Recap

MMSE Estimator

Have rv X which correlates w/ Y.

Have estimation $\hat{X} = g(Y)$

Error $W = X - \hat{X}$

Define MSE as $\mathbb{E}[W^2]$

① For $\hat{X}$ a constant,

$$\hat{X}^*_{MMSE} = \mathbb{E}[X]$$

$$MSE = Var(X)$$

② For $\hat{X}$ a function of Y

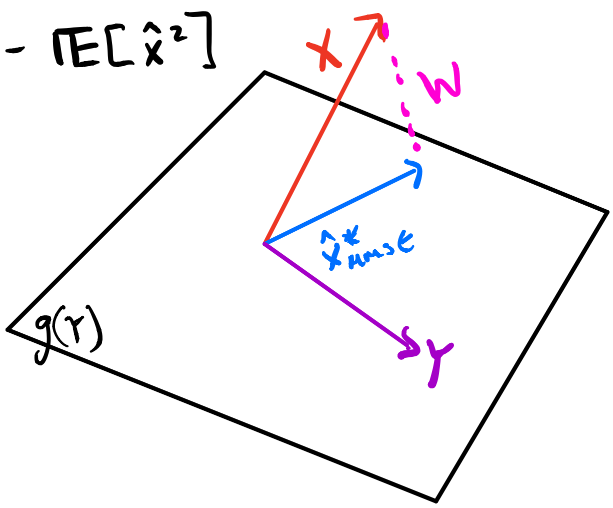
$$\hat{X}^*_{MMSE} = \mathbb{E}[X|Y]$$

$$MSE = \mathbb{E}_Y[Var(X|Y)]$$

Showed that

$$MSE = \mathbb{E}[X^2] - \mathbb{E}[\hat{X}^2]$$

$$W \perp g(Y)$$



Linear Minimum Mean Squared Error

Now our estimator $g(Y)$ takes the form

$$g(Y) = aY + b$$

The MSE becomes

$$MSE = \mathbb{E}[(X - g(Y))^2]$$
$$= \mathbb{E}[(X - aY - b)^2]$$

Want

$$\{a^*, b^*\} = \underset{a,b}{\operatorname{argmin}}\ \mathbb{E}[(X - aY - b)^2]$$

Use orthogonality principle

$$W \perp g(Y) \Rightarrow X - a^*Y - b^* \perp aY + b \quad \underline{\forall\, a, b}$$

Consider the following cases

① $a = 0,\ b = 1$

$X \perp Y \Leftrightarrow \mathbb{E}[XY]$

$$X - a^*Y - b^* \perp 1 \Rightarrow \mathbb{E}[\underbrace{X - a^*Y - b^*}] = 0$$

$W \rightarrow \mathbb{E}[W = 0]$

$\Rightarrow$ UNBIASED Estimator

$$\mathbb{E}[X] - \mathbb{E}[a^*Y] - \mathbb{E}[b^*] = 0$$

$$\rightarrow b^* = \mathbb{E}[X] - a^*\mathbb{E}[Y]$$

② $a = 1,\ b = 0$

$$X - a^*Y - b^* \perp Y \Rightarrow \mathbb{E}[(X - a^*Y - b^*)Y] = 0$$

$$\mathbb{E}[(X - a^*Y - b^*)Y]$$
$$= \mathbb{E}\left[(X - a^*Y - \mathbb{E}[X] - a^*\mathbb{E}[Y])Y\right]$$
$$= \mathbb{E}[(X - \mathbb{E}[X] - a^*(Y - \mathbb{E}[Y]))Y]$$

