Math4826 Lecture Note 4 Appendix

Heng Peng

March 18, 2020

P4 Strictly Stationary process, for any m,

$$p(z_{t_1}, \ldots, z_{t_m}) = p(z_{t_1+k}, \ldots, z_{t_m+k})$$

Weakly Stationary process

$$\mathrm{E}z_t = \mu, \mathrm{Cov}(z_t, z_{t+k}) = \gamma_k, k = 0, 1, \ldots.$$

- Example: white noise process $\{a_t, t = 0, \pm 1, \pm 2, \ldots\}$ are a sequence of uncorrelated variable from fixed distribution with mean $\mathrm{E}(a_t) = 0, \mathrm{Var}(a_t) = \sigma^2$ and $\mathrm{Cov}(a_t, a_{t-k}) = 0$ for all $k \neq 0$.

P8.

$$\begin{aligned}\gamma_0 &= \mathrm{E}(a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots)^2 \\ &= \mathrm{E}a_t^2 + \psi_1^2 \mathrm{E}a_{t-1}^2 + \psi_2^2 \mathrm{E}a_{t-2}^2 + \cdots \\ &= \sigma^2 \sum_{j=0}^{\infty} \psi_j^2\end{aligned}$$

$$\begin{aligned}\gamma_k &= \mathrm{E}(z_t - \mu)(z_{t-k} - \mu) \\ &= \mathrm{E}(a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots)(a_{t-k} + \psi_1 a_{t-k-1} + \cdots) \\ &= 1 \cdot \psi_k \mathrm{E}a_{t-k}^2 + \psi_1 \psi_{k+1} \mathrm{E}a_{t-k-1}^2 + \psi_2 \psi_{k+2} \mathrm{E}a_{t-k-2}^2 + \cdots \\ &= \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}\end{aligned}$$

- Example $\psi_1 = -\theta$ and $\psi_j = 0, j \geq 2$, then $z_t - \mu = a_t - \theta a_{t-1}$, the first order moving average process.
- Example $\psi_j = \phi^j$, the first-order autoregressive process

$$\begin{aligned} z_t - \mu &= a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \cdots \\ &= a_t + \phi(a_{t-1} + \phi a_{t-2} + \phi^2 a_{t-3} + \cdots) \\ &= \phi(z_{t-1} - \mu) + a_t \end{aligned}$$

P9. For AR(1), it is important that $|\phi| < 1$ (Stationary Condition), since otherwise the ψ weight would not converge.

P11.

$$1 - \phi_1 B - \phi_2 B^2 = (1 - G_1 B)(1 - G_2 B) = 0,$$

G_1^{-1} and G_2^{-1} are its roots. For the stationarity, it requires that the roots are such that $|G_1^{-1}| > 1$ and $G_2^{-1}| > 1$.

- Example: $\phi_1 = 0.8, \phi_2 = -0.15$
  The solution of
  $$(1 - 0.8B + 0.15B^2) = (1 - 0.5B)(1 - 0.3B) = 0$$
  is given by $G_1^{-1} = 1/0.5 = 2$ and $G_2^{-1} = 10/3$, which are both larger than 1 in absolute value. Hence the process is stationary.
- Example: $\phi_1 = 1.5, \phi_2 = -0.5$
  The solution of
  $$(1 - 1.5B + 0.5B^2) = (1 - B)(1 - 0.5B) = 0$$
  is given by $G_1^{-1} = 1$ and $G_2^{-1} = 2$, which has one root at 1. Hence the process is not stationary.
- Example: $\phi_1 = 1, \phi_2 = -0.5$
  The solution of
  $$(1 - B + 0.5B^2) = 0$$
  are complex and given by $G_1^{-1} = 1 + i$ and $G_2^{-1} = 1 - i$, $|G_1^{-1}| = |G_2^{-1}| = \sqrt{1+1} = \sqrt{2}$ are both larger than 1. Hence the process is stationary.

