2024S AAI 627-WS1

E Navaneet Kumar

MIDTERM

Q1

MA(2) process is

$$X_t = e_t + a_1 e_{t-1} + a_2 e_{t-2}$$

a) The expectation $E(X_t)$

$$E(X_t) = E(e_t) + a_1 E(e_{t-1}) + a_2 E[e_{t-2})$$

$e_t \sim N(0, \sigma^2)$ is iid process

$$E(e_t) = E[e_{t-1}) = E(e_{t-2}) = 0$$

$$E[X_t] = 0$$

b) The Variance $Var[X_t]$

$$Var(X_t) = E[x_t^2] - \cancel{(E(X_t))^2}^{\,0}$$

$$E[x_t^2] = E[e_t^2 + a_1^2 e_{t-1}^2 + a_2^2 e_{t-2}^2 + 2a_1 e_t e_{t-1} + a_2 e_t e_{t-2} + a_1 a_2 e_{t-1} e_{t-2} + a_2 e_t e_{t-2} + a_2 a_1 e_{t-1} e_{t-2}]$$

$$E[x_t^2] = \sigma^2 + a_1^2 \sigma^2 + a_2^2 \sigma^2$$

$$= \sigma^2 (1 + a_1^2 + a_2^2)$$

$$Var(X_t) = E(X_t^2) = \sigma^2 (1 + a_1^2 + a_2^2)$$

c) The first order auto correlation $\rho(1)$

$$\rho(1) = \frac{Cov(X_t, X_{t-1})}{Var(X_t)}$$

$$Cov(X_t, X_{t-1}) = E[(X_t - \cancel{\mu_{X_t}}^{0})(X_{t-1} - \cancel{\mu_{X_{t-1}}}^{0})]$$

$$= E[X_t X_{t-1}]$$

$$Cov(X_t, X_{t-1}) = E[(a_1 e_{t-1} + a_2 e_{t-2} + e_t)$$
$$(e_{t-1} + a_1 e_{t-2} + a_2 e_{t-3})$$

$$= E[a_1 e_{t-1}^2 + a_2 a_1 e_{t-2}^2]$$

($\because$ all other terms are zero)

