ME 233 Advance Control II

Lecture 5
Least Squares Estimation

(ME233 Class Notes pp. LS1-LS5)

Notation

Let X and Y be continuous random vectors with joint PDF $p_{XY}(x,y)$

Let x and y be respectively outcomes of X and Y and

$$x \in R_x \subseteq R^{n_x} \qquad y \in R_y \subseteq R^{n_y}$$

$$p_{XY} : R_x \times R_y \to R_+$$

Marginal Expectation (review)

Let X and Y be continuous random vectors with joint PDF $p_{XY}(x,y)$

Marginal Expectation (mean) of X

$$m_X = E\{X\}$$

$$= \int_{R_x} \underbrace{\int_{R_y} x\, p_{XY}(x,y)\,dy}_{x p_X(x)}\,dx$$

Marginal Expectation (review)

Let X and Y be continuous random variables with joint PDF $p_{XY}(x,y)$

Marginal Expectation (mean) of X

$$m_X = E\{X\} = \int_{R_x} x\, p_X(x)\,dx$$

$$= \hat{x}$$

new notation
(following the ME233 class notes)

5

Marginal Expectation $\widehat{x}$

$\widehat{x}$ is the minimum least squares **marginal estimator** of X, i.e.

- For any deterministic vector z

$$E\{\|X-\widehat{x}\|^2\} \leq E\{\|X-z\|^2\}$$

Euclidean norm

6

Marginal Expectation $\widehat{x}$

$$E\{\|X-\widehat{x}\|^2\} \leq E\{\|X-z\|^2\}$$

Proof:

$$E\{\|X-z\|^2\} = E\{\|(X-\widehat{x})-(z-\widehat{x})\|^2\}$$

$$= E\{\|X-\widehat{x}\|^2 + \|z-\widehat{x}\|^2 - 2(z-\widehat{x})^T(X-\widehat{x})\}$$

$$= E\{\|X-\widehat{x}\|^2\} + \|z-\widehat{x}\|^2 - 2(z-\widehat{x})^T \cancelto{0}{E\{X-\widehat{x}\}}$$

$$\geq E\{\|X-\widehat{x}\|^2\} \qquad \blacksquare$$

7

Conditional Expectation (review)

Let X and Y be continuous random vectors with joint PDF $p_{XY}(x,y)$

Conditional Expectation (conditional mean) of X given and outcome $Y = y$

$$m_{X|y} = E\{X|Y=y\}$$

$$= \int_{R_x} x\, p_{X|y}(x)dx$$

8

Conditional Expectation (review)

Conditional Expectation (conditional mean) of X given and outcome $Y = y$

$$m_{X|y} = \int_{R_x} x\, p_{X|y}(x)dx$$

$$= \int_{R_x} x \left(\frac{p_{XY}(x,y)}{p_Y(y)}\right) dx$$

$$= \widehat{x}|y$$

new notation (following the ME233 class notes)

Conditional Expectation $\widehat{x}|_y$

Notice that the conditional expectation $\widehat{x}|_y$

$$\widehat{x}|_y \;=\; \int_{R_x} x\,\frac{p_{XY}(x,y)}{p_Y(y)}dx$$

can be interpreted as a function of <u>the random variable Y</u>.

$$\widehat{X}|_Y \;=\; \int_{R_x} x\,\frac{p_{XY}(x,Y)}{p_Y(Y)}dx$$

Conditional Expectation $\widehat{X}|_Y$

Lemma:

For any function $f(\cdot)$ of the random vector Y, with the appropriate dimensions

$$E\{f(Y)\,X\} = E\{f(Y)\widehat{X}|_Y\}$$

we can replace X by its conditional expectation $\widehat{X}|_Y$

Marginal Expectation $\widehat{x}$

$$E\{f(Y)\,X\} = E\{f(Y)\widehat{X}|_Y\}$$

Proof:

First examine the left-hand side:

$$E\{f(Y)X\} = \int_{R_y}\int_{R_x} f(y)x\,\underbrace{p_{XY}(x,y)}\,dx\,dy$$

$$= \int_{R_y}\int_{R_x} f(y)x\,\overbrace{p_{X|y}(x)p_Y(y)}\,dx\,dy$$

$$= \int_{R_y} f(y)\left[\int_{R_x} x\,p_{X|y}(x)dx\right]p_Y(y)dy$$

Marginal Expectation $\widehat{x}$

$$E\{f(Y)\,X\} = E\{f(Y)\widehat{X}|_Y\}$$

Proof:

First examine the left-hand side:

$$E\{f(Y)X\} = \int_{R_y} f(y)\underbrace{\left[\int_{R_x} x\,p_{X|y}(x)dx\right]}_{\hat{x}|_y}p_Y(y)dy$$

$$E\{f(Y)X\} = \int_{R_y} f(y)\widehat{x}|_y\,p_Y(y)dy$$

13

Marginal Expectation $\widehat{x}$

$$E\{f(Y)\,X\} = E\{f(Y)\widehat{X}|_Y\}$$

Proof:

Now examine the right-hand side:

$$E\{f(Y)\widehat{X}|_Y\} = \int_{R_y}\int_{R_x} \underbrace{f(y)\widehat{x}|_y}_{\text{Not a function of } x} p_{XY}(x,y)dx\,dy$$

$$E\{f(Y)\widehat{X}|_Y\} = \int_{R_y} f(y)\widehat{x}|_y \underbrace{\left[\int_{R_x} p_{XY}(x,y)dx\right]}_{p_Y(y)} dy$$

