ME 233 Advanced Control II

Lecture 6
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

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

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

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 $p_{XY}(x,y)$

***<u>Marginal Expectation</u>* (mean)** of X

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

$$= \hat{x}$$

new notation
(following the ME233 class notes)

Marginal Expectation $\widehat{x}$

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

- For any deterministic vector $\boldsymbol{z}$

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

Euclidean norm

Marginal Expectation $\widehat{x}$

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

Proof:

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

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

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

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

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 $Y=y$

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

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

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 $\widehat{x}\big|_y$

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

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

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 $\hat{X}|_Y$

Lemma:

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

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

we can replace $\boldsymbol{X}$ *by its conditional expectation* $\hat{X}|_Y$

Marginal Expectation $\widehat{x}$

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

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:

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

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

Marginal Expectation $\widehat{x}$

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

Proof:

Now examine the right-hand side:

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

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

Marginal Expectation $\widehat{x}$

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

Proof:

Therefore,

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

■

Conditional Expectation $\hat{X}|_Y$

Theorem:

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

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

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

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

Marginal Expectation $\widehat{x}$

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

Proof:

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

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

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

Marginal Expectation $\widehat{x}$

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

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

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

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

Conditional Expectation for Gaussians (review)

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

$$X|_y \sim N(\hat{x}_y, \ \Lambda_{X|yX|y})$$

where

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

Conditional Mean for Gaussians

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

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

$$\begin{aligned} E\{\hat{X}|_Y\} &= \hat{x} + \Lambda_{XY}\Lambda_{YY}^{-1}\cancelto{\mathbf{0}}{E\{Y - \hat{y}\}} \\ &= \hat{x} \end{aligned}$$

Conditional Mean for Gaussians

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

$$
\begin{aligned}
\tilde{X}|_y &= X - (\hat{x} + \Lambda_{XY}\Lambda_{YY}^{-1}(y - \hat{y})) \\
&= \tilde{X} - \Lambda_{XY}\Lambda_{YY}^{-1}(y - \hat{y})
\end{aligned}
$$

$$
\begin{aligned}
\tilde{X}|_Y &= \tilde{X} - \Lambda_{XY}\Lambda_{YY}^{-1}(Y - \hat{y}) \\
&= \tilde{X} - \Lambda_{XY}\Lambda_{YY}^{-1}\tilde{Y}
\end{aligned}
$$

Conditional Mean for Gaussians

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

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

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

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 $\widehat{X}_{|Y}$ are ***orthogonal***

