Final Review Side A

2023年4月18日 23:24

Autoregressive (AR)

* AR always invertible

$\Phi_p(B)x_t = w_t$ where $\Phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 \cdots$ or no solution

Causal condition: $|\text{root to } \Phi_p(z)| > 1$ or $\Phi_p(B)^{-1} = \Psi_\infty(B)$ for some Ψ

- AR(1): $|\phi| < 1$
- AR(2): $|\phi_2| < 1$ $\phi_1 + \phi_2 < 1$ $\phi_2 - \phi_1 < 1$

AR(p) to MA(∞)

$\Phi_p(B)x_t = w_t \Rightarrow x_t = \Psi_\infty(B) w_t$ where $\Psi_\infty(B) = 1 + \psi_1 B + \psi_2 B^2 \cdots$

use $\Phi_p(B) \cdot \Psi_\infty(B) = 1$ we can derive coefficients ψ_i

$\psi_0 = 1$ $\psi_j = \phi_1 \psi_{j-1} + \phi_2 \psi_{j-2} + \cdots$

Moving-average (MA)

* MA always stationary and causal

$x_t = \Theta_q(B) w_t$ where $\Theta_q(B) = 1 + \theta_1 B + \theta_2 B^2 \cdots$ or no solution

Invertible condition: $|\text{root to } \Theta_q(z)| > 1$ or $\Theta_q(B)^{-1} = \Pi_\infty(B)$ for some Π

- MA(1): $|\theta| < 1$
- MA(2): $|\theta_2| < 1$ $\theta_1 + \theta_2 < 1$ $\theta_2 - \theta_1 < 1$

MA(q) to AR(∞)

$x_t = \Theta_q(B) w_t \Rightarrow \Pi_\infty(B) x_t = w_t$ where $\Pi_\infty(B) = 1 - \pi_1 B - \pi_2 B^2 \cdots$

use $\Theta_q(B) \cdot \Pi_\infty(B) = 1$ we can derive coefficients π_i

$\pi_0 = -1$ $\pi_j = -\theta_1 \pi_{j-1} - \theta_2 \pi_{j-2} - \cdots$

ARMA(p.q)

$\Phi_p(B) x_t = \Theta_q(B) w_t$ where $\Phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 \cdots$

$\Theta_q(B) = 1 + \theta_1 B + \theta_2 B^2 \cdots$

Causal condition: $|\text{root to } \Phi_p(z)| > 1$

Invertible condition: $|\text{root to } \Theta_q(z)| > 1$

Identifiability: $\Phi_p(z) = 0$ and $\Theta_q(z) = 0$ has no common roots

Causal ARMA(p.q) to MA(∞)

$\Phi_p(B) x_t = \Theta_q(B) w_t \Rightarrow x_t = \Psi_\infty(B) w_t$

use $\Phi_p(B) \cdot \Psi_\infty(B) = \Theta_q(B)$ we can derive coefficients ψ_i

$\psi_0 = 1$ $\psi_j = \theta_j + \phi_1 \psi_{j-1} + \phi_2 \psi_{j-2} + \cdots$

Invertible ARMA(p.q) to AR(∞)

$\Phi_p(B) x_t = \Theta_q(B) w_t \Rightarrow \Pi_\infty(B) x_t = w_t$

use $\Phi_p(B) = \Theta_q(B) \cdot \Pi_\infty(B)$ we can derive coefficients π_i

$\pi_0 = -1$ $\pi_j = \phi_j - \theta_1 \pi_{j-1} - \theta_2 \pi_{j-2} - \cdots$

ARIMA(p.d.q)

$\Phi_p(B)(1-B)^d x_t = \delta + \Theta_q(B) w_t$ where $\delta = \mu(1 - \phi_1 - \phi_2 \cdots - \phi_p)$

ex: $(1 - \phi_1 B - \phi_2 B^2)(1-B) x_t = (1 + \theta B) w_t$ is ARIMA(2,1,1)

$p = 2$, $d = 1$, $q = 1$

ARIMA(p.d.q)×(P.D.Q)s

$\Phi_p(B) \cdot \Phi_P(B^s) \cdot (1-B)^d (1-B^s)^D x_t = \delta + \Theta_q(B) \cdot \Theta_Q(B^s) w_t$

ex: $(1 - 0.5B)(1 - B^4) x_t = (1 - 0.3B) w_t$ is ARIMA$(1,0,1)\times(0,1,0)_4$

$p = 1$, $s = 4, D = 1$, $q = 1$

* Proof

• AR(1) $x_t - \phi x_{t-1} = w_t$

mean: $E(x_t) = \sum_0^\infty \phi_i w_{t-i} = 0$

variance: $\gamma(0) = E(\sum_0^\infty \phi^i w_{t-i})^2 = \sigma_w^2 \sum_0^\infty \phi^{2i} = \frac{\sigma_w^2}{1-\phi^2}$

covariance: $\gamma(h) = E(x_t x_{t+h}) = \sigma_w^2 \sum_0^\infty \phi^i \phi^{i+h} = \phi^h \sigma_w^2 \sum_0^\infty \phi^{2i} = \sigma_w^2 \cdot \frac{\phi^h}{1-\phi^2}$

ACF: $\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \phi^h$

• AR(2) $x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} = w_t$

covariance: $\gamma(h) = E(x_t x_{t+h}) = \phi_1 \gamma(h-1) + \phi_2 \gamma(h-2)$