14

Marginal Expectation $\widehat{x}$

$$E\{f(Y)\,X\} = E\{f(Y)\widehat{X}|_Y\}$$

Proof:

Therefore,

$$E\{f(Y)X\} = \int_{R_y} f(y)\widehat{x}|_y\, p_Y(y)dy$$
$$= E\{f(Y)\widehat{X}_Y\}$$

■

15

Conditional Expectation $\widehat{X}|_Y$

Theorem:

$\widehat{X}|_Y$ is the least squares minimum estimator of ***X*** given Y, *i.e.*

$$E\{\|X-\widehat{X}|_Y\|^2\} \le E\{\|X-f(Y)\|^2\}$$

for all functions $f(\cdot)$ of Y of appropriate dimensions

$$\|X\|^2 = X^T X$$

16

Marginal Expectation $\widehat{x}$

$$E\{\|X-\widehat{X}|_Y\|^2\} \le E\{\|X-f(Y)\|^2\}$$

Proof:

$$E\{\|X-f(Y)\|^2\} = E\{\|(X-\widehat{X}|_Y)-(f(Y)-\widehat{X}|_Y)\|^2\}$$

$$= E\left\{\left\|X-\widehat{X}|_Y\right\|^2 + \left\|f(Y)-\widehat{X}|_Y\right\|^2 \right.$$
$$\left. -2(f(Y)-\widehat{X}|_Y)^T(X-\widehat{X}|_Y)\right\}$$

$$= E\left\{\left\|X-\widehat{X}|_Y\right\|^2\right\} + E\left\{\left\|f(Y)-\widehat{X}|_Y\right\|^2\right\}$$
$$-2E\{(f(Y)-\widehat{X}|_Y)^T X\} + 2E\{(f(Y)-\widehat{X}|_Y)^T\widehat{X}|_Y\}$$

17

Marginal Expectation $\widehat{x}$

$$E\{\|X - \widehat{X}|_Y\|^2\} \le E\{\|X - f(Y)\|^2\}$$

Proof:

Define $g(Y) := (f(Y) - \hat{X}|_Y)^T$

$$E\left\{\|X - f(Y)\|^2\right\} = E\left\{\left\|X - \hat{X}|_Y\right\|^2\right\} + E\left\{\left\|f(Y) - \hat{X}|_Y\right\|^2\right\} \cancelto{0}{- 2E\{g(Y)X\} + 2E\{g(Y)\hat{X}|_Y\}}$$

Since $\|f(Y) - \hat{X}|_Y\|^2 \ge 0$ for all outcomes,

$$E\{\|f(Y) - \widehat{X}|_Y\|^2\} \ge 0$$

$$\Rightarrow \quad E\{\|X - f(Y)\|^2\} \ge E\{\|X - \hat{X}|_Y\|^2\} \quad \blacksquare$$

18

Conditional Expectation for Gaussians (review)

When $\begin{bmatrix} X \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} m_X \\ m_Y \end{bmatrix}, \begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix}\right)$

$$X|_y \sim N(\widehat{x}_y,\ \Lambda_{X|_yX|_y})$$

where

$$\widehat{x}|_y = \widehat{x} + \Lambda_{XY}\Lambda_{YY}^{-1}(y - \widehat{y})$$
$$\Lambda_{X|_yX|_y} = \Lambda_{XX} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YX}$$

19

Conditional Mean for Gaussians

When $\begin{bmatrix} X \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} m_X \\ m_Y \end{bmatrix}, \begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix}\right)$

$$\widehat{X}|_Y = \widehat{x} + \Lambda_{XY}\Lambda_{YY}^{-1}(Y - \widehat{y})$$

$$E\{\widehat{X}|_Y\} = \widehat{x} + \Lambda_{XY}\Lambda_{YY}^{-1}\cancelto{0}{E\{Y - \widehat{y}\}}$$
$$= \widehat{x}$$

20

Conditional Mean for Gaussians

When $\begin{bmatrix} X \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} m_X \\ m_Y \end{bmatrix}, \begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix}\right)$

$$\tilde{X}|_y = X - (\widehat{x} + \Lambda_{XY}\Lambda_{YY}^{-1}(y - \widehat{y}))$$
$$= \tilde{X} - \Lambda_{XY}\Lambda_{YY}^{-1}(y - \widehat{y})$$

$$\tilde{X}|_Y = \tilde{X} - \Lambda_{XY}\Lambda_{YY}^{-1}(Y - \widehat{y})$$
$$= \tilde{X} - \Lambda_{XY}\Lambda_{YY}^{-1}\tilde{Y}$$

21

Conditional Mean for Gaussians

When $\begin{bmatrix} X \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} m_X \\ m_Y \end{bmatrix}, \begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix} \right)$

$$\tilde{X}|_Y = \tilde{X} - \Lambda_{XY}\Lambda_{YY}^{-1}\tilde{Y}$$

$$\begin{aligned} E\{\tilde{X}|_Y\} &= \cancelto{0}{E\{\tilde{X}\}} - \Lambda_{XY}\Lambda_{YY}^{-1}\cancelto{0}{E\{\tilde{Y}\}} \\ &= 0 \end{aligned}$$

