732A62 Time Series Analysis
Division of Statistics and Machine Learning
Department of Computer and Information Science

Teaching session I

Instructions

The assignments in the first section will be solved by the teacher during the teaching session. You are welcome to ask questions if you cannot follow the derivations. The problems in the second section are take home exercises and the key is given in section 3. The following assignments are hand-in and no solution is given in the key.

Assignment 12
Assignment 18

The hand-in assignment should be solved individually and should be submitted via LISAM in pdf format before the deadline also specified in LISAM.
The solutions are graded pass / insufficient. An insufficient solution can be completed and resubmitted.

1. Assignments solved by the teacher

Assignment 1

Suppose $E(X) = 2$, $var(X) = 9$, $E(Y) = 0$, $var(Y) = 4$, and $corr(X,Y) = 0.25$. Find:
(a) $var(X + Y)$.
(b) $cov(X, X + Y)$.

Assignment 2

Suppose $y_t = 5 + 2t + x_t$, where $\{x_t\}$ is a zero-mean stationary series with autocovariance function γ_k.
(a) Find the mean function for $\{y_t\}$.
(b) Find the autocovariance function for $\{y_t\}$.
(c) Is $\{y_t\}$ stationary? Why or why not?

Assignment 3

Suppose that $\{x_t\}$ is stationary with autocovariance function γ_k. Show that for any fixed positive integer n and any constants $c_1, c_2,..., c_n$, the process $\{y_t\}$ defined by $y_t = \sum_{i=1}^{n} c_i x_{t-i+1}$ is stationary.

Assignment 4

Suppose that $x_t = w_t - w_{t-12}$. Show that $\{x_t\}$ is stationary and that, for $k > 0$, its autocorrelation function is nonzero only for lag $k = 12$.

Assignment 5

Suppose $x_t = \mu + w_t - w_{t-1}$. Find $var(\bar{x})$. Note any unusual results. In particular, compare your answer to what would have been obtained if $x_t = \mu + w_t$.

Assignment 6

Calculate and sketch the autocorrelation functions for AR(1) model with $\phi = 0.6$. Plot for sufficient lags that the autocorrelation function has nearly died out.

Assignment 7

Let $\{x_t\}$ be an AR(2) process $x_t = \phi x_{t-2} + w_t$ Find the range of values of ϕ for which the process is causal.

Assignment 8

For each of the following ARMA models, find the roots of the AR and MA polynomials, identify the values of p and q for which they are ARMA(p,q) (be careful of parameter redundancy), determine whether they are causal, and determine whether they are invertible. In each case, $w_t \sim wn(0,1)$.
.

a) $x_t + 0.81x_{t-2} = w_t + 1/3w_{t-1}$
b) $x_t - x_{t-1} = w_t - 0.5w_{t-1} - 0.5w_{t-2}$

Assignment 9

For those the following model, compute the first four coefficients $\psi_0, \ldots \psi_3$ in the causal linear process representation $x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j}$

a) $x_t + 0.81x_{t-2} = w_t + 1/3w_{t-1}$

2. Take home assignments

Assignment 10

Suppose $E(X) = 2$, $Var(X) = 9$, $E(Y) = 0$, $Var(Y) = 4$, and $Corr(X,Y) = 0.25$. Find:
(a) $Corr(X + Y, X - Y)$.

Assignment 11

Let $\{w_t\}$ be a zero mean white noise process. Suppose that the observed process is $x_t = w_t + \theta w_{t-1}$, where θ is either 3 or 1/3.

(a) Find the autocorrelation function for $\{x_t\}$ both when θ = 3 and when θ = 1/3.
(b) You should have discovered that the time series is stationary regardless of the value of θ and that the autocorrelation functions are the same for θ = 3 and θ =1/3. For simplicity, suppose that the process mean is known to be zero and the variance of y_t is known to be 1. You observe the series $\{y_t\}$ for $t = 1, 2,..., n$ and suppose that you can produce good estimates of the autocorrelations ρ_k. Do you think that you could determine which valueθ is correct (3 or 1/3) based on the estimate of ρ_k? Why or why not?

Assignment 12

Let $\{x_t\}$ be a zero-mean, unit-variance stationary process with autocorrelation function ρ_h. Suppose that μ_t is a nonconstant function and that σ_t is a positive-valued nonconstant function. The observed series is formed as $y_t = \mu_t + \sigma_t X_t$.
(a) Find the mean and covariance function for the $\{y_t\}$ process.
(b) Show that the autocorrelation function for the $\{y_t\}$ process depends only on the time lag. Is the $\{y_t\}$ process stationary?
(c) Is it possible to have a time series with a constant mean and with $Corr(y_t, y_{t+h})$ free of t but with $\{y_t\}$ *not* stationary?