P14. For $k = 1$

$$\rho_1 = \phi_1 \rho_0 + \phi_2 \rho_{-1} = \phi_1 + \phi_2 \rho_1 = \frac{phi_1}{1 - \phi_2}$$

For $k = 2$

$$\rho_2 = \phi_1 \rho_1 + \phi_2 \rho_0 = \frac{\phi_1^2}{1 - \phi_2} + \phi_2$$

- By $\gamma_0(1 - \phi_1 \rho_1 - \phi_2 \rho_2) = \sigma^2$, then

$$\gamma_0 = \frac{1 - \phi_2}{1 + \phi_2} \frac{\sigma^2}{(1 - \phi_2)^2 - \phi_1^2}$$

P16. $\boldsymbol{\rho} = (\rho_1, \rho_2, \ldots, \rho_p)^T, \boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_p)^T$

- For the AR(1) model, the Yule-Walker equation is given by $\rho_1 = \phi$.
- For the AR(2) model the Yule-Walker equation are

$$\rho_1 = \phi_1 + \rho_1 \phi_2$$

$$\rho_2 = \rho_1 \phi_1 + \phi_2$$

which leads to

$$\phi_1 = \frac{\rho_1(1 - \rho_2)}{1 - \rho_1^2}, \phi_2 = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

- $\phi(B)\rho_k = 0, k = 1, 2, \ldots,$ determines the behavior of the authorcorrelation function. It can be shown that its solution is

$$\rho_k = A_1 G_1^k + \cdots + A_p G_p^k, k = 1, 2, \ldots$$

where $G_i^{-1}, i = 1, \ldots, p$ are the distinct roots of $\phi(B) = (1 - G_1 B)(1 - G_2 B) \cdots (1 - G_p B)$.
- The stationary condition imply that $|G_i^{-1}| > 1, i = 1, \ldots, p$. Hence the ACF is described by a mixture of damped exponential (for real roots) and damped sine waves(for complex roots).

P17. Alternative interpretation of Partial Autocorrelations

The partial correlation coefficient between two random variables X and Y, conditional on a third variable W, is the ordinary correlation coefficient calculated from the conditional distribution $p(x, y|w)$. It can be thought of as the correlation between $X - \mathrm{E}(X|W)$ and $Y - \mathrm{E}(Y|W)$, and the assumption of joint normality of (X, Y, W) is given by

$$\begin{aligned}\rho_{XY\cdot W} &= \frac{\mathrm{E}(X - \mathrm{E}(X|W))(Y - \mathrm{E}(Y|W))}{\{\mathrm{E}(X - \mathrm{E}(X|W)^2\mathrm{E}(Y - \mathrm{E}(Y|W)^2\}^{\frac{1}{2}}} \\ &= \frac{\rho_{XY} - \rho_{XW}\rho_{YW}}{[(1 - \rho_{XW}^2)(1 - \rho_{YW}^2]^{\frac{1}{2}}}\end{aligned}$$

- In the context of a lag 2 partial autocorrelation, the variables are $X = z_t$, $Y = z_{t-2}$, $W = z_{t-1}$, and $\rho_{XY} = \rho_2, \rho_{XW} = \rho_{YW} = \rho_1$. Hence

$$\begin{aligned}\rho_{22} &= \rho_{z_t z_{t-2}\cdot z_{t-1}} = \mathrm{Corr}(z_t - \mathrm{E}(z_t|z_{t-1}), z_{t-2} - \mathrm{E}(z_{t-2}|z_{t-1})) \\ &= \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}\end{aligned}$$

P23. Example: Second-Order Moving Average Process[MA(2)]

$$z_t - \mu = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \quad \text{or} \quad z_t - \mu = (1 - \theta_1 B - \theta_2 B^2) a_t$$

- Hence $\psi_0 = 1, \psi_1 = -\theta_1, \psi_2 = -\theta_2, \psi = 0; j > 2$. Then

