

problem 1 : linear MMSE estimation

estimating scalar  $\gamma$

from vector  $X \in \mathbb{R}^n$

And linear MMSE estimator  $\hat{\gamma} = w^T X$   
that minimizes MSE,  $E[(\gamma - \hat{\gamma})^2]$

a)  $K = \begin{bmatrix} 5 & 2 \\ 2 & 4 \end{bmatrix}$  find optimal  $w^*$

$$\hat{\gamma} = w^* X$$

Variance = how spread out values are from the mean.

Variance = avg of  $(\text{value} - \text{mean})^2$

Covariance = if variables move together

$$K = \begin{bmatrix} \text{Var}(X) & \text{Cov}(X, Y) \\ \text{Cov}(X, Y) & \text{Var}(Y) \end{bmatrix}$$

$$\text{Var}(X) = 5$$

$$\text{Var}(Y) = 4$$

$\text{Cov}(X, Y) = 2 \rightarrow$  move together positively

1a) finding optimal weight

$$\hat{Y} = w \times X$$

$\hat{Y}$  = our prediction of  $Y$

$w$  = ? weight

$X$  = value we observe

good weight  $\rightarrow$  close to true values  $\rightarrow$  mean squared error.

$$MSE = \text{Avg } + (Y - \hat{Y})^2$$

$$\text{optimal } w^* = \frac{\text{cov}(X, Y)}{\text{var}(X)}$$

$$w^* = \frac{2}{5} = 0.4$$

1b) calculate min. MMSE

$$\text{MMSE} = \text{Var}(Y) - \frac{[\text{cov}(X, Y)]^2}{\text{var}(X)}$$

$\text{Var}(Y)$  = total uncertainty in  $Y$

$$\frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X)} = \text{how much uncertainty we remove by using } X$$

The difference = leftover uncertainty

$$\text{MMSE} = 4 - \left( \frac{(2)^2}{5} \right)$$

$$= 4 - \frac{4}{5} = \frac{16}{5} = 3.2$$

$$\boxed{\text{MMSE} = 3.2}$$

$$1c) \underbrace{w^* x}_{\text{best predictor of gaussian random variables}} = \underbrace{E[Y|X]}_{\text{conditional expectation average value of } Y, \text{ given we know } x}$$

best predictor  
of gaussian  
random variables

conditional expectation  
average value of  $Y$ ,  
given we know  $x$

$$E[Y|X] = \left( \frac{\text{Cov}(X, Y)}{\text{Var}(X)} \right) x + \bar{Y}$$

$$E[Y|X] = E[Y] + \frac{\text{cov}(X, Y)}{\text{var}(X)} \times (X - E[X])$$

zero mean.

$$E[Y|X] = \frac{2}{5} \times X$$

$w^{*1} = 2/5$

$$E[Y|X] = w^* \times X \quad \checkmark$$

1d) non gaussian

$$\hat{Y} = \tilde{w} X$$

linear MMSE only uses the covariance matrix  
doesn't care abt the actual distribution.

therefore  $w$  stays the same and MMSE  
stays the same, it doesn't matter  
where they come from (uniform, gaussian,  
etc)  
since the matrix is the same.