Assignment 13

Suppose that x is a random variable with zero mean. Define a time series by
$y_t = (-1)^t x$
(a) Find the mean function for $\{y_t\}$.
(b) Find the autocovariance function for $\{y_t\}$.
(c) Is $\{y_t\}$ stationary?

Assignment 14

Suppose $x_t = \mu + w_t + w_{t-1}$. Find $var(\bar{x})$. Note any unusual results. In particular, compare your answer to what would have been obtained if $x_t = \mu + w_t$.

Assignment 15

Calculate and sketch the autocorrelation function for MA(2) model with $\theta_1 = 0.5$. and $\theta_2 = 0.4$

Assignment 16

Describe the important characteristics of the autocorrelation function for the following models: (a) MA(1), (b) MA(2), (c) AR(1), (d) AR(2), and (e) ARMA(1,1).

Assignment 17

Suppose that $\{x_t\}$ is an AR(1) process with $-1 < \varphi < +1$.
(a) Find the autocovariance function for $y_t = \nabla x_t = x_t - x_{t-1}$ in terms of φ and σ_w^2
(b) In particular, show that $var(y_t) = \frac{2\sigma_w^2}{1+\phi}$

Assignment 18

For each of the following ARMA models, find the roots of the AR and MA polynomials, identify the values of *p* and *q* for which they are ARMA(p,q) (be careful of parameter redundancy), determine whether they are causal, and determine whether they are invertible. In each case, $w_t \sim wn(0,1)$.
.

c) $x_t - 3x_{t-1} = w_t + 2w_{t-1} - 8w_{t-2}$
d) $x_t - 2x_{t-1} + 2x_{t-2} = w_t - \frac{8}{9}w_{t-1}$
e) $x_t - 4x_{t-2} = w_t - w_{t-1} + 0.5w_{t-2}$
f) $x_t - \frac{9}{4}x_{t-1} - \frac{9}{4}x_{t-2} = w_t$

Assignment 19

For the following models, compute the first four coefficients $\psi_0, \ldots \psi_3$ in the causal linear process representation $x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j}$

a) $x_t - 2x_{t-1} + 2x_{t-2} = w_t - \frac{8}{9}w_{t-1}$
b) $x_t - \frac{9}{4}x_{t-1} - \frac{9}{4}x_{t-2} = w_t - 3w_{t-1} + \frac{1}{9}w_{t-2} - \frac{1}{3}w_{t-3}$

3. Key

Assignment 10

Approximately 0.39

Assignment 11

$\rho(0) = 1, \rho(1) = 0.3, \rho(h) = 0\ otherwise$

Assignment 13

a) 0
b) $(-1)^h \sigma_x^2$
c) Yes

Assignment 14

$var(\bar{x}) = \frac{2(2n-1)}{n^2}\sigma_w^2$

Assignment 15

$\rho_1 \approx 0.5, \rho_2 \approx 0.28, \rho_i = 0, i > 2$

Assignment 16

a) $-\frac{1-\phi}{1+\phi}\phi^{h-1}\sigma_w^2$

Assignment 17

a) p=1,q=2, neither causal or invertible
b) p=2,q=1, invertible, but not causal
c) p=2,q=2, invertible, but not causal
d) p=2, q=0, invertible, not causal

Assignment 19

a)1, 10/9, 2/9, -16/9

b) 1, -3/4, 1/9+9/16, -1/12-27/64

Teaching session 2: Hand In assignment.

Assignment 1:

$$ARIMA(p,d,q) \times (P,D,Q)_S$$

using the backshift operator:

$$\phi^P(B^S)\,\phi^p(B)\,(1-B^S)^D\,(1-B)^d\,x_t = \theta^Q(B^S)\,\theta^q(B)\,w_t$$

you may use the following

$$p=3,\ d=2,\ q=1,\ P=2,\ D=1,\ Q=1,\ S=5$$

$$z_t = A z_{t-1} + e_t\,,$$

$$x_t = C z_t + v_t$$

e_t and v_t are noise that we want to get rid of.

x_t is the measure.

z_t is the latent variable that we want to find without the noise (i.e we want to remove the noise)