first term W

Note) $Cov(WY) = \mathbb{E}[WY]$ here

$$\Rightarrow Cov((X - \mathbb{E}[X] - a^*(Y - \mathbb{E}[Y]),\ Y) = 0$$

$$Cov(X, Y) - a^*Cov(Y, Y) = 0$$

$$a^* = Cov(X, Y)/Var(Y)$$

$Cov(Y, Y)$

Aside

$$\mathrm{Cov}(X, aY + bZ) = a\,\mathrm{Cov}(X,Y) + b\,\mathrm{Cov}(X,Z)$$

$$\mathrm{Cov}(X - \mathbb{E}[X], Y - \mathbb{E}[Y]) = \mathrm{Cov}(X,Y)$$

So

$$\hat{X}^*_{LMMSE} = a^* Y + b^*$$

$$= \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)} Y + \mathbb{E}[X] - \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)} \mathbb{E}[Y]$$

$$= \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)} (Y - \mathbb{E}[Y]) + \mathbb{E}[X]$$

and the MSE is

$$MSE = \mathbb{E}[W^2] = \mathrm{Cov}(W,W)$$

$$= \mathrm{Cov}\Big((X - \mathbb{E}[X] - a^*(Y - \mathbb{E}[Y]), (X - \mathbb{E}[X] - a^*(Y - \mathbb{E}[Y])\Big)$$

$$= \mathrm{Cov}(X,X) + a^{2*}\mathrm{Cov}(Y,Y) - 2a^*\mathrm{Cov}(X,Y)$$

$$= \mathrm{Var}(X) + \left(\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}\right)^2 \mathrm{Var}(Y) - 2\left(\frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}\right)\mathrm{Cov}(X,Y)$$

$$= \mathrm{Var}(X) - \frac{(\mathrm{Cov}(X,Y))^2}{\mathrm{Var}(Y)}$$

Example

$X \sim U(-1,1) \qquad Y = X^2$

$$f_X(x) = \begin{cases} 1/2, & -1 \le x \le 1 \\ 0, & o/w \end{cases}$$

MSE X to estimate Y

$$\hat{Y}_{MMSE} = \mathbb{E}[Y \mid X] = X^2$$

$$MSE = \mathbb{E}((Y - \hat{Y})^2) = 0$$

$$\hat{Y}_{LMMSE} = \frac{cov(X,Y)}{Var(X)}(X - \mathbb{E}[X]) + \mathbb{E}[Y] \qquad \mathbb{E}[X] = 0$$

$$= \mathbb{E}[Y]$$

$$MSE = \mathbb{E}[(Y - \hat{Y})^2] = \mathbb{E}[(Y - \mathbb{E}[Y])^2] = Var(Y)$$

Shows LMMSE isn't always preferred w/o linear relations

Example - Find LMMSE

$$f_{X,Y}(x,y) = \begin{cases} x+y, & 0 \le x,y \le 1 \\ 0, & \text{o/w} \end{cases}$$

$$\hat{X} = g(Y) = \frac{\text{Cov}(X,Y)^{①}}{\text{Var}(Y)^{②}} (Y - \mathbb{E}[\underset{③}{Y}]) + \mathbb{E}[X]^{④}$$

Need to compute 4 terms

$$\mathbb{E}[X] = \int_{y=0}^{1}\int_{x=0}^{1} x(x+y)\,dx\,dy = \int_{y=0}^{1} \frac{x^3}{3} + \frac{x^2}{2}y \Big|_{x=0}^{1} dy = \frac{1}{3}y + \frac{1}{4}y^2 \Big|_{y=0}^{1} = \frac{7}{12}$$

$$\mathbb{E}[Y] = \int_{y=0}^{1}\int_{x=0}^{1} y(x+y)\,dx\,dy = \frac{7}{12}$$

$$\mathbb{E}[XY] = \int_{x=0}^{1}\int_{y=0}^{1} (xy)(x+y)\,dy\,dx = \frac{1}{3}$$

$$\text{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{1}{3} - \left(\frac{7}{12}\right)^2 = \frac{-1}{144}$$

$$\text{Var}(Y) = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2$$

$$\mathbb{E}[Y^2] = \int_{x=0}^{1}\int_{y=0}^{1} y^2(x+y)\,dy\,dx = \frac{5}{12}$$

$$= \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{11}{144}$$

Thus

$$\hat{X}^*_{\text{LMMSE}} = \frac{-1}{11}\left(Y - \frac{7}{12}\right) + \frac{7}{12}$$

