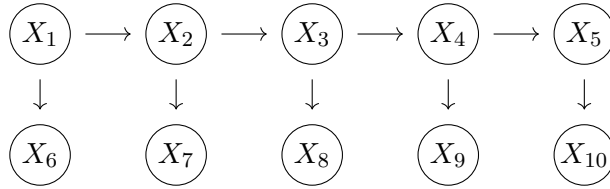


DS-GA 1018: Homework 1

Due Friday September 29th at 5:00 pm

Problem 1 (5 points): Consider the graphical model below:



i. (1 points): Write down a full factorization of $p(X_{1:10})$ implied by the graphical model. Your factorization should be as simple as possible (*simplicity is measured by the number of X_* terms that show up in your final expression*).

Solution: By chain rule of probability we have

$$P(X_{1:10}) = \prod_{i=1}^{10} P(X_i | X_{1:i-1}) P(X_1)$$

By D-separation algorithm, we have: $X_i \perp X_{1:i-2} | X_{i-1}, \forall i \in \{3, 4, 5\}$
 $\therefore P(X_i | X_{1:i-1}) = P(X_i | X_{i-1}), \forall i \in \{3, 4, 5\}$ Also we know that:

1. $i = 6$, X_1 d-separates x_6 from the rest, $X_6 \perp X_{2:5} | X_1$
2. $i = 7$, X_2 d-separates x_7 from the rest, $X_7 \perp X_{1,3:6} | X_2$
3. $i = 8$, X_3 d-separates x_8 from the rest, $X_8 \perp X_{1,2,4:7} | X_3$
4. $i = 9$, X_4 d-separates x_9 from the rest, $X_9 \perp X_{1,3,5:8} | X_4$
5. $i = 10$, X_5 d-separates x_{10} from the rest, $X_{10} \perp X_{1,4,6:9} | X_5$

\therefore Finally we have:

$$P(X_{1:10}) = P(X_1)P(X_2|X_1)P(X_3|X_2)P(X_5|X_4)P(X_6|X_1) \\ P(X_7|X_2)P(X_8|X_3)P(X_9|X_4)P(X_{10}|X_5)$$

ii. (1 points): What is the Markov boundary of X_4 ?

Solution: By definition, X_4 's

1. **Parents:** $\{X_3\}$

2. **Children:** $\{X_5, X_9\}$

3. **Paraent of Children:** $\{X_4\}$ (Excluded)

\therefore Markov boundary is $\{X_3, X_5, X_9\}$.

iii. (1 points): What is the Markov boundary of X_8 ?

Solution: By definition, X_8 's

1. **Parents:** $\{X_3\}$

2. **Children:** \emptyset

3. **Paraent of Children:** \emptyset (Excluded)

\therefore Markov boundary is $\{X_3\}$.

iv. (2 points): Write down a full factorization of $p(X_{1:3}, X_{5:10}|X_4)$ that is as simple as possible.

Solution: We will have the following derivations:

$$\begin{aligned}
P(X_{1:3}, X_{6:8}|X_4) &= P(X_1, X_2, X_{6:8}|X_3, X_4)P(X_3|X_4) && \text{(Bayes Rule)} \\
&= P(X_1, X_2, X_{6:8}|X_3)P(X_3|X_4) && (X_3 \text{ d-separates } X_4 \text{ and the rest}) \\
&= P(X_1, X_2, X_6, X_7|X_8, X_3)P(X_8|X_3)P(X_3|X_4) && \text{(Bayes Rule)} \\
&= P(X_1, X_2, X_6, X_7|X_3)P(X_8|X_3)P(X_3|X_4) && (X_3 \text{ d-separates } X_8 \text{ and the rest}) \\
&= P(X_1, X_6, X_7|X_2, X_3)P(X_2|X_3)P(X_8|X_3)P(X_3|X_4) && \text{(Bayes Rule)} \\
&= P(X_1, X_6, X_7|X_2)P(X_2|X_3)P(X_8|X_3)P(X_3|X_4) && (X_2 \text{ d-separates } X_3 \text{ and the rest}) \\
&= P(X_6|X_1, X_7, X_2)P(X_1|X_2)P(X_7|X_2)P(X_2|X_3)P(X_8|X_3)P(X_3|X_4) && \text{(Bayes rule, } X_1 \perp X_7|X_2) \\
&= P(X_6|X_1)P(X_1|X_2)P(X_7|X_2)P(X_2|X_3)P(X_8|X_3)P(X_3|X_4) && (X_1 \text{ d-separates } X_2 \text{ and } X_7)
\end{aligned}$$