ARIMA models in state space form: $\phi^p(B)\,x_t = \theta^q(B)\,w_t$

let $r = \max(p, q+1)$, $= \max(3, 1+1) = 3$

$$\phi^r(B) = 1 - \phi_1 B - \dots - \phi_r B^r$$

$$\phi^3(B) = 1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3$$

$$\theta^r(B) = 1 + \theta_1 B - \dots - \theta_{r-1} B^{r-1}$$

$$\theta^3(B) = 1 + \theta_1 B - \theta_2 B^2$$

$$\phi^r(B)\,(\theta^r(B))^{-1}\,x_t = w_t$$

Hence, for $z_t = (\theta^r(B))^{-1} x_t$, we can have $\phi^r(B)\,z_t = w_t$

$$\phi^3(B)\,(\theta^3(B))^{-1}\,x_t = w_t$$

$$(\theta^3(B))^{-1}\,x_t = z_t\,, \qquad \phi^3(B)\,z_t = w_t$$

$$z_t = \begin{bmatrix} z_t \\ z_{t-1} \\ z_{t-2} \\ z_{t-3} \\ z_{t-4} \end{bmatrix} = \begin{bmatrix} z_t \\ z_{t-1} \\ \vdots \\ z_{t-r+1} \end{bmatrix}$$

and

$$z_t = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_r \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} z_{t-1} + \begin{bmatrix} w_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\boxed{\begin{aligned} z_t &= \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} z_{t-1} + \begin{bmatrix} w_t \\ 0 \\ 0 \\ 0 \end{bmatrix} \\ x_t &= \begin{bmatrix} 1 & \theta_1 & \theta_2 & \cdots & \theta_r \end{bmatrix} z_t \\ &= \begin{bmatrix} 1 & \theta_1 & \theta_2 & \theta_3 \end{bmatrix} z_t \end{aligned}}$$

$$\phi^p(B)\, x_t = \theta^q(B)\, w_t$$

$\cancel{z_t} =$

Time Series Analysis

Teaching Session - 1

Assignment 1:

(a) $\text{var}(x+y) = E\{(x+y-E(x+y))^2\}$

$= E\{(x+y-2)^2\} = E\{x^2+y^2+4+2xy-4x-4y\}$

$= E(x^2) + \underbrace{E(y^2)}_{\text{var}(y)} + 4 + 2E(xy) - \underbrace{4E(x)}_{8} - \underbrace{4E(y)}_{0}$

$\rightarrow E(x+y) = E(x) + E(y) = 2$

OR

$\text{var}(x+y) = E\{((x+y) - E(x+y))^2\}$

$= E\{((x-E(x)) + (y-E(y)))^2\}$

$(a+b)^2 = a^2 + 2ab + b^2$

$= E\{(x-E(x))^2 + (y-E(y))^2 + 2(x-E(x))(y-E(y))\}$

$= \text{var}(x) + \text{var}(y) + 2\,\text{cov}(x,y)$

Note:

$\text{corr}(x,y) = \dfrac{\text{cov}(x,y)}{\sqrt{\text{var}(x)\,\text{var}(y)}}$

~~∴ To find cov(x+y, x−y)~~

cov(x+y, x−y

ASSIGNMENT 10:

To compute corr $(x+y, x-y)$, we get cov.

$$\text{cov}(x+y, x-y) = E\{(x+y-E(x+y))\cdot(x-y-E(x-y))\}$$
$$= E\{(\underbrace{(x-E(x))}_{a} + \underbrace{(y-E(y))}_{b})\cdot$$
$$(\underbrace{(x-E(x))}_{a} - \underbrace{(y-E(y))}_{b})\}$$

$$(a+b)(a-b) = a^2 - b^2$$

ASSIGNMENT 2:

(a) $y_t = 5 + 2t + x_t$

$\hookrightarrow E(y_t) = 5 + 2t + E(x_t)$

$= 5 + 2t \qquad \rightarrow y_t - E(y_t) = x_t$

$E(x_t) = 0 \qquad \{x_t\}$ stationary $\qquad \gamma_x k$

(b) $\gamma_y(s,t) = \text{cov}(y_s, y_t) = E\{(y_s - E(y_s))(y_t - E(y_t))\}$

$= \cancel{E\{(x_s - E(x_s))(x_t - E(x_t))\}}$

$= E(x_s x_t) \Rightarrow E\{(x_s - E(x_s))(x_t - E(x_t))\}$

$= \text{cov}(x_s, x_t) \Rightarrow \gamma_x(t,s)$

(c) $\{y_t\}$ is not stationary inspite of having a decent $\gamma_y(t,s)$ due to its $2t$.