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

Least Squares Estimation: Property 1

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

Proof

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

$\blacksquare$

Least Squares Estimation: Property 1

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

Proof

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

$\blacksquare$

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

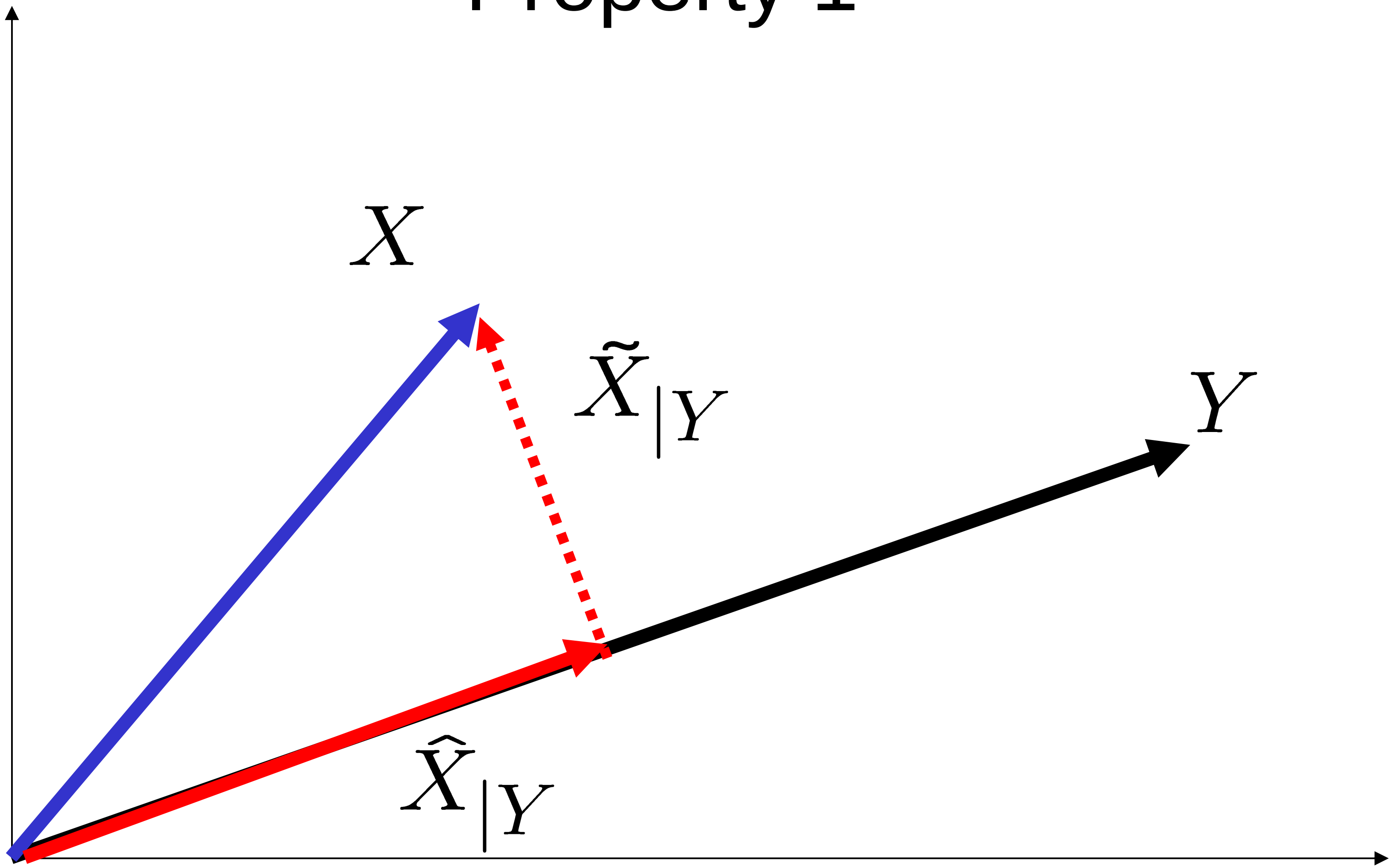
$$\Rightarrow \quad E\{\tilde{X}_{|Y}^T \hat{X}_{|Y}\} = E\{\text{trace}(\tilde{X}_{|Y} \hat{X}_{|Y}^T)\}$$