Similarly:

$$\begin{aligned}
P(X_5, X_9, x_{10}|x_4) &= P(X_5, X_{10}|X_9, X_4)P(X_9|X_4) && \text{(Bayes Rule)} \\
&= P(X_5, X_{10}|X_4)P(X_9|X_4) && (X_4 \text{ d-separates } X_9 \text{ and the rest}) \\
&= P(X_{10}|X_5, X_4)P(X_5|X_4)P(X_9|X_4) && \text{(Bayes Rule)} \\
&= P(X_{10}|X_5)P(X_5|X_4)P(X_9|X_4) && (X_5 \text{ d-separates } X_4 \text{ and the rest})
\end{aligned}$$

Since $P(X_{1:3}, X_{5:10}|X_4) = P(X_{1:3}, X_{6:8}|X_4)P(X_5, X_9, X_{10}|X_4)$, we get the final expression:

$$\begin{aligned}
P(X_{1:3}, X_{5:10}|X_4) &= P(X_6|X_1)P(X_1|X_2)P(X_7|X_2)P(X_2|X_3)P(X_8|X_3) \\
&\quad P(X_3|X_4)P(X_{10}|X_5)P(X_5|X_4)P(X_9|X_4)
\end{aligned}$$

Problem 2 (13 points): Consider the two following MA(4) processes:

$$X_t = W_t + \theta_3 W_{t-3} + \theta_4 W_{t-4} + \theta_c \quad (1)$$

$$Y_t = W_t + \theta_1 W_{t-1} + \theta_4 W_{t-4}, \quad (2)$$

where W_t is drawn from $\mathcal{N}(0, \sigma_W^2)$ and all the θ_* are constants.

i. (2 points): What is the mean, $\mu_X(t)$, of the $\{X_t\}$ process? Justify your answer.

Solution: Given that $\mathbb{E}[W_t] = 0, \forall t$, we have:

$$\begin{aligned} \mu_X(t) &= \mathbb{E}[X_t] \\ &= \mathbb{E}[W_t + \theta_3 W_{t-3} + \theta_4 W_{t-4} + \theta_c] \\ &= \theta_c \end{aligned}$$

ii. (3 points): What is the covariance, $\gamma_X(t, s)$, of the $\{X_t\}$ process?

Solution: From the definition of autocovariance function, we have:

$$\begin{aligned} \gamma_X(|h|) &= \mathbb{E}\left[\left(\sum_{j=0}^4 \theta_j W_{t+h-j}\right)\left(\sum_{j=0}^4 \theta_j W_{t-j}\right)\right] \\ &= \begin{cases} \sigma_W^2(1 + \theta_3^2 + \theta_4^2) & |h| = 0 \\ \sigma_W^2 \theta_3 \theta_4 & |h| = 1 \\ \sigma_W^2 \theta_3 & |h| = 3 \\ \sigma_W^2 \theta_4 & |h| = 4 \\ 0 & |h| > 4 \text{ or } |h| = 2 \end{cases} \end{aligned}$$

iii. (1 points): Is $\{X_t\}$ drawn from a weakly stationary process?

Solution: Since the auto-covariance is only dependent on $|h|$, and by (i) we have $\mathbb{E}[X_t] = \text{constant}$, thus by definition $\{X_t\}$ is **weakly stationary**.

iv. (5 points): What is the cross-covariance, $\gamma_{X,Y}(t, s)$, between X_t and Y_s ?