22

Least Squares Estimation: Property 1

- The conditional estimation error $\tilde{X}_{|Y}$ and Y are ***uncorrelated***

$$E\{\tilde{X}_{|Y}\tilde{Y}^T\} = 0$$

- $\tilde{X}_{|Y}$ and $\hat{X}_{|Y}$ are ***orthogonal***

$$E\{\tilde{X}_{|Y}\hat{X}_{|Y}^T\} = 0 \quad \text{and} \quad E\{\tilde{X}_{|Y}^T\hat{X}_{|Y}\} = 0$$

23

Least Squares Estimation: Property 1

$$E\{\tilde{X}_{|Y}\tilde{Y}^T\} = 0$$

Proof

$$\begin{aligned} E\{\tilde{X}_{|Y}\tilde{Y}^T\} &= E\{(\tilde{X} - \Lambda_{XY}\Lambda_{YY}^{-1}\tilde{Y})\tilde{Y}^T\} \\ &= E\{\tilde{X}\tilde{Y}^T\} - \Lambda_{XY}\Lambda_{YY}^{-1}E\{\tilde{Y}\tilde{Y}^T\} \\ &= \Lambda_{XY} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YY} \\ &= 0 \end{aligned}$$

■

24

Least Squares Estimation: Property 1

$$E\{\tilde{X}_{|Y}\hat{X}_{|Y}^T\} = 0$$

Proof

$$\begin{aligned} E\{\tilde{X}_{|Y}\hat{X}_{|Y}^T\} &= E\{\tilde{X}_{|Y}(\hat{x} + \Lambda_{XY}\Lambda_{YY}^{-1}\tilde{Y})^T\} \\ &= \cancelto{0}{E\{\tilde{X}_{|Y}\}}\hat{x}^T + \cancelto{0}{E\{\tilde{X}_{|Y}\tilde{Y}^T\}}\Lambda_{YY}^{-1}\Lambda_{XY}^T \\ &= 0 \end{aligned}$$

■

Least Squares Estimation: Property 1

$$E\{\tilde{X}_{|Y}^T \hat{X}_{|Y}\} = 0$$

Proof

$$\tilde{X}_{|Y}^T \hat{X}_{|Y} = (\tilde{X}_{|Y}^T \hat{X}_{|Y})^T = \hat{X}_{|Y}^T \tilde{X}_{|Y}$$

$$= \text{trace}(\hat{X}_{|Y}^T \tilde{X}_{|Y}) = \text{trace}(\tilde{X}_{|Y} \hat{X}_{|Y}^T)$$

scalar

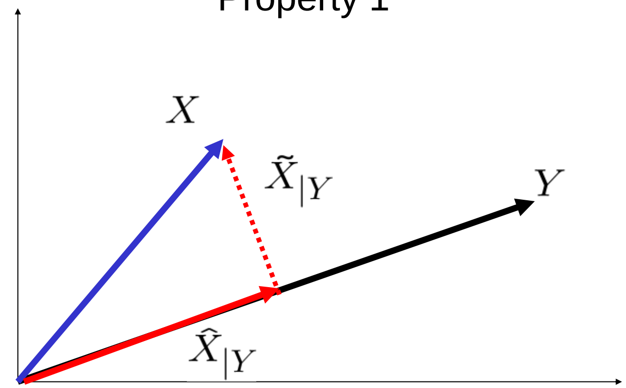
$$\Rightarrow \quad E\{\tilde{X}_{|Y}^T \hat{X}_{|Y}\} = E\{\text{trace}(\tilde{X}_{|Y} \hat{X}_{|Y}^T)\}$$

$$= \text{trace}(E\{\tilde{X}_{|Y} \hat{X}_{|Y}^T\})$$

$$= \text{trace}(0) = 0 \quad \blacksquare$$

Why does trace commute with expectation?

Deterministic interpretation of Property 1



Recursive LS Estimation

Let X, Y and Z be jointly Gaussian R.V.s

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \sim N\left(\begin{bmatrix} m_X \\ m_Y \\ m_Z \end{bmatrix}, \begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} & \Lambda_{XZ} \\ \Lambda_{YX} & \Lambda_{YY} & \Lambda_{YZ} \\ \Lambda_{ZX} & \Lambda_{ZY} & \Lambda_{ZZ} \end{bmatrix} \right)$$

$X \in \mathcal{R}^n$ $\}\ n$

$Y \in \mathcal{R}^M$ $\}\ M >> n,\ p$

$Z \in \mathcal{R}^p$ $\}\ p$

Recursive LS Estimation

1. Assume that we already know of outcome $Y = y$

and we have obtained

$$\hat{x}_{|y} = E\{X|Y = y\}$$

$$\hat{x}_{|y} = \hat{x} + \Lambda_{XY} \underbrace{\Lambda_{YY}^{-1}}_{\text{inverse of an } M \times M \text{ matrix}} (y - \hat{y})$$

n M

Recursive LS Estimation

1. Assume that we already know of outcome $Y = y$

and we have obtained $\hat{x}_{|y} = \hat{x} + \Lambda_{XY} \Lambda_{YY}^{-1} (y - \hat{y})$

2. Now we also know the outcome $Z = z$

How do we efficiently compute

$$\hat{x}_{|yz} = E\{X|Y = y,\, Z = z\} \quad ?$$

Non-Recursive LS Estimation

1) Define the vector $W = \begin{bmatrix} Z \\ Y \end{bmatrix} \qquad \hat{w} = \begin{bmatrix} \hat{z} \\ \hat{y} \end{bmatrix}$

