

## Recall

All r.v.'s  $(\Omega, \mathcal{F}, \mathbb{P})$

Linear Space: addition  $X+Y$   
scalar multiplication  $aX$   $bY$

Hilbert space is the same as a linear space w/ one more operation.

$$\text{inner product } \langle X, Y \rangle = \mathbb{E}[XY]$$

inner product definition allows for ("induces")

$$\text{length/norm: } \|X\| = \sqrt{\langle X, X \rangle}$$

$$\theta : \cos \theta = \frac{\langle X, Y \rangle}{\|X\| \|Y\|}$$

$$\text{distance: } d(X, Y) = \|X - Y\| = \sqrt{\langle X - Y, X - Y \rangle}$$

$$\text{orthogonality: } X \perp Y \Leftrightarrow \cos \theta = 0$$

$$\Rightarrow \langle X, Y \rangle = 0$$

$$\Rightarrow \mathbb{E}[XY] = 0$$

Have rv  $X$ , estimate  $\hat{X}$ .

Define error  $W = X - \hat{X}$

Want error<sup>2</sup> since  $-1, +1$  should have same error

Want to minimize average error.

Thus, want

$$\min_{\hat{X}} \mathbb{E}[W^2] = \min_{\hat{X}} \mathbb{E}[(X - \hat{X})^2]$$

① Restrict  $\hat{X}$  to be a constant.

So,

$$a^* = \operatorname{argmin}_{a \in \mathbb{R}} \mathbb{E}[(X - a)^2]$$

$$\text{MSE}_{a^*} = \text{Var}(X)$$

② Have rv  $Y$  whose observations determine  $X$ .

Restrict  $\hat{X}$  to be a function of  $Y$ ,  $g(Y)$ .  
ie

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

want

$$g^*(Y) = \operatorname{argmin}_g \mathbb{E}[(X - g(Y))^2]$$

Given  $Y=y$ , estimate  $\hat{X} = g^*(y)$  — a constant

So,

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

How we use  $Y$ ?

$$f_X(x) \rightarrow f_{X|Y=y}(X|Y=y)$$

So  $g^*$  is

$$g^*(Y) = \mathbb{E}[X|Y]$$

Statistic completely determined by variable you condition on

and

$$\text{MMSE} = \mathbb{E}_Y[\text{Var}(X|Y)] \quad g(Y)$$

The error

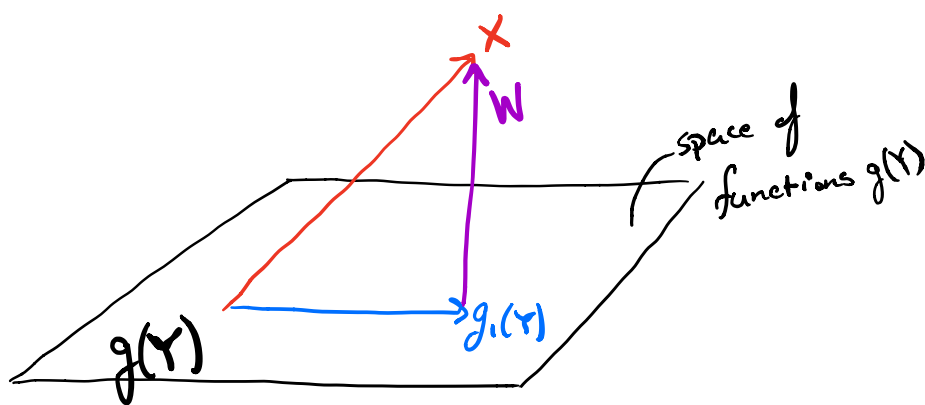
$$W = X - \mathbb{E}[X|Y]$$

has

$$\mathbb{E}[W] = \mathbb{E}[X - \mathbb{E}[X|Y]] = \mathbb{E}[X] - \mathbb{E}[X] = 0$$

we call this type of estimator UNBIASED

## VISUALIZATION



$$\begin{aligned} X, Y, \hat{X} \\ W = X - \hat{X}^* \\ \min_{\hat{X}} \mathbb{E}[W^2] \end{aligned}$$

- Orthogonality principle:

$$W \perp g(Y) \quad \forall g \quad \left. \begin{array}{l} \text{necessary} \\ \text{sufficient} \end{array} \right\} \text{AND condition}$$