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

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

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

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

$$X \in \mathcal{R}^n \qquad \Big\} \; n$$

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

$$Z \in \mathcal{R}^p \qquad \Big\} \; p$$

Recursive LS Estimation

1. Assume that we already know of outcome $\boldsymbol{Y} = \boldsymbol{y}$

and we have obtained

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

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

n

M

Recursive LS Estimation

1. Assume that we already know of outcome $\boldsymbol{Y} = \boldsymbol{y}$

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

2. Now we also know the outcome $\boldsymbol{Z} = \boldsymbol{z}$

How do we efficiently compute

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

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,

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

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

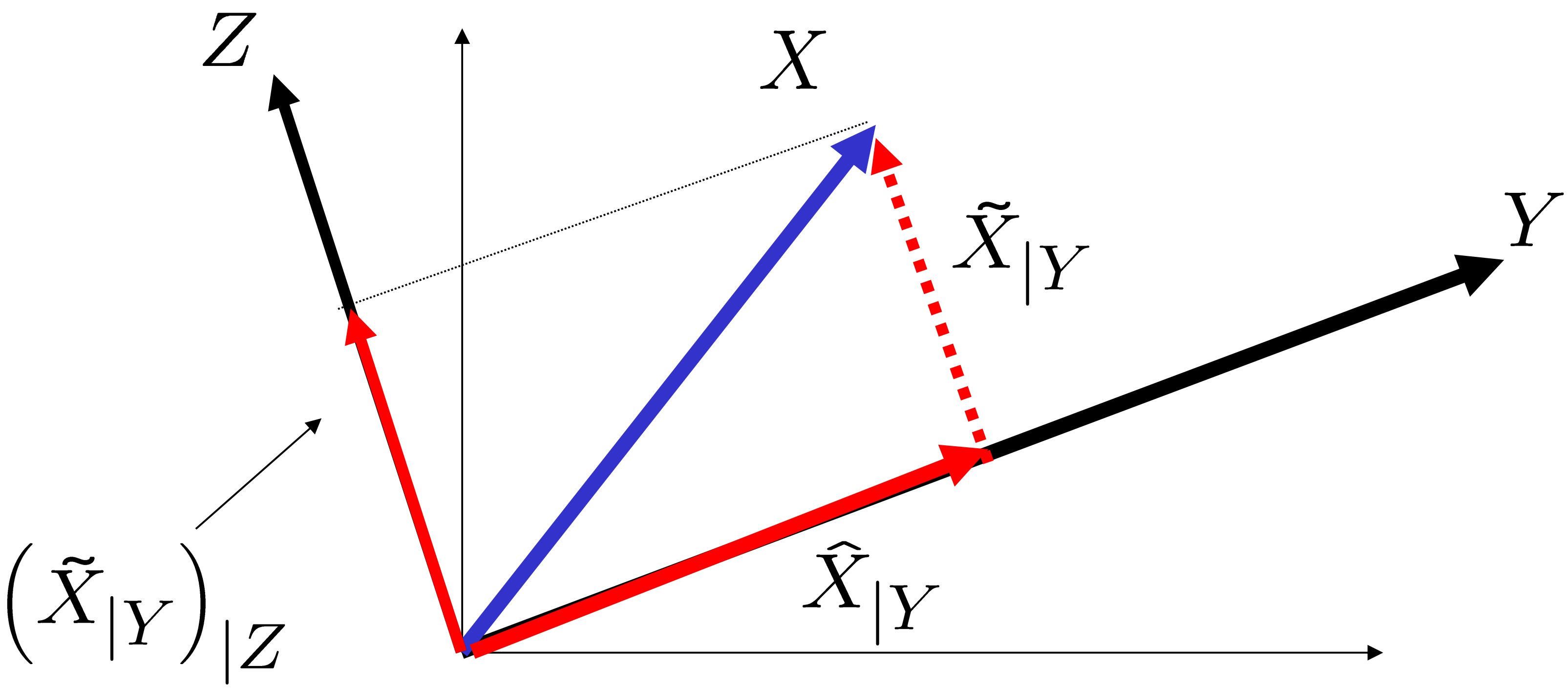
where

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

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

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

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:

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

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

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

because $\boldsymbol{Z}$ and $\boldsymbol{Y}$ are uncorrelated

Least Squares Estimation: Property 2

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

Proof:

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

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

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:

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

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

$\blacksquare$

Least Squares Estimation : Property 3

What happens when $\boldsymbol{Z}$ and $\boldsymbol{Y}$ are **correlated**?

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

Then,

$$\boxed{\hat{X}_{|YZ} = \hat{X}_{|Y} + \underbrace{\left(\tilde{X}_{|Y}\right)_{|(\tilde{Z}_{|Y})}}_{\textit{This warrants further explanation...}}}$$

Recursive LS Estimation

Using $\boldsymbol{Y}$, 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:

$$\tilde{X}_{|Y} = X - \hat{X}_{|Y}$$

$$\tilde{Z}_{|Y} = Z - \hat{Z}_{|Y}$$

<u>Uncorrelated</u> with $\boldsymbol{Y}$ (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:

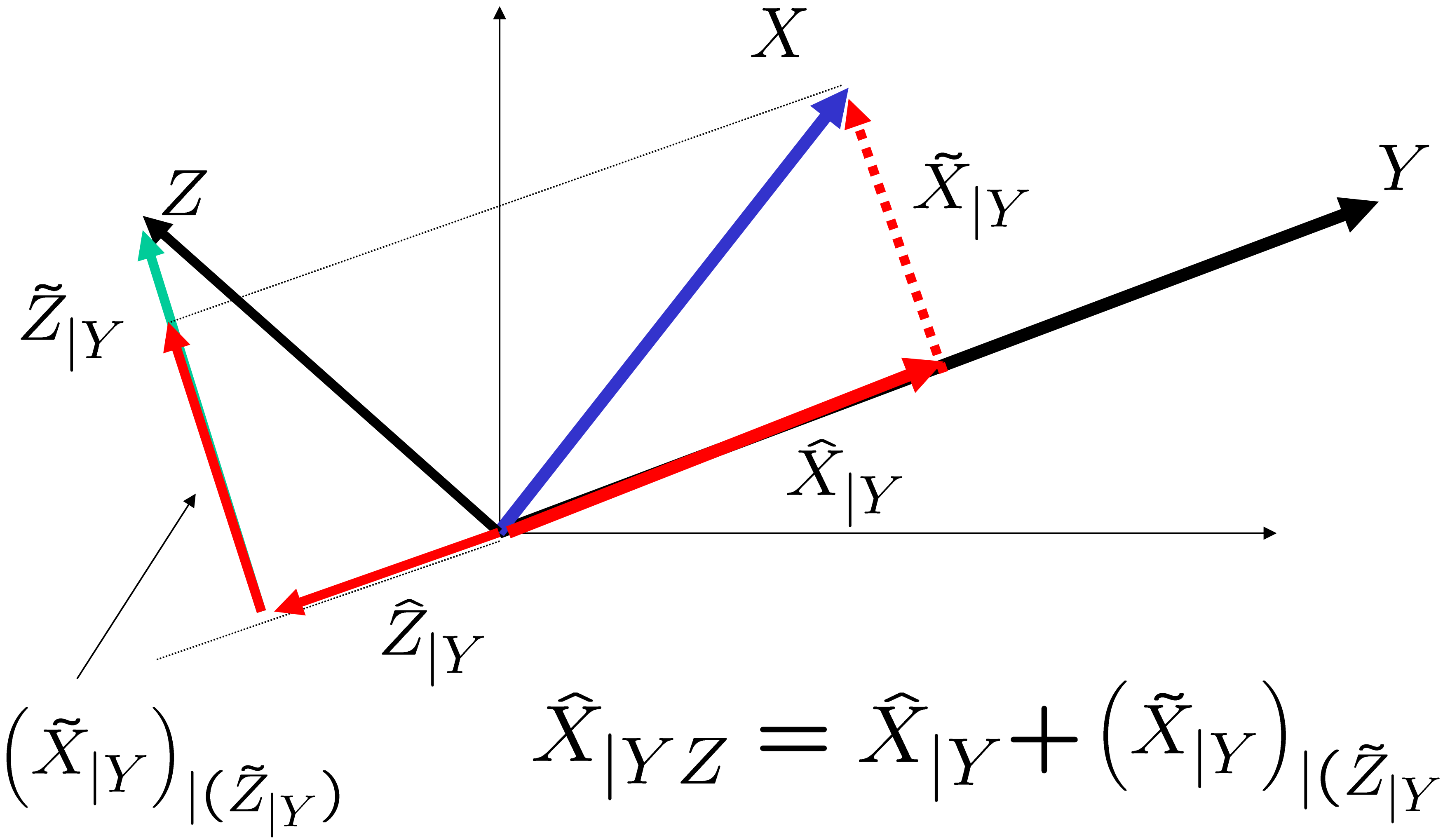
$$\tilde{X}_{|_y} = X - \hat{x}_{|_y}$$

This is still random

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

This is now an outcome

Deterministic interpretation of Property 3



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

Computation of $\left(\tilde{X}_{|Y}\right)_{|(\tilde{Z}_{|Y})}$

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

where:

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

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

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

Least Squares Estimation : Property 3

a) Recursive estimate

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

where:

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

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

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

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:

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

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

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

Derivation of Recursive LS Estimation

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

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

$$\widehat{x}_{|yz} = \widehat{x} + \Lambda_{XW} \underbrace{\Lambda_{WW}^{-1}}_{\text{inversion of an } (p+M) \times (p+M) \text{ matrix}} (w - \widehat{w})$$