2) Compute $\hat{x}_{|w} = E\{X|Y = y,\, Z = z\}$

$$\hat{x}_{|w} = \hat{x} + \Lambda_{XW} \underbrace{\Lambda_{WW}^{-1}}_{\textit{inverse of an } (p+M) \times (p+M) \textit{ matrix}} (w - \hat{w})$$

n

$p + M$

Least Squares Estimation: Property 2

Assume that $\Lambda_{ZY} = E\{\tilde{Z}\tilde{Y}^T\} = 0$

Then,

$$\hat{X}_{|YZ} = \hat{X}_{|Y} + \left(\tilde{X}_{|Y}\right)_{|Z}$$

$$\Lambda_{\tilde{X}_{|YZ}\tilde{X}_{|YZ}} = \Lambda_{\tilde{X}_{|Y}\tilde{X}_{|Y}} - \Lambda_{XZ}\Lambda_{ZZ}^{-1}\Lambda_{ZX}$$

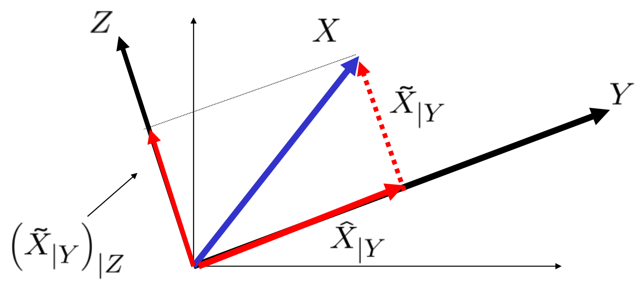
where

$$\hat{X}_{|Y} = \hat{x} + \Lambda_{XY}\Lambda_{YY}^{-1}(Y - \hat{y})$$

$$\left(\tilde{X}_{|Y}\right)_{|Z} = \Lambda_{XZ}\Lambda_{ZZ}^{-1}(Z - \hat{z})$$

$$\Lambda_{\tilde{X}_{|Y}\tilde{X}_{|Y}} = \Lambda_{XX} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YX}$$

Deterministic interpretation of Property 2



$$\hat{X}_{|YZ} = \hat{X}_{|Y} + \left(\tilde{X}_{|Y}\right)_{|Z}$$

Least Squares Estimation: Property 2

$$\left(\tilde{X}_{|Y}\right)_{|Z} = \Lambda_{XZ}\Lambda_{ZZ}^{-1}(Z-\hat{z})$$

Proof:

$$\left(\tilde{X}_{|Y}\right)_{|Z} = \cancelto{0}{E\{\tilde{X}_{|Y}\}} + \Lambda_{\tilde{X}_{|Y}Z}\Lambda_{ZZ}^{-1}(Z-\hat{z})$$

$$\Lambda_{\tilde{X}_{|Y}Z} = E\{\tilde{X}_{|Y}\tilde{Z}^T\} = E\left\{\left[\tilde{X} - \Lambda_{XY}\Lambda_{YY}^{-1}\tilde{Y}\right]\tilde{Z}^T\right\}$$

$$= \underbrace{E\{\tilde{X}\tilde{Z}^T\}}_{\Lambda_{XZ}} - \Lambda_{XY}\Lambda_{YY}^{-1}\cancelto{0}{E\{\tilde{Y}\tilde{Z}^T\}} \qquad \blacksquare$$

because Z and Y are uncorrelated

Least Squares Estimation: Property 2

$$\hat{X}_{|YZ} = \hat{X}_{|Y} + \left(\tilde{X}_{|Y}\right)_{|Z}$$

Proof:

$$\hat{X}_{|YZ} = \hat{x} + \underbrace{\Lambda_{XW}}_{\left[\Lambda_{XZ} \quad \Lambda_{XY}\right]}\underbrace{\Lambda_{WW}^{-1}}_{\begin{bmatrix}\Lambda_{ZZ}^{-1} & 0 \\ 0 & \Lambda_{YY}^{-1}\end{bmatrix}}\underbrace{(W-\hat{w})}_{\begin{bmatrix}\tilde{Z} \\ \tilde{Y}\end{bmatrix}}$$

$$\hat{X}_{|YZ} = \underbrace{\hat{x} + \Lambda_{XY}\Lambda_{YY}^{-1}\tilde{Y}}_{\hat{X}_{|Y}} + \underbrace{\Lambda_{XZ}\Lambda_{ZZ}^{-1}\tilde{Z}}_{\left(\tilde{X}_{|Y}\right)_{|Z}} \qquad \blacksquare$$

Least Squares Estimation: Property 2

$$\Lambda_{\tilde{X}_{|YZ}\tilde{X}_{|YZ}} = \Lambda_{\tilde{X}_{|Y}\tilde{X}_{|Y}} - \Lambda_{XZ}\Lambda_{ZZ}^{-1}\Lambda_{ZX}$$

Proof:

$$\Lambda_{\tilde{X}_{|YZ}\tilde{X}_{|YZ}} = \Lambda_{XX} - \underbrace{\Lambda_{XW}}_{\left[\Lambda_{XZ} \quad \Lambda_{XY}\right]}\underbrace{\Lambda_{WW}^{-1}}_{\begin{bmatrix}\Lambda_{ZZ}^{-1} & 0 \\ 0 & \Lambda_{YY}^{-1}\end{bmatrix}}\underbrace{\Lambda_{WX}}_{\begin{bmatrix}\Lambda_{ZX} \\ \Lambda_{YX}\end{bmatrix}}$$

$$\Lambda_{\tilde{X}_{|YZ}\tilde{X}_{|YZ}} = \underbrace{\Lambda_{XX} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YX}}_{\Lambda_{\tilde{X}_{|Y}\tilde{X}_{|Y}}} - \Lambda_{XZ}\Lambda_{ZZ}^{-1}\Lambda_{ZX} \qquad \blacksquare$$