Assignment 5:

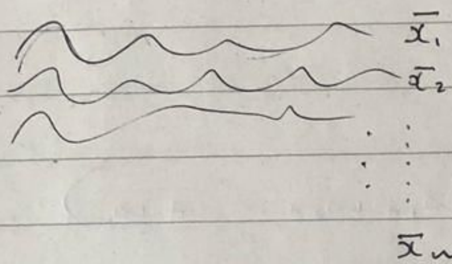
E of w_t & w_{t-1} will be 0

$$x_t = \mu + w_t - w_{t-1} \longrightarrow E[x_t] = \mu$$

Samples: $x_1, \ldots\ldots x_n$

Sample mean $\boxed{\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i}$

If you take infinitely many samples, what will be the variance of their average?



$$\bar{\bar{x}} = \frac{1}{m}\sum_{j=1}^{m}(\bar{x}_m) \Rightarrow \frac{1}{m}\sum_{j=1}^{m}\frac{1}{n}\sum_{i=1}^{n} x_i \Rightarrow \frac{1}{mn}\sum_{j=1}^{m}\sum_{i=1}^{n} x_i$$

$$E(\bar{x}) = E\left\{\frac{1}{n}\sum_{i=1}^{n} x_i\right\} \Rightarrow \frac{1}{n}\sum_{i=1}^{n} E x_i$$

$$= \frac{1}{n}\sum_{i=1}^{n}\mu \Rightarrow \frac{1}{n} n\mu \Rightarrow \mu$$

$$\text{var}(\bar{x}) = E\left\{(\bar{x} - E(\bar{x}))^2\right\} = E\left\{(\bar{x} - \mu)^2\right\}$$

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n)$$

$$= \frac{1}{n}\Big(\mu + \cancel{w_1} - w_0$$
$$+ \mu + \cancel{w_2} - \cancel{w_1}$$
$$+ \mu + \cancel{w_3} - \cancel{w_2} \qquad = \frac{1}{n}(n\mu + w_n - w_0)$$
$$\vdots$$
$$+ \mu + w_n - \cancel{w_{n-1}}\Big) \qquad = \mu + \frac{w_n - w_0}{n}$$

Hence, $\bar{x} - \mu = \dfrac{w_n - w_0}{n}$

$$\therefore \ var(\bar{x}) = E\left\{\left(\frac{w_n - w_0}{n}\right)^2\right\}$$

$$= \frac{1}{n^2} E\left\{ w_n^2 + w_0^2 - \underbrace{2 w_n w_0}_{\text{because white noise}} \right\}$$

$$= \frac{1}{n^2}(2\sigma_w^2) \Rightarrow \frac{2\sigma_w^2}{n^2}$$

If $x_t = \mu + w_t$,

$$\bar{x} = \frac{1}{n}(\mu + w_1 + \mu + w_2 + \cdots + \mu + w_n)$$

$$= \mu + \frac{1}{n}\sum_{i=1}^{n} w_i$$

$$\bar{x} - \mu = \frac{1}{n}\sum_{i=1}^{n} w_i \Rightarrow \frac{1}{n}(w_1 + w_2 + \cdots + w_n)$$

$$var(\bar{x}) = E\left\{\left(\frac{1}{n}(w_1 + w_2 + \cdots + w_n)\right)^2\right\}$$

$$= \frac{1}{n^2} E\left(\underbrace{(w_1 + w_2 + \cdots + w_n)^2}\right)$$

$$= \frac{1}{n^2} var(w_1 + w_2 + \cdots + w_n)$$

$$(w_1^2 + w_2^2 + \cdots + w_n^2 + \underbrace{2w_1w_2 + 2w_1w_3 + \cdots}\,)$$

The expected value of all cross terms are 0.

$$\therefore \ = \frac{1}{n^2}\cdot n\sigma_w^2 \Rightarrow \frac{\sigma_w^2}{n}$$

Assignment 7:

$$x_t = \phi x_{t-2} + w_t$$

$$x_t = \sum_{j=-\infty}^{\infty} \psi_j w_{t-j} \leftarrow \text{From causality}$$

$$x_{t-2} = \sum_{j=-\infty}^{\infty} \psi_j w_{t-2-j}$$

Let $\underline{i = 2+j}$, then

$$x_{t-2} = \sum_{i=-\infty}^{\infty} \psi_{i-2}$$

$x_t = \phi x_{t-2} + w_t$ ← Backshift operator

$x_t - \phi x_{t-2} = w_t$

$\underbrace{(1-\phi B^2)}_{\Phi B} x_t = w_t$ from ARMA causality, $\phi(B)x_t = \theta(B)w_t$

Roots of $\Phi(z) = \pm\sqrt{\frac{1}{\phi}}$ and must be outside the unit circle for causality.

$$\Rightarrow \left|\frac{1}{\phi}\right| > 1 \quad \Rightarrow |\phi| < 1$$

For example: $\phi = +0.9$ or $\phi = -0.8$ etc.