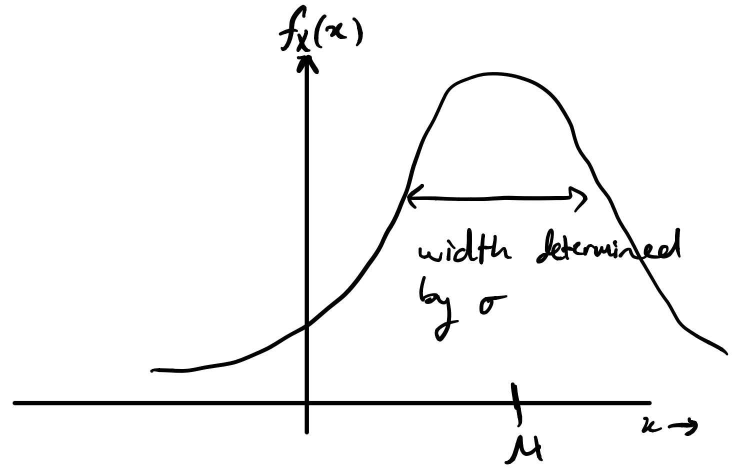
Jointly Gaussian Random Variables

If

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

then

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{\sigma^2}}$$



Definition: X, Y are jointly Gaussian if $aX + bY$ is Gaussian $\forall\, a, b \in \mathbb{R}$.

Definition: Joint pdf of such a distribution is

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{|K|}} \exp\left(-\frac{1}{2}\begin{pmatrix} x-\mu_x \\ y-\mu_y \end{pmatrix}^{T} K^{-1} \begin{pmatrix} x-\mu_x \\ y-\mu_y \end{pmatrix}\right)$$

<u>Note)</u> K is our covariance matrix.

Here,

$$K = \begin{pmatrix} Cov(X,X) & Cov(X,Y) \\ Cov(Y,X & Cov(Y,Y) \end{pmatrix}$$

$$= \mathbb{E}\left[\begin{pmatrix} X-\mu_X \\ Y-\mu_Y \end{pmatrix}\begin{pmatrix} X-\mu_X & Y-\mu_Y \end{pmatrix}\right]$$

Implications

① Jointly Gaussian $\Rightarrow$ Marginally Gaussian
$X, Y \quad \not\Leftarrow \quad X, Y$

Example

$$X \sim \mathcal{N}(0,1)$$
$$Y = \begin{cases} X & \text{if } |X| > 1 \\ -X & \text{if } |X| < 1 \end{cases}$$

$Pr(A) = Pr(A \cap B) \cup Pr(A \cap B^c)$ disjoint

$$F_Y(y) = Pr(Y \le y) = Pr(Y \le y, |X| > 1) + Pr(Y \le y, |X| < 1)$$

$$= Pr(x \le y \cap (x > 1 \text{ or } x < -1)) + Pr(x \ge -y, -1 \le x \le 1)$$
$$= Pr(x \le y, x > 1) + Pr(x \le y, x < -1) + Pr(x \ge -y, -1 \le x \le 1)$$

$$= \begin{cases} Pr(1 \le x \le y) + Pr(x < -1) + Pr(-1 \le x \le 1) = Pr(x \le y), & y \ge 1 \\ 0 + Pr(x < -1) + Pr(-y \le x \le 1) = Pr(-1 < x \le y), & -1 < y < 1 \\ 0 + Pr(x \le y) + 0 = Pr(x \le y), & y \le -1 \end{cases}$$

Thus

$$F_Y(y) = F_X(y) \Rightarrow X, Y \text{ marginally Gaussian}$$

But X, Y NOT jointly Gaussian.

Assume towards a contradiction $X + Y = 1/2$ (uses definition 1)

then

$$X + Y = \begin{cases} 2X, & |X| > 1 \\ 0, & |X| < 1 \end{cases} \Rightarrow |X| > 1 \Rightarrow X = \frac{1}{4} \quad \text{Contradiction!}$$