Least Squares Estimation : Property 3

What happens when Z and Y are **correlated**?

$$\Lambda_{ZY} = E\{\tilde{Z}\tilde{Y}^T\} \neq 0$$

Then,

$$\boxed{\hat{X}_{|YZ} = \hat{X}_{|Y} + \underbrace{\left(\tilde{X}_{|Y}\right)_{|(\tilde{Z}_{|Y})}}}$$

This warrants further explanation…

Recursive LS Estimation

Using Y, we can estimate X and Z by their conditional means:

The conditional mean of X

$$\hat{X}_{|_Y} = \hat{x} + \Lambda_{XY}\Lambda_{YY}^{-1}(Y - \hat{y})$$

The conditional mean of Z

$$\hat{Z}_{|_Y} = \hat{z} + \Lambda_{ZY}\Lambda_{YY}^{-1}(Y - \hat{y})$$

The corresponding conditional estimation errors are:

$$\tilde{X}_{|_Y} = X - \hat{X}_{|_Y} \qquad \tilde{Z}_{|_Y} = Z - \hat{Z}_{|_Y}$$

<u>Uncorrelated</u> with Y (by Least Squares Property 1)

Recursive LS Estimation

We have:

The conditional mean of X

$$\hat{X}_{|_Y} = \hat{x} + \Lambda_{XY}\Lambda_{YY}^{-1}(Y - \hat{y})$$

The conditional mean of Z

$$\hat{Z}_{|_Y} = \hat{z} + \Lambda_{ZY}\Lambda_{YY}^{-1}(Y - \hat{y})$$

If we get the outcomes $Y=y$ and $Z=z$

The corresponding conditional estimation errors become:

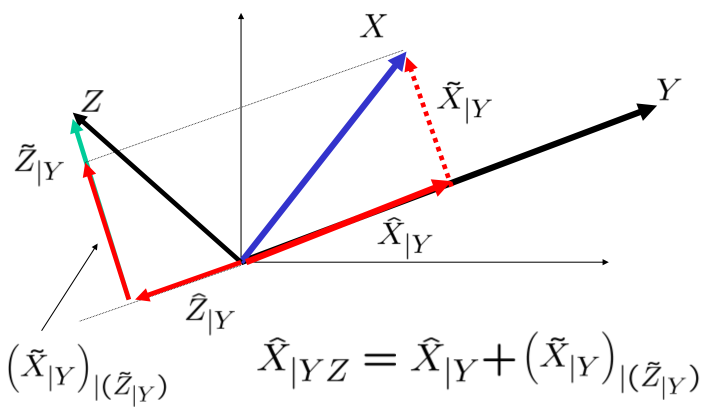
$$\tilde{X}_{|_y} = X - \hat{x}_{|_y}$$

This is still random

$$\boxed{\tilde{z}_{|_y} = z - \hat{z}_{|_y}}$$

This is now an outcome

Deterministic interpretation of Property 3



$$\hat{X}_{|YZ} = \hat{X}_{|Y} + \left(\tilde{X}_{|Y}\right)_{|(\tilde{Z}_{|Y})}$$

Least Squares Estimation : Property 3

a) Recursive estimate

$$\boxed{\hat{X}_{|YZ} = \hat{X}_{|Y} + \left(\tilde{X}_{|Y}\right)_{|(\tilde{Z}_{|Y})}}$$

where:

$$\hat{X}_{|Y} = \hat{x} + \Lambda_{XY}\Lambda_{YY}^{-1}(Y - \hat{y})$$
$$\hat{Z}_{|Y} = \hat{z} + \Lambda_{ZY}\Lambda_{YY}^{-1}(Y - \hat{y})$$
$$\left(\tilde{X}_{|Y}\right)_{|(\tilde{Z}_{|Y})} = \Lambda_{\tilde{X}_{|Y}\tilde{Z}_{|Y}}\Lambda^{-1}_{\tilde{Z}_{|Y}\tilde{Z}_{|Y}}(Z - \hat{Z}_{|Y})$$

$$\Lambda_{\tilde{X}_{|Y}\tilde{Z}_{|Y}} = \Lambda_{XZ} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YZ} \qquad \Lambda^{-1}_{\tilde{Z}_{|Y}\tilde{Z}_{|Y}} = \left[\Lambda_{ZZ} - \Lambda_{ZY}\Lambda_{YY}^{-1}\Lambda_{YZ}\right]^{-1}$$