Solution: use Schur complement

- Given

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

- Compute the Schur complement of Λ_{YY}

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

which is the conditional covariance

Solution: use Schur complement of Λ_{YY}

- Given

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

- Then

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

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

Non-Recursive LS Estimation

$$\hat{x}_{|yz} = \hat{x} + \underbrace{\Lambda_{XW}}_{\begin{bmatrix} \Lambda_{XZ} & \Lambda_{XY} \end{bmatrix}} \underbrace{\Lambda_{WW}^{-1}}_{\begin{bmatrix} \Lambda_{ZZ} & \Lambda_{ZY} \\ \Lambda_{YZ} & \Lambda_{YY} \end{bmatrix}^{-1}} \underbrace{(w - \hat{w})}_{\tilde{w}}$$

$$\tilde{w} = \begin{bmatrix} \tilde{z} \\ \tilde{y} \end{bmatrix}$$

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

$$\begin{bmatrix} \Lambda_{ZZ} & \Lambda_{ZY} \\ \Lambda_{YZ} & \Lambda_{YY} \end{bmatrix}^{-1}$$

$$\begin{bmatrix} \Lambda_{XZ} & \Lambda_{XY} \end{bmatrix}$$

Use Schur complement

$$\widehat{x}\big|_{yz} = \widehat{x} + \begin{bmatrix} \Lambda_{XZ} & \Lambda_{XY} \end{bmatrix} \underbrace{\begin{bmatrix} \Lambda_{ZZ} & \Lambda_{ZY} \\ \Lambda_{YZ} & \Lambda_{YY} \end{bmatrix}^{-1}} \begin{bmatrix} \tilde{z} \\ \tilde{y} \end{bmatrix}$$

$$\downarrow$$

$$\begin{bmatrix} \Lambda_{Z|Y}^{-1} & -\Lambda_{Z|Y}^{-1} F \\ -F^T \Lambda_{Z|Y}^{-1} & \Lambda_{YY}^{-1} + F^T \Lambda_{Z|Y}^{-1} F \end{bmatrix}$$

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

Use Schur complement

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

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

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

Use Schur complement

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

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

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

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

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

The conditional mean of $\boldsymbol{Z}$ given $\boldsymbol{Y} = \boldsymbol{y}$:

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

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

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

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

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

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

Using Gaussian least squares results:

$$E\{\tilde{X}_{|Y} | \tilde{z}_{|y}\} = \cancelto{\boldsymbol{0}}{E\{\tilde{X}_{|Y}\}}$$

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

Estimation errors always have zero means

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

Using Gaussian least squares results:

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

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

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

the conditional covariance

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,

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

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

Using Gaussian least squares results:

$$E\{\tilde{X}_{|Y} | \tilde{z}_{|y}\} = \underbrace{E\{\tilde{X}_{|Y} \tilde{Z}^T_{|Y}\}} \Lambda^{-1}_{Z|Y} \, \tilde{z}_{|y}$$

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

$$E\{\tilde{X} \tilde{Z}^T_{|Y}\} \; + \; \Lambda_{XY} \Lambda^{-1}_{YY} \cancelto{0}{E\{\tilde{Y} \tilde{Z}^T_{|Y}\}}$$

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

Using Gaussian least squares results:

$$E\{\tilde{X}_{|Y}|\tilde{z}_{|y}\} \quad = \quad \underbrace{E\{\tilde{X}\tilde{Z}^T_{|Y}\}}\Lambda^{-1}_{Z|Y}\,\tilde{z}_{|y}$$

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

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

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

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

Therefore,

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

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

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

$$\downarrow$$

$$\begin{bmatrix} \Lambda_{Z|Y}^{-1} & -\Lambda_{Z|Y}^{-1}F \\ -F^T\Lambda_{Z|Y}^{-1} & \Lambda_{YY}^{-1} + F^T\Lambda_{Z|Y}^{-1}F \end{bmatrix}$$

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

Use Schur complement

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

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

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

■

Summary

- The conditional mean is the least squares estimator:

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

- For Gaussians, the conditional mean is an affine function

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

Summary

The conditional mean can be computed recursively:

1. If we first know of outcome $\boldsymbol{Y} = \boldsymbol{y}$

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

Summary

The conditional mean can be computed recursively:

2 If we afterwards know of outcome $\boldsymbol{Z} = \boldsymbol{z}$

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

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

then

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

Course Outline

Finished

- Unit 0: Probability

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