OR $\Rightarrow \underbrace{\sum_{j=-\infty}^{\infty} \psi_j w_{t-j}}_{x_t} - \sum_{i=-\infty}^{\infty} \phi \underbrace{\psi_{i-2} w_{t-i}}_{x_{t-2}} - w_t = 0$ → because of $x_t = \phi x_{t-2} + w_t$ Taking ϕx_{t-2} & w_t to LHS

$$\sum_{j=-\infty}^{\infty} (\psi_j - \phi\psi_{j-2}) w_{t-j} - w_t = 0$$

☆ Anything multiplied to w_t should be 0.

At $j=0$, $\psi_0 - \phi\psi_{-2} - 1 = 0$

Otherwise, $\psi_j - \phi\psi_{j-2} = 0$ for all $j \neq 0$.

$\psi_j = 0$ for $j<0$

$$x_t = \sum_{j=-\infty}^{\infty} \psi_j w_{t-j} \quad \text{where} \sum_{j=0}^{\infty} |\psi_j| < \infty$$

for non-exclusiveness

03-09-19

ADVANCED MACHINE LEARNING

LECTURE 1

- Probabilistic graphical models

<u>Slide 4</u>.

- Arrows ⟵ Cause to effect
[Causal relationships]
Causal model.

- $$\begin{tikzcd} S \arrow[dr] & & \\ & WG \arrow[r, leftrightarrow] & WS \end{tikzcd}$$
← Intermediary
~~If not modeled~~
there may be something in between which leads them to be bi-related.

- Causal model for Temperature, Icecream sales & Soda sales:

$$\begin{tikzcd} & T \arrow[dl] \arrow[dr] & \\ ICS & & SS \end{tikzcd}$$

- For Boyle's law,
Pressure × Volume = Constant

$$\begin{tikzcd} P \arrow[r, bend left] & V \arrow[l, bend left] \end{tikzcd}$$
If P ↑s, then V should ↓s.

$P — V$ ⟸ P & V are correlated

↑
undirected edge

- Bayesian Networks:
- DAG represents causal relationships.
- A ~~no~~ binary no. represents the strength of these relationships. [Binary random var.]
-

$r_0 \leftarrow$ no rain , $r_1 \leftarrow$ rain

$s_0 \leftarrow$ sprinkler off , $s_1 \leftarrow$ sprinkler on

- $q(S) = (0.3, 0.7)$

 ↳ if sprinkler on, ~~wet~~ grass is wet by prob. 30%
 & street is wet with prob. 70%

- Sprinkler $\leftarrow$ Normal var.
 Rain $\leftarrow$ Gaussian var.
 Wet Street $\leftarrow$ Gaussian rand. var. whose mean depends on Rain

- $p(S, R, WG, WS)$ = Product of all [no. btw. 0 & 1]
 ↳ joint distr. of whole domain : $q(S)\,q(R)\,q(WG \mid S, R)\,q(WS \mid R)$

for,

$P(s_0, r_1, ws_0, wg_1) = q(s_0)\,q(r_1)\,q(ws_0 \mid r_1)\,q(wg_1 \mid s_0, r_1)$

$\boxed{2^{4} - 1 = 15 \text{ parameters}}$

- $pa_G \leftarrow$ parent [incoming variable]

- Sum of all = 1

- Ex:-

$$A \rightarrow B \rightarrow C$$

$$\sum_{ABC} P(ABC) = \sum_{ABC}\left[q(A) \cdot q(B|A) \cdot q(C|B)\right]$$

$$= \sum_{A} \underbrace{q(A)}_{1} \left[\sum_{B} \left[\underbrace{q(B|A)}_{1} \left[\sum_{C} \underbrace{q(C|B)}_{\substack{1 \\ \text{because it's a} \\ \text{distribution}}} \right]\right]\right]$$

Do the sums in order from the last to the first.

$$\sum_{ABC} P(ABC) = \sum_{AB} \sum_{C} \underbrace{q(A) \cdot q(B|A)}_{\text{These 2 don't depend on C, so move them out}} \underbrace{\cancel{q(C|B)}}_{1}$$

- $P(X_{1:n}) = \overbrace{\prod_{i=1}^{n} q(\underbrace{X_i}_{\text{var.}} \mid \underbrace{Pa(X_i)}_{\text{Parents of the var.}})}^{\text{Local distribution}}$

$$\underbrace{P(X_i \mid Pa(X_i))}_{\text{Cond. distr.}} = q(X_i \mid Pa(X_i))$$

Ex:-

$$q(\underline{WG \mid S, R}) = P(\underline{WG \mid S, R})$$

- <u>Bayesian Networks : Separation</u> :
- P can be computed from the graph without numerical calc.

$$\begin{tikzcd}
\text{Age} \arrow[r] & \text{Battery} \arrow[dl] \arrow[dr] & & \text{Bulb} \arrow[dl] \\
\text{Radio} & & \text{Lights} &
\end{tikzcd}$$

- Is the distr. of Age independent of Bulb?