Solution: By definition of the covariance-function we have:

$$\begin{aligned}\gamma_{X,Y}(s, t) &= \mathbb{E}[(W_t + \theta_3 W_{t-3} + \theta_4 W_{t-4})(W_s + \theta_1 W_{s-1} + \theta_4 W_{s-4})] \\ &= \begin{cases} \mathbb{E}[W_t^2 + \theta_4^2 W_{t-4}^2] & s = t \\ \mathbb{E}[\theta_1 \theta_3 W_{t-3} W_{s-1}] & s = t - 2 \\ \mathbb{E}[\theta_1 \theta_4 W_{t-4} W_{s-1}] & s = t - 3 \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} \sigma_W^2(1 + \theta_4^2) & s = t \\ \sigma_W^2 \theta_1 \theta_3 & s = t - 2 \\ \sigma_W^2 \theta_1 \theta_4 & s = t - 3 \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

v. (2 points): Is it possible for $\gamma_X(t, t) = 0$? If so, what is one value of $\theta_1, \theta_2, \theta_3, \theta_4$ that satisfied this? Limit yourself to the real numbers.

We know from (iii) that

$$\gamma_X(t, t) = \gamma_X(0) = \sigma_W^2(1 + \theta_3^2 + \theta_4^2)$$

Thus we can conclude that:

$$\begin{cases} \text{If } \sigma_w^2 \neq 0, \text{ then } r_x(t, t) > 0 & \forall \theta_3, \theta_4 \in R \\ \text{If } \sigma_w^2 = 0, \text{ then } r_x(t, t) = 0 & \forall \theta_3, \theta_4 \in R \end{cases}$$

Problem 3 (10 points): Consider the following two models:

$$X_t = 2.5X_{t-1} - X_{t-2} + W_t - 2W_{t-1} \quad (3)$$

$$Y_t = 0.7Y_{t-1} + 0.3Y_{t-2} + W_t - 0.4W_{t-1}, \quad (4)$$

where W_t is drawn from $\mathcal{N}(0, \sigma_W^2)$.

i. (3 points): Identify $\{X_t\}$ as ARMA(p, q). *Watch out for parameter redundancy.*

Solution: We follow the definition and could get:

$$X_t - 2.5X_{t-1} - (-X_{t-2}) = W_t - 2W_{t-1}$$

where. $\phi_1 = 2.5, \phi_2 = -1, \theta_1 = -2$ and $\phi(B) = B^2 - 2.5B + 1, \theta(B) = 1 - 2B$

$$\therefore \phi(B)X_t = \theta(B)\omega_t$$

notice that $\phi(B) = (B - 2)(B - 0.5)$ and $\theta(B) = -2(B - 0.5)$

\therefore There is parameter redundancy. Thus we cancel them out and could obtain

$$\phi(B) = B - 2 \text{ and } \theta(B) = -2$$

\therefore We might just fit a new ARMA (p, q) model:

$$\begin{aligned} -2X_t + X_{t-1} &= -2W_t \\ \Leftrightarrow X_t - \frac{1}{2}X_{t-1} &= W_t \end{aligned}$$

where $p = 1, q = 0$

ii. (1 points): Is $\{X_t\}$ causal? Justify your response.

Solution: From the causality criteria we compute the root of its AR's characteristic polynomial

$$z^2 - 2.5z + 1 = 0$$

where $z_{1,2} = 2, 0.5$

Since $z_2 = 0.5$ is inside unit circle, the process **isn't causal**.

iii. (1 points): Is $\{X_t\}$ invertible? Justify your answer.

Solution: From the invertibility critend ur compute the root of its MA's charctenistic polynomial

$$1 - 2z = 0$$

where $z = 0.5$ Since $z = 0.5$ is inside the unit circle, the process **isn't invertible**.

iv. (3 points): Identify $\{Y_t\}$ as ARMA(p, q).

