/

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.

$$\mathrm{cov}(x+y, x-y) = E\{(x+y-E(x+y)) \cdot (x-y-E(x-y))\}$$

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

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

Assignment 2:

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

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

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

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

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

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

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

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

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

Assignment 5:

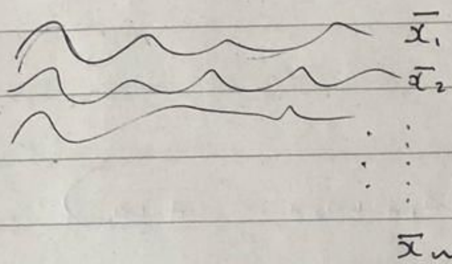
$\xrightarrow{}$ E of w_t & w_{t-1} will be 0

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

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?



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

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

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

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

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

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

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

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

$$= \frac{1}{n^2} E\{ \underbrace{w_n^2 + w_0^2 - 2w_nw_0}_{} \}$$

$\{$ because white noise

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

If $x_t = \mu + w_t$,

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

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

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

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

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

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

$$\rightarrow (w_1^2 + w_2^2 + \dots + w_n^2 + \underbrace{2w_1w_2 + 2w_1w_3 + \dots}_{})$$

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:

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

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

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

Let $i = 2+j$, then

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

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

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

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

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

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

☆ Anything multiplied to ω_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 $j < 0$

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

for non-exclusiveness

03-09-19

ADVANCED MACHINE LEARNING

LECTURE 1

- Probabilistic graphical models

Slide 4.

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

-
$$\begin{tikzcd} S \arrow[dr] & & \\ & WG \arrow[r, leftrightarrow, "V"] & WS \end{tikzcd}$$
← 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 ← Normal var.
  Rain ← Gaussian var.
  Wet Street ← 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:-

$$A \rightarrow B \rightarrow C$$

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

$$= \sum_{A} \underbrace{q(A)}_{1} \left[\sum_{B} \left[\underbrace{q(B|A)}_{1} \left[\sum_{C} \underbrace{q(C|B)}_{1} \right] \right] \right]$$

because it's a distribution

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

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

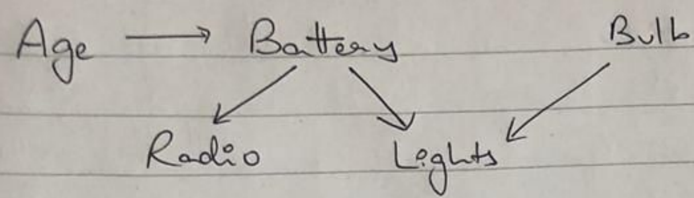
- $P(x_{1:n}) = \prod_{i=1}^{n} \overbrace{q(\underbrace{x_i}_{\text{var.}} | \underbrace{Pa(x_i)}_{\text{Parents of the var.}})}^{\text{Local distribution}}$

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

Ex:-

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

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



- Is the distr. of Age independent of Bulb ?

  $P(A, Bulb) = P(A) . P(Bulb)$

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

CHAIN

- Age → Battery → Radio

  ⇒ Age $\perp$ Radio | Battery $\not\perp$ ← not independent

  independent

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

  ↳ the guy in the middle blocks

  Age from Radio.

FORK

- Radio ← Battery ← Lights

COLLIDER

- Battery → Lights ← Bulb

  ↳ battery & bulb are marginally independent

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

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

- Age → Battery → Lights ← Bulb

  Age & bulb are independent marginally.

  Bulb has no dependence on age.

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

- 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 ⟶ 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 ← intervention operator.

- Probabilistic Reasoning:
  - Marginalizing by Σ's.

- Markov Networks:
  - Graph is undirected.