$$Cov(x_t, x_{t-1}) = a_1\sigma^2 + a_1 a_2 \sigma^2$$

$$\rho(1) = \frac{a_1\sigma^2(1+a_2)}{\sigma^2(1+a_1^2+a_2^2)}$$

$$= \frac{a_1(a_2+1)}{1+a_1^2+a_2^2}$$

d) $$\rho(3) = \frac{Cov(x_t, x_{t-3})}{Var(x_t)}$$

$$Cov(x_t, x_{t-3}) = E[x_t \cdot x_{t-3}]$$

$$= E[(a_1 e_{t-1} + a_2 e_{t-2} + e_t)(e_{t-3} + a_1 e_{t-4} + a_2 e_{t-5})]$$

$$Cov(x_t, x_{t-3}) = 0$$

$$\rho(3) = 0 \quad //$$

Q2

Given

$$X_t = a X_{t-1} + W_t$$

$\Rightarrow$ $X_2 = a X_1 + W_2$

$X_3 = a X_2 + W_3$

$\rightarrow$ The mean of X_t is 0

The variance of X_t at stationary is

$\frac{\sigma_w^2}{1-a^2}$, $\sigma_w^2 = 1$

Covariance

$Cov(X_1, X_2) = Cov(X_1, aX_1 + W_2) = a \cdot Var(X_1)$

$Cov(X_1, X_3) = Cov(X_1, aX_2 + W_3) = a \, Cov(X_1, X_2)$

$= a^2 \, Var(X_1)$

$Var(X_1) = Var(X_2) = Var(X_3) = \frac{1}{1-a^2}$

$\rightarrow$ The best linear Predictor

The best linear predictor $\hat{X}_2$ given X_1 & X_3

$$\hat{X}_2 = E[X_2 \mid X_1, X_3]$$

$$= \beta_1 X_1 + \beta_2 X_3$$

$$\beta_1 = \frac{Cov(X_2, X_1)}{Var(X_1) + Var(X_3) - 2Cov(X_1, X_3)}$$

$$\beta_2 = \frac{Cov(X_2, X_3)}{Var(X_1) + Var(X_3) - 2Cov(X_1, X_3)}$$

$$\Rightarrow \quad \beta_1 = \frac{a \frac{1}{1-a^2}}{\frac{1}{1-a^2} + \frac{1}{1-a^2} - 2a^2 \frac{1}{1-a^2}}$$

$$\beta_2 = \frac{a \frac{1}{1-a^2}}{\frac{1}{1-a^2} + \frac{1}{1-a^2} - 2a^2 \frac{1}{1-a^2}}$$

by simplifying.

we find that $\beta_1 = \beta_2 = \dfrac{a}{1+a^2}$

$$\Rightarrow \hat{X}_2 = \frac{a}{1+a^2} X_1 + \frac{a}{1+a^2} X_3$$

$$\Rightarrow \hat{X}_2 = \frac{a}{1+a^2} (X_1 + X_3)$$

This linear combination of X_1 & X_3 that predicts X_2 based on the least squares in the content of a Gaussian dist.

Q3

$X_t = 2w_t + w_{t-1} + n_t$

$\gamma_{w,x}(h) = E[w_{t+h} X_t]$

$X_t = 2w_t + w_{t-1} + n_t$

$\gamma_{w,x}(h) = E[w_{t+h}(2w_t + w_{t-1} + n_t)]$

$\gamma_{w,x}(h) = 2E[w_{t+h} w_t] + E[w_{t+h} \cdot w_{t-1}] + E[w_{t+h} n_t]$

w_t & n_t are uncorrelated for any time shift

$E[w_{t+h} n_t] = 0$ for all h

$\gamma_{w,x}(h) = 2E[w_{t+h} w_t] + E[w_{t+h} w_{t-1}]$

Now lets compute $E[w_{t+h} w_t]$ & $E[w_{t+h} w_{t-1}]$ for different values of h.

w_t a white noise process

$E[w_t w_t] = Var(w_t) = 3$ (given)

$E[w_t w_s] = 0$ for any $t \neq s$ because the white noise is uncorrelated at different times

So,
for $h = 0$
$E[w_t w_t] = 3$ & $E[w_t w_{t-1}] = 0$
because t & $t-1$ are different times.

for $h = -1$
$E[w_{t-1} w_t] = 0$ because t & $t-1$ are different times, but

$E[w_{t-1} w_{t-1}] = 3$ because its variance of w_{t-1}

$h = 1$
$E[w_{t+1} w_t] = 0$ because t & $t+1$ are different times, & $E[w_{t+1} w_{t-1}] = 0$

because $t+1$ & $t-1$ are different times

for $|h| > 1$: both terms are 0 because the times are different.