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2)\sigma^2, \gamma_1 = (-\theta_1 + \theta_1\theta_2)\sigma^2, \gamma_2 = -\theta_2\sigma^2$$

$$\gamma_k = 0, \text{for } k > 2$$

- The autocorrelations are

$$\rho_1 = \frac{-\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2}, \rho_2 = \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}, \rho_k = 0, \text{ for } k > 2$$

- If MA(2) process is in term of an infinite autoregressive rerpesentation,

$$z_t - \mu = \pi_1(z_{t-1} - \mu) + \pi_2(z_{t-2} - \mu) + \cdots + a_t,$$

then the π weight can be obtained from

$$\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \cdots = (1 - \theta_1 B - \theta_2 B^2)^{-1}$$

and

$$(1 - \pi_1 B - \pi_2 B^2 - \cdots)(1 - \theta_1 B - \theta_2 B^2) = 1$$

and are given by

$$B^1 : -\pi_1 - \theta_1 = 0,\ \pi_1 = -\theta_1$$

$$B^2 : -\pi_2 + \theta_1\pi_1 - \theta_2 = 0,\ \pi_2 = \theta_1\pi_1 - \theta_2 = -\theta_1^2 - \theta_2$$

$$B^j : -\pi_j + \theta_1\pi_{j-1} + \theta_2\pi_{j-2} = 0,\ \pi_j = \theta_1\pi_{j-1} + \theta_2\pi_{j-1}, j > 2.$$

- MA(2) continuous. For invertibility, the π weight is required converged, which in turn implies conditions on the parmater θ_1 and θ_2, the roots of $(1-\theta_1 B - \theta_2 B^2 = 0 = (1-H_1 B)(1-H_2 B)$ lie outside the unit circle. Hence

$$\theta_1+\theta_2<1, \theta_2-\theta_1<1, -1<\theta_2<1.$$

- Partial autocorrelation function for MA(1)

$$\phi_{11}=\rho_1=\frac{-\theta}{1+\theta^2}=\frac{-\theta(1-\theta^2)}{1-\theta^4}$$

$$\phi_{22}=\frac{\rho_2-\rho_1^2}{1-\rho_1^2}=\frac{-\rho_1^2}{1-\rho_1^2}=\frac{-\theta^2}{1+\theta^2+\theta^4}=\frac{-\theta^2(1-\theta^2)}{1-\theta^6}$$

$$\phi_{33}=\frac{\begin{vmatrix}1&\rho_1&\rho_1\\\rho_1&1&\rho_2\\\rho_2&\rho_1&\rho_3\end{vmatrix}}{\begin{vmatrix}1&\rho_1&\rho_2\\\rho_1&1&\rho_1\\\rho_2&\rho_1&\rho_1\end{vmatrix}}=\frac{\begin{vmatrix}1&\rho_1&\rho_1\\\rho_1&1&0\\0&\rho_1&0\end{vmatrix}}{\begin{vmatrix}1&\rho_1&0\\\rho_1&1&\rho_1\\0&\rho_1&\rho_1\end{vmatrix}}=\frac{\rho_1^3}{1-2\rho_1^2}=\frac{-\theta^3(1-\theta^2)}{1-\theta^8}$$

P26. ARMA(1,1) Model, ψ weight

$$z_t - \mu = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots = \frac{1-\theta B}{1-\phi B} a_t$$

or

$$(1-\phi B)(a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots) = 1 - \theta B$$

Hence

$$B^1 : \psi_1 - \phi = -\theta, \psi_1 = \phi - \theta$$

$$B^2 : \psi_2 - \phi\psi_1 = 0, \psi_2 = \phi\psi_1 = (\phi - \psi)\phi$$

$$B^j : \psi_j - \phi\psi_{j-1} = 0, \psi_j = \phi\psi_{j-1} = (\phi - \theta)\phi^{j-1}, j > 0$$

