LMMSE estimation, orthogonality

6.011, Spring 2018

Lec 14

LMMSE estimator:
first step (obtaining unbiasedness)

Linear estimator: $\widehat{Y}_\ell = aX + b$, with a and b picked to minimize $E[(Y - \widehat{Y}_\ell)^2]$ over joint density of X and Y

$$\Rightarrow \quad \min_{a,b} E[(\underbrace{Y - aX}_{Z} - b)^2]$$

First $\min_b E[(Z - b)^2] \quad \Rightarrow \quad b = \mu_Z = \mu_Y - a\mu_X$

This yields an **unbiased** estimator: $E[\widehat{Y}_\ell] = E[Y] = \mu_Y$

LMMSE estimator: second step (solve reduced problem)

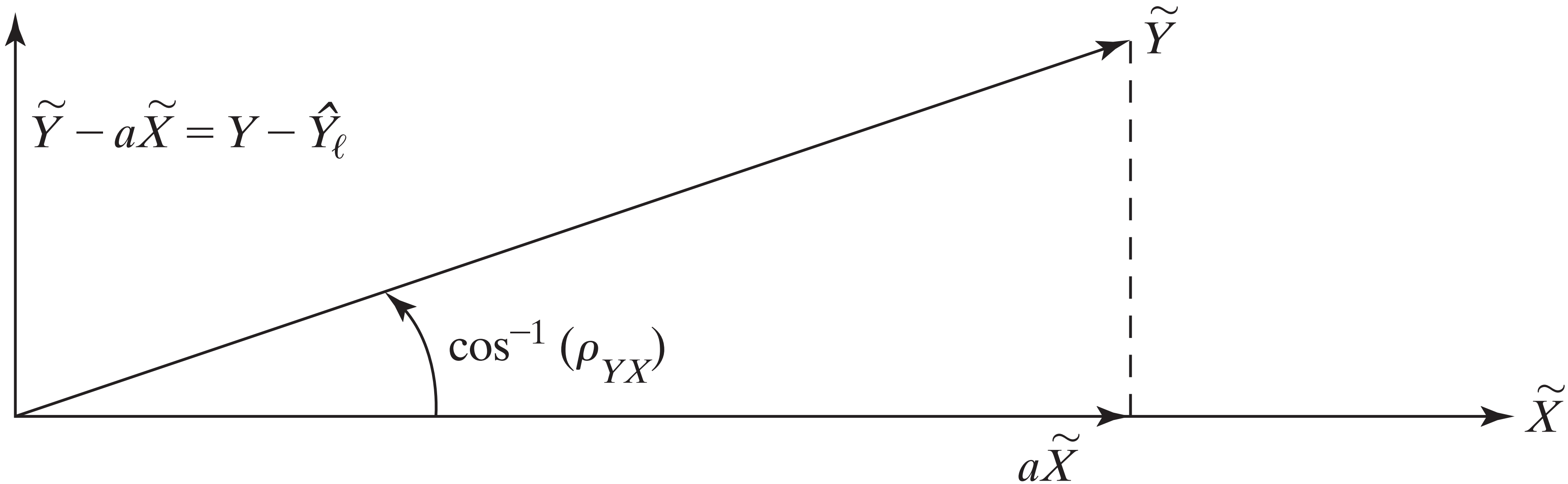
$$\text{Now } \min_a E[(Y - aX - b)^2] = E[(\underbrace{\{Y - \mu_Y\}}_{\widetilde{Y}} - a\underbrace{\{X - \mu_X\}}_{\widetilde{X}})^2]$$

$$\text{i.e.} \quad \min_a E[(\widetilde{Y} - a\widetilde{X})^2]$$

$$\Rightarrow \quad a = \frac{\sigma_{YX}}{\sigma_X^2} = \rho_{YX} \frac{\sigma_Y}{\sigma_X}$$

(can be shown in different ways, e.g., by vector picture)

LMMSE estimator as projection



$$\text{For the optimum } a, \quad (\widetilde{Y} - a\widetilde{X}) \perp \widetilde{X}$$

$$\text{i.e.,} \quad E[(\widetilde{Y} - a\widetilde{X})\widetilde{X}] = 0$$

$$\Rightarrow \quad a = \frac{\sigma}{\sigma_X^2} = \rho_{YX} \frac{\sigma}{\sigma_X}$$

Putting it all together

$$\widehat{Y}_\ell = \widehat{y}_\ell(X) = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(X - \mu_X)$$

or equivalently $$\frac{\widehat{Y}_\ell - \mu_Y}{\sigma_Y} = \rho \, \frac{X - \mu_X}{\sigma_X}$$

Also, the resulting MMSE is $\sigma_Y^2(1 - \rho^2)$

Orthogonality relations

Unbiasedness condition can be written as $Y - \widehat{Y}_\ell \perp 1$

(or $\perp$ to *any* constant)

We also know $(\widetilde{Y} - a\widetilde{X}) \perp \widetilde{X}$

or equivalently $Y - \widehat{Y}_\ell \perp \widetilde{X}$

or equivalently $Y - \widehat{Y}_\ell \perp X$

Conversely, first + last above yield equations for a, b

Extension to multivariate case

$$\min_{a_0,\ldots,a_L} E[(Y \quad \underbrace{\{a_0 + \Sigma_{j=1}^{L} a_j X_j\}}_{\widehat{Y}_\ell})^2]$$

$$\text{First } \min_{a_0} \quad \Rightarrow \quad a_0 = \mu_Y - \Sigma_{j=1}^{L} a_j \mu_{X_j}$$

This ensures unbiasedness of the estimator.

$$\text{Now } \min_{a_1,\ldots,a_L} E[(\widetilde{Y} - \Sigma_{j=1}^{L} a_j \widetilde{X}_j)^2]$$

Applying orthogonality gives the "normal equations"

$$E\left[\left(\widetilde{Y} - \Sigma_{j=1}^{L} a_j \widetilde{X}_j\right)\widetilde{X}_i\right] = 0$$

$$\begin{bmatrix} \sigma_{X_1 X_1} & \sigma_{X_1 X_2} & \cdots & \sigma_{X_1 X_L} \\ \sigma_{X_2 X_1} & \sigma_{X_2 X_2} & \cdots & \sigma_{X_2 X_L} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{X_L X_1} & \sigma_{X_L X_2} & \cdots & \sigma_{X_L X_L} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_L \end{bmatrix} = \begin{bmatrix} \sigma_{X_1 Y} \\ \sigma_{X_2 Y} \\ \vdots \\ \sigma_{X_L Y} \end{bmatrix}$$

$$(\mathbf{C}_{\mathbf{XX}})\,\mathbf{a} = \mathbf{c}_{\mathbf{X}Y}$$

$$\text{MMSE:} \quad \sigma_Y^2 - \mathbf{c}_{Y\mathbf{X}}(\mathbf{C}_{\mathbf{XX}})^{-1}\mathbf{c}_{\mathbf{X}Y} = \sigma_Y^2 - \mathbf{c}_{Y\mathbf{X}}.\mathbf{a}$$