So the cross-correlation function $\gamma_{w,x}(h)$ is

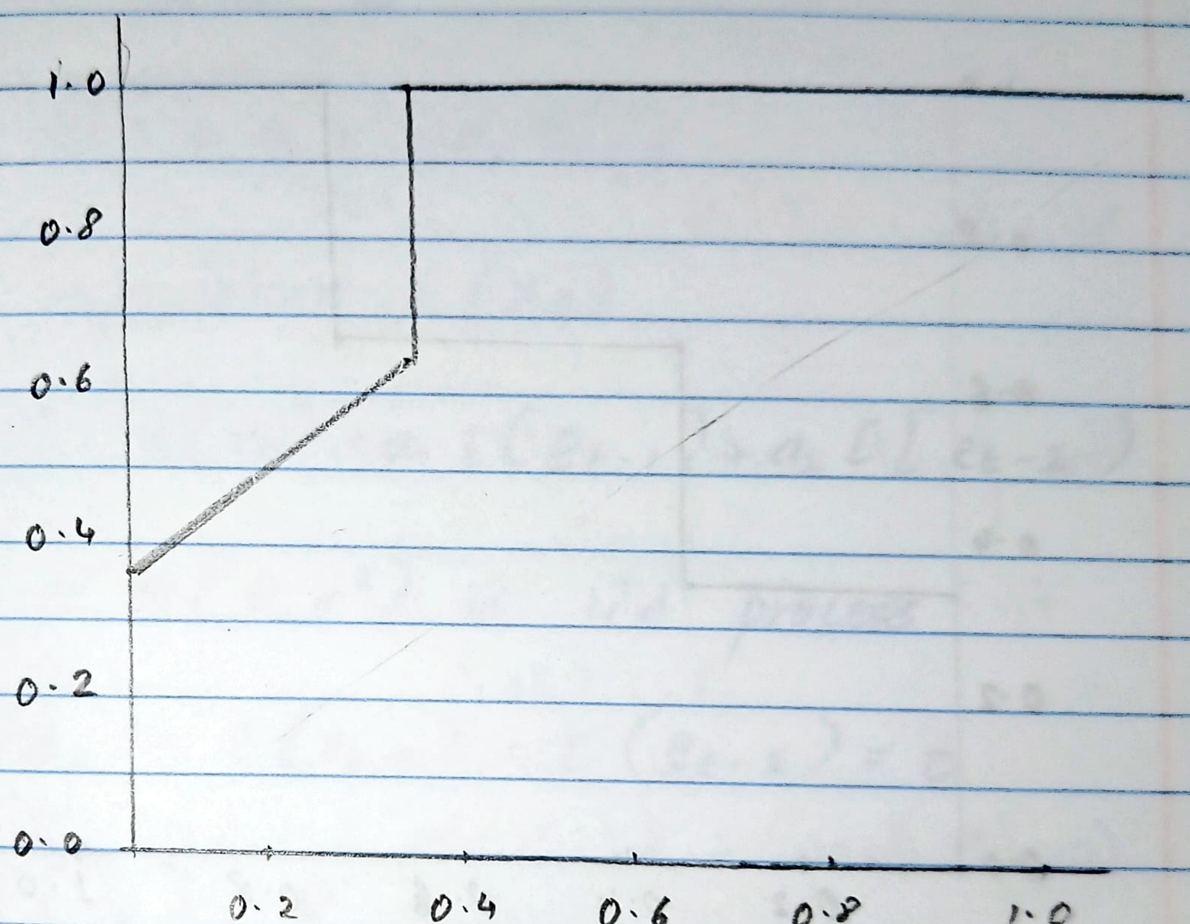
$$\gamma_{w,x}(h) = \begin{cases} 2 \cdot 3 + 0 = 6 & \text{if } h = 0 \\ 0 + 3 = 3 & \text{if } h = -1 \\ 0 + 0 = 0 & \text{if } h = 1 \\ 0 & \text{if } |h| > 1 \end{cases}$$

2) Cross correlation function b/w $\{w_t\}$ & $\{x_t\}$

$$\gamma_{w,x}(h) = \begin{cases} 6 & \text{if } h = 0 \\ 3 & \text{if } h = -1 \\ 0 & \text{otherwise} \end{cases}$$