- π representation for ARMA(1,1)

$$a_t = z_t - \mu - \pi_1(z_{t-1} - \mu) - \pi_2(z_{t-2} - \mu) - \cdots = \pi(B)(z_t - \mu) = \frac{1-\phi B}{1-\theta B}(z_t - \mu)$$

$$B^1 : -\pi_1 - \theta = -\phi, \pi_1 = \phi - \theta$$

$$B^2 : -\pi_2 + \theta\pi_1 = 0, \pi_2 = \theta\pi_1 = (\phi - \theta)\theta$$

$$B^j : -\pi_j + \theta\pi_{j-1}, \pi_j = \theta\pi_{j-1} = (\phi - \theta)\theta^{j-1}$$

- Autocorrelation function of ARMA(1,1), set $\mathrm{E} z_t = \mu = 0$

$$\gamma_k = \phi\gamma_{k-1} + \mathrm{E}(a_t z_{t-k}) - \theta \mathrm{E}(a_{t-1} z_{t-k})$$

If $k > 1$, $\mathrm{E}(a_t z_{t-k}) = \theta \mathrm{E}(a_{t-1} z_{t-k}) = 0$, Therefore

$$\gamma_k = \phi\gamma_{k-1}, \text{ for } k > 1$$

$$\mathrm{E}(a_t z_t) = \mathrm{E}[a_t(a_t + \psi_1 z_{t-1} + \psi_2 z_{t-2} + \cdots)] = \mathrm{E} a_t^2 = \sigma^2$$

$$\mathrm{E}(a_t z_{t-1}) = \mathrm{E}[a_{t-1}(a_t + \psi_1 z_{t-1} + \psi_2 z_{t-2} + \cdots)] = \mathrm{E}\psi_1 a_{t-1}^2 = (\phi - \theta)\sigma^2$$

Then

$$k = 0 : \gamma_0 = \phi\gamma_1 + \sigma^2 - \theta(\phi - \theta)\sigma^2$$

$$k = 1 : \gamma_1 = \phi\gamma_0 - \theta\sigma^2$$

Solve this equation system to get γ_0 and γ_1. (*check by yourself*)

P49. AR(1) process: Suppose we are given past observations $z_n, z_{n-1}, \ldots$ and wish to predict z_{n+l}. For $\ell = 1$

$$\begin{aligned} z_n(l) &= \mathrm{E}(z_{n+1}|z_n, z_{n-1}, \ldots) \\ &= \mathrm{E}\{[\mu + \phi(z_n - \mu) + a_{n+1}]|z_n, z_{n-1}, \ldots\} \\ &= \mu + \phi(z_n - \mu). \end{aligned}$$

Since $\mathrm{E}(z_n|z_n, z_{n-1}, \ldots) = z_n, \mathrm{E}(a_{n+1}|z_n, z_{n-1}, \ldots) = 0$.
For $\ell = 2$,

$$\begin{aligned} z_n(2) &= \mathrm{E}(z_{n+2}|z_n, z_{n-1}, \ldots) \\ &= \mathrm{E}\{[\mu + \phi(z_{n+1} - \mu) + a_{n+2}]|z_n, z_{n-1}, \ldots\} \\ &= \mu + \phi[z_n(1) - \mu] = \mu + \phi^2(z_n - \mu) \end{aligned}$$

The ℓ-step-ahead prediction can be written as

$$\begin{aligned} z_n(\ell) &= \mathrm{E}(z_{n+\ell}|z_n, z_{n-1}, \ldots) \\ &= \mathrm{E}\{[\mu + \phi(z_{n+\ell-1} - \mu) + a_{n+\ell}]|z_n, z_{n-1}, \ldots\} \\ &= \mu + \phi[z_n(\ell - 1) - \mu] = \cdots = \mu + \phi^\ell(z_n - \mu) \end{aligned}$$

- AR(1) continuous. The forecast errors corresponding to the above forecasts are

$$e_n(1) = z_{n+1} - z_n(1) = \mu + \phi(z_n - \mu) + a_{n+1} - [\mu + \phi(z_n - \mu)] = a_{n+1}.$$

$$\begin{aligned} e_n(2) &= z_{n+2} - z_n(2) \\ &= \mu + \phi(z_{n+1} - \mu) + a_{n+2} - [\mu + \phi^2(z_n - \mu)] \\ &= a_{n+2} + \phi[(z_{n+1} - \mu) - \phi(z_n - \mu)] = a_{n+2} + \phi a_{n+1} \end{aligned}$$