42

Least Squares Estimation : Property 3

b) Recursive estimation error

$$\boxed{\Lambda_{\tilde{X}_{|YZ}\tilde{X}_{|YZ}} = \Lambda_{\tilde{X}_{|Y}\tilde{X}_{|Y}} - \Lambda_{\tilde{X}_{|Y}\tilde{Z}_{|Y}}\Lambda^{-1}_{\tilde{Z}_{|Y}\tilde{Z}_{|Y}}\Lambda_{\tilde{Z}_{|Y}\tilde{X}_{|Y}}}$$

where:

$$\Lambda_{\tilde{X}_{|Y}\tilde{X}_{|Y}} = \Lambda_{XX} - \Lambda_{XY}\Lambda^{-1}_{YY}\Lambda_{YX}$$

$$\Lambda_{\tilde{X}_{|Y}\tilde{Z}_{|Y}} = \Lambda_{XZ} - \Lambda_{XY}\Lambda^{-1}_{YY}\Lambda_{YZ}$$

$$\Lambda_{\tilde{Z}_{|Y}\tilde{Z}_{|Y}} = \Lambda_{ZZ} - \Lambda_{ZY}\Lambda^{-1}_{YY}\Lambda_{YZ}$$

43

Derivation of Recursive LS Estimation

1) Define the vector $W = \begin{bmatrix} Z \\ Y \end{bmatrix}$ $\quad \widehat{w} = \begin{bmatrix} \widehat{z} \\ \widehat{y} \end{bmatrix}$

2) Compute $\widehat{x}_{|yz} = E\{X|Y = y,\, Z = z\}$

$$\widehat{x}_{|yz} = \widehat{x} + \Lambda_{XW}\,\underbrace{\Lambda^{-1}_{WW}}_{}\,(w - \widehat{w})$$

inversion of an (p+M) × (p+M) matrix

44

Solution: use Schur complement

- Given

$$\Lambda_{WW} = \begin{bmatrix} \Lambda_{ZZ} & \Lambda_{ZY} \\ \Lambda_{YZ} & \Lambda_{YY} \end{bmatrix} \qquad \text{and} \qquad \Lambda^{-1}_{YY}$$

- Compute the Schur complement of Λ_{YY}

$$\Delta = \Lambda_{ZZ} - \Lambda_{ZY}\Lambda^{-1}_{YY}\Lambda_{YZ}$$
$$= \Lambda_{\tilde{Z}_{|Y}\tilde{Z}_{|Y}} := \Lambda_{Z|Y}$$

which is the conditional covariance

45

Solution: use Schur complement of Λ_{YY}

- Given

$$\Lambda_{WW} = \begin{bmatrix} \Lambda_{ZZ} & \Lambda_{ZY} \\ \Lambda_{YZ} & \Lambda_{YY} \end{bmatrix} \qquad \Lambda_{Z|Y} = \Lambda_{ZZ} - \Lambda_{ZY}\Lambda^{-1}_{YY}\Lambda_{YZ}$$

- Then

$$\Lambda^{-1}_{WW} = \begin{bmatrix} \Lambda^{-1}_{Z|Y} & -\Lambda^{-1}_{Z|Y}F \\ -F^T\Lambda^{-1}_{Z|Y} & \Lambda^{-1}_{YY} + F^T\Lambda^{-1}_{Z|Y}F \end{bmatrix}$$

$$F = \Lambda_{ZY}\Lambda^{-1}_{YY}$$

46

Non-Recursive LS Estimation

$$\widehat{x}_{|yz} = \widehat{x} + \underbrace{\Lambda_{XW}}_{\left[\begin{array}{cc}\Lambda_{XZ} & \Lambda_{XY}\end{array}\right]} \underbrace{\Lambda_{WW}^{-1}}_{\left[\begin{array}{cc}\Lambda_{ZZ} & \Lambda_{ZY} \\ \Lambda_{YZ} & \Lambda_{YY}\end{array}\right]^{-1}} \underbrace{(w - \widehat{w})}_{\tilde{w} = \left[\begin{array}{c}\tilde{z} \\ \tilde{y}\end{array}\right]}$$

$$W = \left[\begin{array}{c} Z \\ Y \end{array}\right]$$

47

Use Schur complement

$$\widehat{x}_{|yz} = \widehat{x}$$

$$+ \left[\begin{array}{cc}\Lambda_{XZ} & \Lambda_{XY}\end{array}\right] \underbrace{\left[\begin{array}{cc}\Lambda_{ZZ} & \Lambda_{ZY} \\ \Lambda_{YZ} & \Lambda_{YY}\end{array}\right]^{-1}}_{\left[\begin{array}{cc}\Lambda_{Z|Y}^{-1} & -\Lambda_{Z|Y}^{-1}F \\ -F^T\Lambda_{Z|Y}^{-1} & \Lambda_{YY}^{-1} + F^T\Lambda_{Z|Y}^{-1}F\end{array}\right]} \left[\begin{array}{c}\tilde{z} \\ \tilde{y}\end{array}\right]$$

$$\Lambda_{Z|Y} = \Lambda_{ZZ} - \Lambda_{ZY}\Lambda_{YY}^{-1}\Lambda_{YZ} \qquad F = \Lambda_{ZY}\Lambda_{YY}^{-1}$$

48

Use Schur complement

$$\widehat{x}_{|yz} = \widehat{x} + \Lambda_{XY}\Lambda_{YY}^{-1}\tilde{y}$$

$$+ \left(\Lambda_{XZ} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YZ}\right)\Lambda_{Z|Y}^{-1}\left(\tilde{z} - \Lambda_{ZY}\Lambda_{YY}^{-1}\tilde{y}\right)$$

$$\Lambda_{Z|Y} = \Lambda_{ZZ} - \Lambda_{ZY}\Lambda_{YY}^{-1}\Lambda_{YZ}$$

49

Use Schur complement

$$\widehat{x}_{|yz} = \underbrace{\widehat{x} + \Lambda_{XY}\Lambda_{YY}^{-1}\tilde{y}}_{\widehat{x}_{|y}}$$