Solution: Rewrite we could get:

$$\begin{aligned} Y_t &= 0.7Y_{t-1} - 0.3Y_{t-2} = W_t - 0.4W_{t-1} \\ \phi(B) &= 1 - 0.7B - 0.3B^2, \quad \theta(B) = 1 - 0.4B \\ &= -(0.3B + 1)(B - 1) \end{aligned}$$

There is no parameter redundancy and we can identify it as ARMA(2, 1) model.

v. (1 points): Is $\{Y_t\}$ causal? Justify your answer.

Solution: Evaluating the characteristic polynomial: $1 - 0.7z - 0.3z^2 = 0$

$$z_1 = 1, z_2 = -\frac{10}{3}$$

Since $z_1 = 1$ is on the unit circle, it's **not causal**.

vi. (1 points): Is $\{Y_t\}$ invertible? Justify your answer.

Solution: Evaluating the characteristic polynomial $1 - 0.4z = 0$ $z = 2.5$ which is outside the unit circle
 \therefore The process **is invertible**.

Problem 4 (7 points): Consider an AR(2) process with the equations:

$$P(B) = (1 - 0.4B)(1 + 0.4B). \quad (5)$$

Please answer the following questions:

i. (1 points): Is the process causal?

Solution: Evaluating

$$(1 - 0.4z)(1 + 0.4z) = 0$$

$$z_1 = 2.5 \quad z_2 = -2.5$$

All roots are outside the unit circle, therefore the process is causal.

ii. (6 points): What is the correlation function $\rho(t, t + h) = \rho(h)$?
Hint: remember that $\rho(0) = 1$.

Solution: We could read the model definition and get:

$$X_t - 0.16X_{t-2} = W_t$$

This can be rewritten as:

$$X_t = 0.16X_{t-2} + W_t$$

Note that $0.16 < 1$, thus there exists a unique solution. We solve it using the idea of a difference equation. first, we multiply each side of the model by $X_{t-h} \forall h > 0$

$$E[X_t X_{t-h}] = 0.16E[X_{t-2} X_{t-h}] + E[W_t X_{t-h}]$$

$\therefore X_{t-h} = \sum_{j=0}^{\infty} \psi_j W_{t-h-j}$ where ψ_j is some constant coefficient.

$$\therefore E \left[W_t \sum_{j=0}^{\infty} \psi_j W_{t-h-j} \right] = \sum_{j=0}^{\infty} \psi_j E[W_t W_{t-h-j}]$$

$$= 0 \quad \forall h > 0$$

$$\therefore \gamma(h) = 0.16\gamma(h-2), h = 1, 2, \dots$$

Dividing $r(0)$ on both sides we have

$$\rho(h) = 0.16\rho(h-2)$$

Let z_1, z_2 be the roots of associated homogeneous difference equation of order, $\phi(z) = 1 - 0.16z^2$

$$z_1 = 2.5 \quad z_2 = -2.5$$

$\therefore z_1, z_2$ are real and distinct, we have:

$$\rho(h) = c_1 z_1^{-h} + c_2 z_2^{-h}$$

Now we figure out the initial conditions:

$$\begin{cases} \rho(0) = 1 \\ \rho(1) = 0.16\rho(-1) \end{cases} \Rightarrow \begin{cases} \rho(0) = 1 \\ \rho(1) = 0 \end{cases} \text{ since } \rho(1) = \rho(-1)$$

We solve for c_1, c_2 :

$$\begin{cases} c_1 + c_2 = 1 \\ 0.4c_1 - 0.4c_2 = 0 \end{cases} \Rightarrow \begin{cases} c_1 = 0.5 \\ c_2 = 0.5 \end{cases}$$

$$\therefore \rho(h) = 0.5 \cdot 0.4^h + 0.5 \cdot (-0.4)^h, h = 0, 1, 2, \dots$$