Similarly, it can be shown that

$$e_n(\ell) = a_{n+\ell} + \phi a_{n+\ell-1} + \cdots + \phi^{\ell-1} a_{n+1}$$

and

$$V[e_n(\ell)] = \sigma^2(1 + \phi^2 + \cdots + \phi^{2(\ell-1)}) = \sigma^2 \frac{1 - \phi^{2\ell}}{1 - \phi^2}$$

- AR(1) continuous. Consider the yield series. It is shown that this series can be described by an AR(1) model with $\hat{\mu} = 0.97, \hat{\phi} = 0.85$ and $\hat{\sigma}^2 = 0.024$. Since the last observation is $z_{156} = 0.49$, the forecasts are

$$\hat{z}_{156}(1) = 0.97 + 0.85(0.49 - 0.97) = 0.56$$

$$\hat{z}_{156}(2) = 0.97 + 0.85^2(0.49 - 0.97) = 0.62$$

$$\hat{z}_{156}(3) = 0.97 + 0.85^3(0.49 - 0.97) = 0.68$$

and their variance are

$$\mathrm{Var}[e_{156}(1)] = 0.024$$

$$\mathrm{Var}[e_{156}(2)] = 0.024\frac{1 - .85^4}{1 - .85^2} = 0.041$$

$$\mathrm{Var}[e_{156}(3)] = 0.024\frac{1 - .85^6}{1 - .85^4} = 0.054$$

- AR(2) process $z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t$ with $\mu = 0$. The one-step ahead ($\ell = 1$ forecast given the observations $z_n, z_{n-1}, \ldots$ can be expressed as

$$\begin{aligned} z_n(1) &= \mathrm{E}(z_{n+1} | z_n, z_{n-1}, \ldots) \\ &= \mathrm{E}[(\phi_1 z_n + \phi_2 z_{n-1} + a_{n+1}) | z_n, z_{n-1}, \ldots] \\ &= \phi_1 z_n + \phi_2 z_{n-1} \end{aligned}$$

  For $\ell = 2$

$$\begin{aligned} z_n(2) &= \mathrm{E}(z_{n+2} | z_n, z_{n-1}, \ldots) \\ &= \mathrm{E}[(\phi_1 z_{n+1} + \phi_2 z_n + a_{n+2}) | z_n, z_{n-1}, \ldots] \\ &= \phi_1 z_n(1) + \phi_2 z_n. \end{aligned}$$

  In general,

$$z_n(\ell) = \phi_1 z_n(\ell-1) + \phi_2 z_n(\ell-2), \text{ or } (1 - \phi_1 B - \phi_2 B^2) z_n(\ell) = 0, \ell > 0$$

  .

- The forecast error and its weight can be calculated by substituting the ψ weight of the AR(2) model. It is easily to seen that the ψ weights are

$$\psi_1 = \phi_1, \psi_2 = \phi_1^2 + \phi_2, \psi_j = \phi_1 \psi_{j-1} + \phi_2 \psi_{j-2}, j \geq 2$$

- ARIMA(0,1,1) process. Given the observations $z_n, z_{n-1}, \ldots$. the predictions from the model $z_t = z_{t-1} + a_t - \theta a_{t-1}$ can be obtained from the conditional expectation form:

$$
\begin{aligned}
z_n(1) &= \mathrm{E}(Z_{n+1}|z_n, z_{n-1}, \ldots) \\
&= \mathrm{E}[(z_n + a_{n+1} - \theta a_n)|z_n, z_{n-1}, \ldots] = z_n - \theta a_n
\end{aligned}
$$