  $P(A, \text{Bulb}) = P(A) \cdot P(\text{Bulb})$

- $P(R, \text{Bulb} \mid \text{Battery}) = P(R \mid \text{Battery}) \cdot P(B \mid \text{Battery})$

CHAIN

- Age $\rightarrow$ Battery $\rightarrow$ Radio

  $\Rightarrow$ Age $\perp$ Radio | Battery

  independent

  $\not\perp$ ← not independent

  marginal

  $\Rightarrow$ Age $\not\perp$ Radio | $\emptyset$ [if something is dependent, we will have to look into numbers]

  ↳ the guy in the middle blocks Age from Radio.

FORK

- Radio $\leftarrow$ Battery $\leftarrow$ Lights

COLLIDER

- Battery $\rightarrow$ Lights $\leftarrow$ Bulb

  ↳ battery & bulb are marginally independent

  If lights don't work & battery works, bulb will not work.

  ↳ Battery $\perp$ Bulb | $\emptyset$

- Age $\rightarrow$ Battery $\rightarrow$ Lights $\leftarrow$ Bulb

  Age & bulb are independent marginally.

  Bulb has no dependence on age.

  $\therefore P(\text{Age} \mid \cancel{\text{Bulb}}, \text{Light}, \text{Battery}) = P(\text{Age} \mid \text{Light}, \text{Battery})$

- Given, $A \longrightarrow B$
$P(A)$?
$P(B|A)$?
On estimating,
$P(A)$ = uniform
$P(B|A)$ = uniform

- Separation $\Leftarrow$ to read independences.

- In a path, *direction of arrow is irrelevant.

* Z is open if no non-collider in path ρ.
Non-colliders = Age, battery, bulb

If path is closed, then INDEPENDENT.

- Age and bulb are separated because the path is not open between the 2. $\leftarrow$ independence in joint distr.

- How many free parameters in wet ~~gra~~ grass Bayesian N/w?
When using BN $\Rightarrow$ 8
On computing P,
it will be $= 2^n - 1$
$= 2^4 - 1 \Rightarrow 15$

Independence reduces no. of parameters.

- **Causal reasoning:**
  - What is the distr. of a var. when it is forced to take a particular state?
  - Rain $\longrightarrow$ Wet Street

But if ~~forced~~ street is forced to be wet by throwing water

$$P(R|WS) \neq P(R) \leftarrow \text{without force}$$

$$P(R|\underline{do(WS)}) = P(R)$$

forcing street to be wet [forcing rain to happen by throwing water or something]

On forcing, take the model & remove Rain from the model.
Remove all tables with r_0.
Rain will definitely happen (r_1) as it is being forced.

do $\leftarrow$ intervention operator.

- **Probabilistic Reasoning:**
  - Marginalizing by Σ's.

- **Markov Networks:**
  - Graph is undirected.

Time Series Analysis - Teaching assignment 1

Assignment 12:

Let $\{x_t\}$ be a zero mean unit variance stationary process with ACF $= \rho_h$

$$y_t = \mu_t + \sigma_t x_t$$

a) Find mean and covariance function of $\{y_t\}$ process

$$E(y_t) = E(\mu_t + \sigma_t x_t)$$

$\because$ μ_t and y_t are non-stochastic

$$\mu_t + \sigma_t E(x_t) = \mu_t$$

0

$$\text{cov}(y_s, y_t) = E\left[(y_s - E(y_s))(y_t - E(y_t))\right]$$

$$y_t - E(y_t) = \cancel{\mu_t} + \sigma_t x_t - \cancel{\mu_t} = \sigma_t x_t$$

$$y_s - E(y_s) = \mu_s - \sigma_s x_s - \mu_s = \sigma_s x_s$$

$$E(\sigma_s \sigma_t x_s x_t) = \sigma_s \sigma_t E(x_s x_t) = \sigma_s \sigma_t E\left[(x_s - E(x_s))(x_t - E(x_t))\right]$$

$$= \sigma_s \sigma_t \gamma_x(s,t) = \sigma_s \sigma_t \text{cov}(x_s, x_t)$$

b) Show that the ACF for the $\{y_t\}$ only on time lag.