$\widehat{x}_{|y} \longleftarrow$ *expected value of X given outcome y*

$$+ \left(\Lambda_{XZ} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YZ}\right)\Lambda_{Z|Y}^{-1}\left(\tilde{z} - \Lambda_{ZY}\Lambda_{YY}^{-1}\tilde{y}\right)$$

50

Use Schur complement

We will now show that

$$\hat{x}_{|yz} = \hat{x}_{|y} + \underbrace{(\Lambda_{XZ} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YZ})\,\Lambda_{Z|Y}^{-1}\,(\tilde{z} - \Lambda_{ZY}\Lambda_{YY}^{-1}\tilde{y})}_{E\{\tilde{X}_{|Y}|\tilde{z}_{|y}\}}$$

The expected value of $\tilde{X}_{|y}$ given the outcome $\tilde{z}_{|y}$

51

Computation of $\tilde{z}_{|y}$

The conditional mean of Z given $Y = y$:

$$\hat{z}_{|y} = \hat{z} + \Lambda_{ZY}\,\Lambda_{YY}^{-1}\,\tilde{y}$$

$$\tilde{z}_{|y} = z - \hat{z}_{|y}$$

$$\tilde{z}_{|y} = \underbrace{z - \hat{z}}_{\tilde{z}} - \Lambda_{ZY}\Lambda_{YY}^{-1}\tilde{y}$$

52

Computation of $E\{\tilde{X}_{|Y}|\tilde{z}_{|y}\}$

Therefore, $\tilde{z}_{|y} = \tilde{z} + \Lambda_{ZY}\,\Lambda_{YY}^{-1}\,\tilde{y}$

We will now compute $E\{\tilde{X}_{|Y}|\tilde{z}_{|y}\}$ using the LS result:

$$E\{\tilde{X}_{|Y}|\tilde{z}_{|y}\} = E\{\tilde{X}_{|Y}\} + E\{\tilde{X}_{|Y}\tilde{Z}_{|Y}^{T}\}E\{\tilde{Z}_{|Y}\tilde{Z}_{|Y}^{T}\}^{-1}\,\tilde{z}_{|y}$$

to verify that

$$E\{\tilde{X}_{|Y}|\tilde{z}_{|y}\} = (\Lambda_{XZ} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YZ})\,\Lambda_{Z|Y}^{-1}\,\underbrace{(\tilde{z} - \Lambda_{ZY}\Lambda_{YY}^{-1}\tilde{y})}_{\tilde{z}_{|y}}$$

53

Computation of $E\{\tilde{X}_{|Y}|\tilde{z}_{|y}\}$

Using Gaussian least squares results:

$$E\{\tilde{X}_{|Y}|\tilde{z}_{|y}\} = \underbrace{\cancel{E\{\tilde{X}_{|Y}\}}}_{\boldsymbol{0}} + E\{\tilde{X}_{|Y}\tilde{Z}_{|Y}^{T}\}\,E\{\tilde{Z}_{|Y}\tilde{Z}_{|Y}^{T}\}^{-1}\,\tilde{z}_{|y}$$

Estimation errors always have zero means

54

Computation of $E\{\tilde{X}_{|Y} | \tilde{z}_{|y}\}$

Using Gaussian least squares results:

$$E\{\tilde{X}_{|Y} | \tilde{z}_{|y}\} = \; E\{\tilde{X}_{|Y} \tilde{Z}_{|Y}^T\} \underbrace{E\{\tilde{Z}_{|Y} \tilde{Z}_{|Y}^T\}}^{-1} \tilde{z}_{|y}$$

$$E\{\tilde{Z}_{|Y} \tilde{Z}_{|Y}^T\} = \; \Lambda_{\tilde{Z}_{|Y} \tilde{Z}_{|Y}} \; = \; \Lambda_{Z|Y}$$

$$= \Lambda_{ZZ} - \Lambda_{ZY} \Lambda_{YY}^{-1} \Lambda_{YZ}$$

the conditional covariance

55

Computation of $E\{\tilde{X}_{|Y} | \tilde{z}_{|y}\}$

Using Gaussian least squares results:

$$E\{\tilde{X}_{|Y} | \tilde{z}_{|y}\} = \; E\{\tilde{X}_{|Y} \tilde{Z}_{|Y}^T\} \; \Lambda_{Z|Y}^{-1} \; \tilde{z}_{|y}$$

Notice that, from the Schur complements result,

$$E\{\tilde{X}_{|Y} | \tilde{z}_{|y}\} \; = (\Lambda_{XZ} - \Lambda_{XY} \Lambda_{YY}^{-1} \Lambda_{YZ}) \, \Lambda_{Z|Y}^{-1} \, \tilde{z}_{|y}$$

56

Computation of $E\{\tilde{X}_{|Y} | \tilde{z}_{|y}\}$

Using Gaussian least squares results:

$$E\{\tilde{X}_{|Y} | \tilde{z}_{|y}\} = \; \underbrace{E\{\tilde{X}_{|Y} \tilde{Z}_{|Y}^T\}} \; \Lambda_{Z|Y}^{-1} \; \tilde{z}_{|y}$$

$$E\{(\tilde{X} - \Lambda_{XY} \Lambda_{YY}^{-1} \tilde{Y}) \tilde{Z}_{|Y}^T\}$$

$$E\{\tilde{X} \tilde{Z}_{|Y}^T\} \; + \; \Lambda_{XY} \Lambda_{YY}^{-1} E\{\cancel{\tilde{Y} \tilde{Z}_{|Y}^T}\}$$