$$
z_n(2) = \mathrm{E}(z_{n+2}|z_n, z_{n=1}, \ldots) = z_n - \theta a_n
$$

and in general

$$
z_n(\ell) = z_n(\ell - 1), \text{or } (1 - B)z_n(\ell) = 0.
$$

The ψ weight can be obtained from $\psi(B) = (1 - \theta B)/(1 - B)$ and it is given by $\psi_j = 1 - \theta$ for all $j > 0$. Hence the forecast error is given by

$$
e_n(\ell) = a_{n+\ell} + (1 - \theta)(a_{n+\ell-1} + \cdots + a_{n+1})
$$

and its variance by

$$
V[e_n(\ell)] = \sigma^2[1 + (\ell - 1)(1 - \theta)^2].
$$

- ARIMA(0,1,1) process continuous. Alternatively, the forecast can be expressed as a linear combination of the past observations. Write the model in its autoregressive representation

$$z_t = \sum_{j=1}^{\infty} \pi_j z_{t-j} + a_t$$

where $\pi_j = (1-\theta)\theta^{j-1}, j \geq 1$ are coefficients in $\pi(B) = (1-B)/(1-\theta B)$. Hence $z_t = (1-\theta)(z_{t-1} + \theta z_{t-2} + \theta^2 z_{t-3} + \cdots) + a_t$. Taking the conditional expectation of z_{n+1} given $z_n, z_{n-1}, \ldots$, we find that

$$z_n(1) = (1-\theta)(z_n + \theta z_{n-1} + \theta^2 z_{n-2} + \cdots).$$

This forecast is an exponentially weighted average of present and past observation and is the same as that obtained from single exponential smoothing with a smoothing constant $\alpha = 1 - \theta$.

- ARIMA(1,1,1) process: $(1-\phi B)(1-B)z_t = \theta_0 + (1-\theta B)a_t$ or

$$z_t = \theta_0 + (1+\phi)z_{t-1} - \phi z_{t-2} + a_t - \theta a_{t-1}$$

  Taking conditional expectation, we can calculate the forecasting according to

$$z_n(1) = \mathrm{E}(z_{n+1}|z_n, z_{n-1}, \ldots) = \theta_0 + (1+\phi)z_n - \phi z_{n-1} - \theta a_n$$

$$z_n(2) = \mathrm{E}(z_{n+2}|z_n, z_{n-1}, \ldots) = \theta_0 + (1+\phi)z_n(1) - \phi z_n$$

  and so on

$$z_n(\ell) = \mathrm{E}(z_{n+\ell}|z_n, z_{n-1}, \ldots) = \theta_0 + (1+\phi)z_n(\ell-1) - \phi z_n(\ell-2)$$

  or For $\ell \geq 2$

$$[(1-(1+\phi)B + \phi B^2]z_n(\ell) = (1-\phi B)(1-B)z_n(\ell) = \theta_0.$$

- ARIMA(0,2,2) process: $(1-B)^2 z_t = (1-\theta_1 B - \theta_2 B^2) a_t$ or

$$z_t = 2z_{t-1} - z_{t-2} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$

Given the observations $z_n, z_{n-1}, \ldots,$ the forecasts are

$$z_n(1) = \mathrm{E}(z_{n+1}|z_n, z_{n-1}, \ldots) = 2z_n - z_{n-1} - \theta_1 a_n - \theta_2 a_{n-1}$$

$$z_n(2) = \mathrm{E}(z_{n+2}|z_n, z_{n-1}, \ldots) = 2z_n(1) - z_n - \theta_2 a_n$$
$$z_n(3) = \mathrm{E}(z_{n+3}|z_n, z_{n-1}, \ldots) = 2z_n(2) - z_n(1)$$

and

$$z_n(\ell) = \mathrm{E}(z_{n+\ell}|z_n, z_{n-1}, \ldots) = 2z_n(\ell-1) - z_n(\ell-2), \ell \geq 3$$

or

$$(1-2B+B^2)z_n(\ell) = (1-B)^2 z_n(\ell) = 0$$