Is $\{y_t\}$ stationary?

$$\rho(s,t) = \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\,\gamma(t,t)}}$$

$$\rho_y(s,t) = \frac{\gamma(y_s, y_t)}{\sqrt{\gamma(s,s)\,\gamma(t,t)}} = \gamma(s,t) = \sigma_s \sigma_t \text{cov}(x_t, x_s)$$

$$= \sigma_s \sigma_t \gamma(s-t)$$

$$\left.\begin{array}{l} E(y_t) = \text{constant} \\ \gamma(s,t) = \gamma(s-t) \\ \text{var}(y_t) < \infty \end{array}\right\} \Rightarrow \text{stationary}$$

c) Yes, It is possible where $\text{var}(y_t) = \infty$

Assignment 18

For each of ARMA models, find the roots of the AR and MA polynomials, identify the values of p and q for which they are ARMA(p, q) (be careful of parameter redundancy), determine whether they are causal, and determine whether they are invertible. In each case, $w_t \sim wn(0,1)$

c) $x_t - 3x_{t-1} = w_t + 2w_{t-1} - 8w_{t-2}$

$1 - 3B = 0$

$3B = 1$

$B = 1/3$

$B < 1$, $\therefore$ not causal

$1 + 2B - 8B^2 = 0$

$8B^2 - 2B - 1 = 0$

$D = 4 - 4(-1) \cdot 8 = 36$

$B = \left(\frac{2+6}{16}, \frac{2-6}{16}\right) = \left(\frac{1}{2}, \frac{-1}{4}\right)$

$B < 1$, $\therefore$ not invertible

d) $x_t - 2x_{t-1} + 2x_{t-2} = w_t - \frac{8}{9} w_{t-1}$

$1 - 2B + 2B^2 = 0$

$2B^2 - 2B + 1 = 0$

$D = 4 - 4 \cdot 2 \cdot 1 = -4$

$B = (0.5 + 0.5i, 0.5 - 0.5i)$

$\therefore$ not causal

$1 - \frac{8}{9}B = 0$

$\frac{8}{9}B = 1 \Rightarrow B = \frac{9}{8} > 1$

$\therefore$ invertible

f) There are no roots.

$\therefore$ Invertible

$1 - 9/4\,B - 9/4\,B^2 = 0$

$B = (-1.33, 0.33)$. Non causal

e) $x_t - 4x_{t-2} = w_t - w_{t-1} + 0.5w_{t-2}$

$1 - 4B^2 = 0$

$(1 - 2B)(1 + 2B) = 0$

$B = (1/2, -1/2)$

$\therefore$ not causal

$1 - B + 0.5B^2 = 0$

$0.5B^2 - B + 1 = 0$

$D = 1 - 4 \cdot 0.5 \cdot 1 = 1 - 2 = -1$

$B = (1 + i, 1 - i)$

$\therefore$ non-invertible

Teaching session 3 : Hand In assignment.

Assignment 1:

Prove the Kalman filtering recursion for the following state space model with initial prior on state $f(z_1) = N(z; m_0, P_0)$

where: $e_t \sim N(0, Q_t)$

$v_t \sim N(0, R_t)$

$$z_t = A_{t-1} z_{t-1} + e_t \quad (1)$$

$$x_t = C_t z_t + v_t \quad (2)$$

Property 1: $f(y_1) f(y_2 | y_1) = f(y_1, y_2)$

$$N(y_1; u_1, \Sigma) \; N(y_2; By_1, R) = N\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}; \begin{bmatrix} u \\ Bu \end{bmatrix} \begin{bmatrix} \Sigma & \Sigma B^T \\ B\Sigma & B\Sigma B^T + R \end{bmatrix}\right)$$

Property 2: Marginalisation and conditioning

$$f(y_1, y_2) = N\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}; \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right)$$

then: $f(y_1) = N(y_1; u_1, \Sigma_{11})$

$f(y_2) = N(y_2; u_2, \Sigma_{22})$

$f(y_1 | y_2) = N(y_1; u_1 + \Sigma_{12} \Sigma_{22}^{-1} (y_2 - u_2), \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21})$

$f(y_2 | y_1) = N(y_2; u_2 + \Sigma_{21} \Sigma_{11}^{-1} (y_1 - u_1), \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12})$

Assume that the posterior distribution of previous step is gaussian.

$$p(z_{t-1} | x_{1:t-1}) = N(z_{t-1} | m_{t-1}, P_{t-1})$$

Chapman - Kolmogorov eq:

$$p(z_t | x_{1:t-1}) = \int p(z_t | z_{t-1}) \, p(z_{t-1} | x_{1:t-1}) \, dx_{t-1}$$

$$= \int N(z_t | A_{t-1} z_{t-1}, Q_{t-1}) \, N(z_{t-1} | m_{t-1}, P_{t-1})$$

Using gaussian distribution computation rules from previous slides, we get the prediction step.