Q4

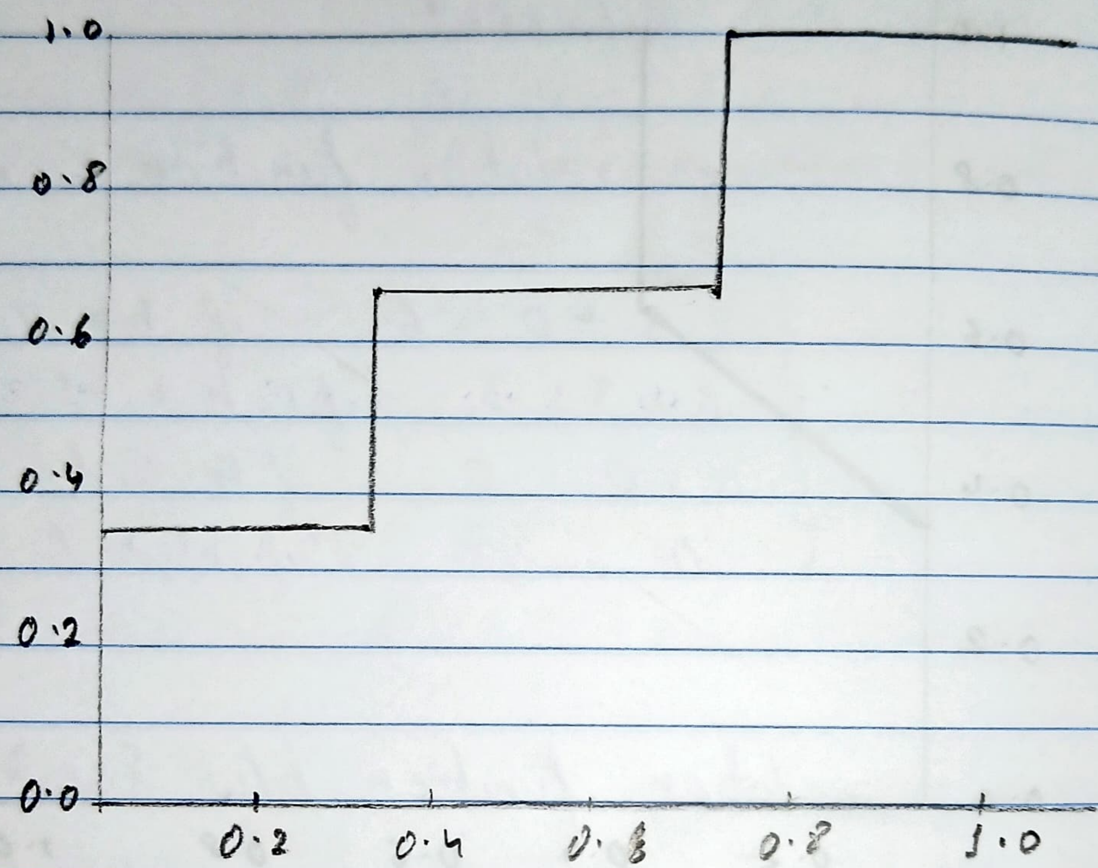
ROC for Method 1



Method 1

AUC = 0.83

ROC for Method 2



Method 2

AUC = 0.67

⇒ Based on the AUC value which is

AUC of Method 1 = 0.83

AUC of Method 2 = 0.67

∴ Method 1 is relatively better in terms of AUC value.

Q5

a) Autoregressive, AR(2)

order $q = 0$, $p = 2$

b) Moving Average, MA(1)

order $p = 0$, $q = 1$

c) Autoregressive, AR(1)

order, $q = 0$, $p = 1$

Q6

a) Normalizing the predictors in a logistic regression model will not change the AUC performance of the model. The reason is that normalization is simply a linear transformation of the feature scale & does not change the rank order of the predictions.

Since AUC is a rank-order metric ie it depends only on the order in which the probabilities are ranked, not on their absolute values, So the performance as measured by the AUC remains the same.

b) Transforming each predictor by taking the exponential (as in $y_i = e^{x_i}$) will change the relationship b/w the predictors & the response variable in a non-linear way. This transformation can lead to different decision boundaries in the logistic regression model.

Unlike normalization, this not a linear transformation and can change the rank ordering of the predicted probabilities, which in turn can increase or decrease the AUC performance. The increase or decrease of AUC depends on how well the transformed predictors allign with the log-odds of the response variable. If the exponential transformation makes the predictors better at seperating the two classes, the AUC will improve. Conversely, if the transformation makes the seperation worse, the AUC will decrease.