Minimum mean square error (MMSE) estimator

Consider two random variables X and Y with a joint pdf $f_{X,Y}(x,y)$.

Let's say we want an estimator for Y given $X = x_0$ that minimizes the mean squared error:

$$ \mathrm{min} E[(Y-\hat{y})^2|X = x_0] $$

This is satisfied by the choosing $\hat{y}$ to be the conditional expectation of Y.

$$ \hat{y} = E[Y|X=x_0] $$

When X is a random variable, the estimator is also a random variable $\hat{Y}$:

$$ \hat{Y} = E[Y|X] $$

Linear MMSE

We can pick a line $\hat{y}_l(x) = ax + b$ that minimizes the mean squared error. That is:

$$ \underset{a, b}{\min} E[(Y - \hat{y}_l (X))^2] $$

We want this to be an unbiased esimator, so we can set b:

$$ b = \mu_Y - a\mu_X $$

When we plug in this b back into the above expression, we get:

$$ E[(Y - aX - \mu_Y + a\mu_X)^2] = E[((Y - \mu_Y) - (aX - \mu_X))^2] $$

After defining $\tilde{Y} = Y - \mu_Y$, $\tilde{X} = X - \mu_X$, we can rewrite the above expression as:

$$ E[(\tilde{Y} - a\tilde{X})^2] = \sigma_y^2 - 2a\sigma_{YX} + a^2\sigma_X^2 $$

The value of a that minimizes this expression is:

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

Resulting MSE

The resulting MSE from a linear MMSE estimator is:

$$ E[(\tilde{Y} - a\tilde{X})^2] = \sigma_Y^2 (1 - \rho^2) = \sigma_Y^2 \left[1 - \left(

\frac{\sigma_{XY}}{\sigma_X \sigma_Y} \right)^2 \right] $$

Orthogonalities

The following orthogonalities hold for the linear MSE estimator:

$$ (Y - \tilde{Y}_l) \perp 1 $$

(This means that that value has zero expected value.)

$$ (Y - a\tilde{X}) \perp \tilde{X} $$

$$ (Y - \tilde{Y}_l) \perp \tilde{X} $$

$$ (Y - \tilde{Y}_l) \perp X $$

Multivariate LMMSE estimator

$$ \hat{Y}_l = a_0 + a_1X_1 + \dots + a_LX_L $$

We set a_0 to make $\hat{Y}_l$ unbiased:

$$ a_0 = \mu_Y - \sum_{j=1}^L a_j\mu_{X_j} $$

The rest of the coefficients can be found using the following equation, also known as the normal equations:

$$ \mathbf{C}_{XX} \mathbf{a} = \mathbf{c}_{XY} $$

where

$$ \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_L \end{bmatrix} $$

$$ \mathbf{C}_{\mathbf{XX}} = \begin{bmatrix} \sigma_{X_1}^2 & \sigma_{X_1 X_2} & \dots & \sigma_{X_1 X_L} \\ \sigma_{X_2 X_1} & \sigma_{X_2}^2 & \dots & \sigma_{X_2 X_L} \\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{X_L X_1} & \sigma_{X_L X_2} & \dots & \sigma_{X_L}^2 \end{bmatrix} $$

$$ \mathbf{c}_{\mathbf{X}Y} = \begin{bmatrix} \sigma_{X_1 Y} \\ \sigma_{X_2 Y} \\ \vdots \\ \sigma_{X_L Y} \end{bmatrix} $$

Mean square error of multivariate LMMSE estimator

$$ \sigma_Y^2 - \mathbf{c}_{Y\mathbf{X}}\mathbf{a} $$

where $\mathbf{c}_{Y\mathbf{X}} = \mathbf{c}_{\mathbf{X}Y}^T$