$W$  induced by  $\hat{X}^*$  (best estimator)

Claim  $X - \mathbb{E}[X|Y] \perp g(Y) \forall g$

Proof  $\mathbb{E}[(X - \mathbb{E}[X|Y])g(Y)] = 0$  inner product  $\langle X - \mathbb{E}[X|Y], g(Y) \rangle$

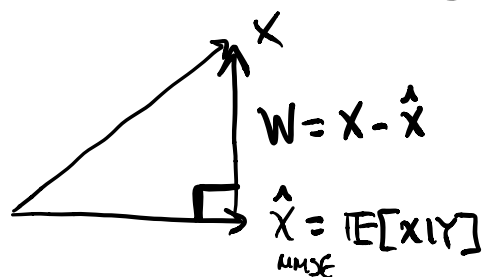
LHS

$$\mathbb{E}_{xy}[Xg(Y)] - \mathbb{E}_Y[\mathbb{E}_{xy}[X|Y]g(Y)]$$

$$\mathbb{E}_Y[\mathbb{E}_{x|Y}[Xg(Y)|Y]] - \mathbb{E}_Y[\mathbb{E}_{x|Y}[X|Y]g(Y)]$$

$$\mathbb{E}_Y[g(Y)\mathbb{E}_{x|Y}[X|Y]] - \mathbb{E}_Y[g(Y)\mathbb{E}_{x|Y}[X|Y]] = 0 \checkmark$$

MSE Using Pythagorean Theory



$$\mathbb{E}[X] = \mathbb{E}[\hat{X}]$$

$\|w\|^2 = \langle w, w \rangle$  in Hilbert Space

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

$$= \text{Var}(X) + (\mathbb{E}(X))^2 - (\text{Var}(\hat{X}_{MSE}) + (\mathbb{E}[\hat{X}_{MSE}])^2)$$

$$= \text{Var}(X) - \text{Var}(\hat{X}_{MSE})$$

Example (Like Prelim 1):

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

$\hat{X}_{MSE}$  using  $Y$ ?  $MSE$ ?

Know  $\hat{X}_{MSE}$  is  $E[X|Y]$  and  $MSE$  is  $E_Y[\text{Var}(X|Y)]$

Step 1: Find  $f_Y(y)$ .  $f_Y(y) = \int_x f_{X,Y}(x,y) dx = \begin{cases} \frac{1}{2} + y & , 0 \leq y \leq 1 \\ 0 & , \text{o/w} \end{cases}$

Step 2: Find  $f_{X|Y}(x|y)$ .  $f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} \frac{x+y}{\frac{1}{2}+y} & , 0 \leq x,y \leq 1 \\ 0 & , \text{o/w} \end{cases}$

$$\begin{aligned} \text{Step 3: } E[X|Y=y] &= \int_{x=0}^1 x f_{X|Y}(x|y) dx \\ &= \int_{x=0}^1 \frac{x+y}{\frac{1}{2}+y} x dx = \int_{x=0}^1 \frac{x^2 + xy}{\frac{1}{2}+y} dx \\ &= \frac{2}{1+y} \left( \frac{1}{3} x^3 + \frac{1}{2} x^2 y \bigg|_0^1 \right) \\ &= \frac{2}{1+2y} \left( \frac{1}{3} + \frac{1}{2} y \right) \\ &= \frac{2\left(\frac{1}{3} + \frac{1}{2} y\right)}{1+2y} = \frac{2+3y}{3+6y} \end{aligned}$$

thus

$$E[X|Y] = \frac{2+3Y}{3+6Y}$$

$$\begin{aligned}\text{Step 4: } \text{Var}(X|Y=y) &= \mathbb{E}[X^2|Y=y] - (\mathbb{E}[X|Y=y])^2 \\ &= \int_{x=0}^1 x^2 f_{X|Y}(x|y) dx - \left(\frac{2+3y}{3+6y}\right)^2\end{aligned}$$

$$= \frac{3+4y}{6+12y} - \left(\frac{2+3y}{3+6y}\right)^2$$

Then

$$\text{Var}(X|Y) = \frac{3+4Y}{6+12Y} - \left(\frac{2+3Y}{3+6Y}\right)^2$$

$$\text{Step 5: } \mathbb{E}_Y[\text{Var}(X|Y)] = \int_y \text{Var}(X|Y=y) f_Y(y) dy$$