$$p(z_{t+1} \mid x_{1:t}) = \mathcal{N}(z_{t} \mid A_{t-1} m_{t-1}, A_{t-1} P_{t-1} A_{t-1}^T + Q_{t})$$

$$= \mathcal{N}(z_{t} \mid m_{t|t-1}, P_{t|t-1})$$

The joint distribution is given by x_t and z_t.

$$p(z_{t}, x_{t} \mid x_{1:t-1}) = p(x_{t} \mid z_{t})\, p(z_t \mid x_{1:t-1})$$

$$= \mathcal{N}\left(\begin{bmatrix} z_{t} \\ x_{t} \end{bmatrix} \middle| \begin{pmatrix} m_{t|t-1} \\ C m_{t|t-1} \end{pmatrix}, \begin{pmatrix} P_{t|t-1} & P_{t|t-1} C^T \\ C P_{t|t-1} & C P_{t|t-1} C^T + R \end{pmatrix}\right)$$

The conditional distribution of z_t and x_t is given as

$$p(z_t \mid x_t, y_{1:t-1}) = p(z_t \mid x_{1:t})$$

$$= \mathcal{N}(z_t \mid m_{t|t}, P_{t|t})$$

where

$$m_{t|t} = m_{t|t-1} + K_t (x_t - C_t m_{t|t-1})$$

$$P_{t|t} = (I - K_t C_t) P_{t|t-1}$$

$$K_t = P_{t|t-1} C_t^T (C_t P_{t|t-1} C_t^T + R)^{-1}$$

The state space model is given by,

$$z_t = A z_{t-1} + e_t \quad , e_t \sim N(0,R)$$
$$x_t = C z_t + v_t \quad ; v_t \sim N(0,Q)$$

To find the state space representation for ARIMA $(p,d,q) \times (P,D,Q)_s$,

$$\Phi^P(B^s)\phi^p(B)(1-B^s)^D(1-B)^d x_t = \Theta^Q(B^s)\theta^q(B) w_t$$
$$\Phi^P(B^s)\phi^p(B)(1-B^s)^D(1-B)^d \underbrace{\left[\Theta^Q(B^s)\theta^q(B)\right]^{-1} x_t}_{z_t} = w_t$$

$$\therefore \Phi^P(B^s)\phi^p(B)(1-B^s)^D(1-B)^d z_t = w_t \quad \leftarrow AR$$

Matrix A will be,

$$A = \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 & \cdots & \phi_r \\ 1 & 0 & \cdots & \cdots & 0 \\ \vdots & 1 & \cdots & \cdots & \vdots \\ \vdots & & \cdots & \cdots & \vdots \\ 0 & 0 & \cdots & \cdots & 1 \end{bmatrix}$$

where $r = \max(P + p + d + Ds, sQ + q + 1)$

$\therefore$ The AR equation can be written as,

$$\begin{bmatrix} z_t \\ z_{t-1} \\ \vdots \\ z_{t-r+1} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_r \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \cdots & \cdots & 1 \end{bmatrix} \begin{bmatrix} z_{t-1} \\ \vdots \\ \vdots \\ z_{t-r} \end{bmatrix} + \begin{bmatrix} w_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Matrix C will be,

$$C = [1 \;\; \theta_1 \;\; \theta_2 \;\; \cdots \;\; \theta_{r-1}]$$

~~where R = max(p, q+1)~~

$\therefore$ The MA equation can be written as,

$$x_t = [1 \;\; \theta_1 \;\; \theta_2 \;\; \cdots \;\; \theta_{r-1}] \begin{bmatrix} z_t \\ z_{t-1} \\ \vdots \\ z_{t-r+1} \end{bmatrix}$$

On substituting,

$p=3, \ d=2, \ q=1, \ P=2, \ D=1, \ Q=1, \ s=5$

$$r = \max(3+2+2+1+5, \ 5+2+1+1)$$
$$= \max(13, 9)$$
$$= 13$$

$$\begin{bmatrix} Z_t \\ Z_{t-1} \\ \vdots \\ Z_{t-14} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{13} \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 1 \end{bmatrix} \begin{bmatrix} Z_{t-1} \\ Z_{t-2} \\ \vdots \\ Z_{t-13} \end{bmatrix} + \begin{bmatrix} w_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

And,

$$x_t = \begin{bmatrix} 1 & \theta_1 & \theta_2 & \cdots & \theta_{12} \end{bmatrix} \begin{bmatrix} Z_t \\ Z_{t-1} \\ \vdots \\ Z_{t-14} \end{bmatrix}$$