0

57

Computation of $E\{\tilde{X}_{|Y} | \tilde{z}_{|y}\}$

Using Gaussian least squares results:

$$E\{\tilde{X}_{|Y} | \tilde{z}_{|y}\} \; = \; \underbrace{E\{\tilde{X} \tilde{Z}_{|Y}^T\}} \Lambda_{Z|Y}^{-1} \; \tilde{z}_{|y}$$

$$E\{\tilde{X} \tilde{Z}_{|Y}^T\} \; = \; E\{\tilde{X} (\tilde{Z} - \Lambda_{ZY} \Lambda_{YY}^{-1} \tilde{Y})^T\}$$

$$= E\{\tilde{X} \tilde{Z}^T\} - \; E\{\tilde{X} \tilde{Y}^T\} \Lambda_{YY}^{-1} \Lambda_{YZ}$$

$$= \Lambda_{XZ} - \; \Lambda_{XY} \Lambda_{YY}^{-1} \Lambda_{YZ}$$

Computation of $E\{\tilde{X}_{|Y}|\tilde{z}_{|y}\}$

Therefore,

$$E\{\tilde{X}_{|Y}\tilde{Z}_{|Y}^T\} = \Lambda_{XZ} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YZ}$$

and

$$E\{\tilde{X}_{|Y}|\tilde{z}_{|y}\} = (\Lambda_{XZ} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YZ})\,\Lambda_{Z|Y}^{-1}\,\tilde{z}_{|y}$$

Non-Recursive LS Estimation Error

$$\Lambda_{\tilde{X}_{|W}\tilde{X}_{|W}} = \Lambda_{XX} - \underbrace{\Lambda_{XW}}_{\begin{bmatrix}\Lambda_{XZ} & \Lambda_{XY}\end{bmatrix}}\underbrace{\Lambda_{WW}^{-1}}_{\begin{bmatrix}\Lambda_{ZZ} & \Lambda_{ZY}\\ \Lambda_{YZ} & \Lambda_{YY}\end{bmatrix}^{-1}}\Lambda_{WX}$$

$$W = \begin{bmatrix} Z \\ Y \end{bmatrix}$$

Use Schur complement

$$\Lambda_{\tilde{X}_{|YZ}\tilde{X}_{|YZ}} = \Lambda_{XX} - \begin{bmatrix}\Lambda_{XZ} & \Lambda_{XY}\end{bmatrix}\underbrace{\begin{bmatrix}\Lambda_{ZZ} & \Lambda_{ZY}\\ \Lambda_{YZ} & \Lambda_{YY}\end{bmatrix}^{-1}}_{\begin{bmatrix}\Lambda_{Z|Y}^{-1} & -\Lambda_{Z|Y}^{-1}F\\ -F^T\Lambda_{Z|Y}^{-1} & \Lambda_{YY}^{-1}+F^T\Lambda_{Z|Y}^{-1}F\end{bmatrix}}\begin{bmatrix}\Lambda_{ZX}\\ \Lambda_{ZY}\end{bmatrix}$$

$$\Lambda_{Z|Y} = \Lambda_{ZZ} - \Lambda_{ZY}\Lambda_{YY}^{-1}\Lambda_{YZ} \qquad F = \Lambda_{ZY}\Lambda_{YY}^{-1}$$

Use Schur complement

$$\Lambda_{\tilde{X}_{|YZ}\tilde{X}_{|YZ}} = \Lambda_{XX} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YX}$$

$$-(\Lambda_{XZ} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YZ})\Lambda_{Z|Y}^{-1}(\Lambda_{ZX} - \Lambda_{ZY}\Lambda_{YY}^{-1}\Lambda_{YX})$$

$$\Lambda_{Z|Y} = \Lambda_{ZZ} - \Lambda_{ZY}\Lambda_{YY}^{-1}\Lambda_{YZ}$$

■

62

Summary

- The conditional mean is the least squares estimator:

$$E\{\|X - \widehat{X}|_Y\|^2\} \leq E\{\|X - f(Y)\|^2\}$$

- For Gaussians, the conditional mean is an affine function

$$\widehat{x}|_y = \widehat{x} + \Lambda_{XY}\,\Lambda_{YY}^{-1}\,(y - \widehat{y})$$

63

Summary

The conditional mean can be computed recursively:

1. If we first know of outcome $Y = y$

$$\widehat{x}_{|y} = \widehat{x} + \Lambda_{XY}\,\Lambda_{YY}^{-1}\,\tilde{y}$$

64

Summary

The conditional mean can be computed recursively:

2 If we afterwards know of outcome $Z = z$

$$\widehat{z}_{|y} = \widehat{z} + \Lambda_{ZY}\,\Lambda_{YY}^{-1}\,\tilde{y}$$
$$\tilde{z}_{|y} = z - \widehat{z}_{|y}$$

then

$$\boxed{\widehat{x}_{|yz} = \quad \widehat{x}_{|y} \quad + \quad E\{\tilde{X}_{|Y}|\tilde{z}_{|y}\}}$$

65

Course Outline

- Unit 0: Probability

Finished

- Unit 1: State-space control,estimation
- Unit 2: Input/output control
- Unit 3: Adaptive control