} = \begin{bmatrix} -1.3169 \\ 1.4368 \end{bmatrix}$$

5.3

MVE estimation of x when $P = 100I$:

$$\hat{x} = \begin{bmatrix} -1.3163 \\ 1.4365 \end{bmatrix}$$

MVE estimation of x when $P = 10^6 I$:

$$\hat{x} = \begin{bmatrix} -1.3169 \\ 1.4368 \end{bmatrix}$$

5.4

1. BLUE equals to standard least squares when the inverse of the covariance of the noise is identity.
2. MVE reduces to BLUE when the covariance of the x approaches infinity.

6

Recall

$$\widehat{x} = \bar{x} + PC^\top \left(CPC^\top + Q\right)^{-1} (y - \bar{y})$$

and

$$E\left\{(x - \hat{x})(x - \hat{x})^\top\right\} = P - PC^\top \left(CPC^\top + Q\right)^{-1} CP$$

where $\bar{x} = E\{x\}, \bar{\epsilon} = E\{\epsilon\}$ and $\bar{y} = C\bar{x} + \bar{\epsilon}$.

Given

$$\bar{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ and } \bar{\epsilon} = 0,$$

We have

$$\hat{x} = \begin{bmatrix} -0.8836 \\ 1.0802 \end{bmatrix}$$