ACF: $\rho(1) = \frac{\phi_1}{1-\phi_2}$, $\rho(2) = \frac{\phi_1^2}{1-\phi_2} + \phi_2$

$\rho(h) = \phi_1 \rho(h-1) + \phi_2 \rho(h-2)$ for $h \geq 3$

Derivation -

$E(x_t x_{t-h}) = \phi_1 E(x_{t-1} x_{t-h})$
$+ \phi_2 E(x_{t-2} x_{t-h}) + E(w_t w_{t-h})$

$\gamma(h) = \phi_1 \gamma(h-1) + \phi_2 \gamma(h-2)$

divide by $\gamma(0)$

$\gamma(h) = \phi_1 \rho(h-1) + \phi_2 \rho(h-2)$

use $\rho(0) = 1$, $\rho(1) = \rho(-1)$ get result

• AR(p)

ACF: $\rho(h) = \phi_1 \rho(h-1) + \phi_2 \rho(h-2) + \cdots$

PACF: $\phi_{hh} = \phi_h$ for $h \leq p$, $= 0$ for $h > p$

$\phi_{hh} = \text{corr}(x_{t+h} - \hat{x}_{t+h}, x_t - \hat{x}_t)$

$\phi_{11} = \rho(1)$ $\phi_{pp} = \phi_p$ * for any model. AR, MA, ARMA

• MA(1) $x_t = w_t + \theta w_{t-1}$

trend of scatterplot = covariance

mean: $E(x_t) = 0$

variance: $\gamma(0) = E(w_t^2 + 2\theta w_t w_{t-1} + \theta^2 w_{t-1}^2) = (1+\theta^2)\sigma_w^2$

covariance: $\gamma(h) = E(w_t w_{t-h}) + \theta E(w_{t-1} w_{t-h}) + \theta E(w_t w_{t-h}) + \theta^2 E(w_{t-1} w_{t+h-1})$

$= (1+\theta^2)\sigma_w^2$ for $h = 0$, $\theta \sigma_w^2$ for $h = \pm 1$, 0 for $h = \pm 2, \pm 3 \cdots$

ACF: $\rho(h) = \frac{\gamma(h)}{\gamma(0)} = 1$ for $h = 0$. $\frac{\theta}{1+\theta^2}$ for $h = \pm 1$, 0 for $h = \pm 2, \pm 3 \cdots$

PACF: $\phi_{hh} = (-\theta)^h (1-\theta^2) \cdot (1 - \theta^{2(h+1)})^{-1}$ for $h \geq 1$

• MA(2) $x_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2}$

ACF: $\rho(1) = \frac{\theta_1(1+\theta_2)}{1+\theta_1^2+\theta_2^2}$. $\rho(2) = \frac{\theta_2}{1+\theta_1^2+\theta_2^2}$. $\rho(h) = 0$ for $h > 2$

• MA(q)

mean: $E(x_t) = \sum_0^q \theta_j E(w_{t-j}) = 0$

variance: $\gamma(0) = (1 + \theta_1^2 + \theta_2^2 \cdots)\sigma_w^2$

covariance: $\gamma(h) = \sigma_w^2 \sum_h^q \theta_i \theta_{i-h}$ for $h = 0 \cdots \pm q$, 0 for $h > q$

ACF: $\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \sum_h^q \theta_i \theta_{i-h} \cdot (\sum_0^q \theta_j^2)^{-1}$ for $h = 0 \cdots \pm q$, 0 for $h > q$

• ARMA(1,1) $x_t - \phi x_{t-1} = w_t + \theta w_{t-1}$

covariance: $\gamma(0) = \sigma_w^2 \frac{1 + 2\phi\theta + \theta^2}{1-\phi^2}$ $\gamma(h) = \sigma_w^2 \frac{(1+\phi\theta)(\phi+\theta)}{1-\phi^2} \phi^{h-1}$ for $h \geq 1$

ACF: $\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{(1+\phi\theta)(\phi+\theta)}{1+2\phi\theta+\theta^2} \phi^{h-1}$ for $h \geq 1$

Derivation :

$\gamma(h) = Cov(x_{t+h}, x_t) = E(x_{t+h} x_t)$

$= \phi E(x_{t+h-1} x_t) + E(w_{t+h} x_t) + \theta E(w_{t+h-1} x_t)$

$E(w_{t+h-1} \sum_0^\infty \psi_j w_{t-j})$: $\begin{cases} \psi_1 \sigma_w^2 & h = 0 \\ \psi_0 \sigma_w^2 & h = 1 \\ 0 & h \geq 2 \end{cases}$

$E(w_{t+h} \sum_0^\infty \psi_j w_{t-j})$: $\begin{cases} \psi_0 \sigma_w^2 & h = 0 \\ 0 & h \geq 1 \end{cases}$

$\gamma(h) = \begin{cases} \phi\gamma(1) + \sigma_w^2(1 + \phi\theta + \theta^2) & h = 0 \\ \phi\gamma(0) + \sigma_w^2\theta & h = 1 \\ \phi\gamma(h-1) & h \geq 2 \end{cases}$ * $\gamma(h) = \phi^{h-1}\gamma(1)$

Initial condition $\begin{cases} \gamma(0) = \phi\gamma(1) + \sigma_w^2(1 + \phi\theta + \theta^2) \\ \gamma(1) = \phi\gamma(0) + \sigma_w^2\theta \end{cases}$

solve for $\gamma(0)$. $\gamma(1)$, and get $\gamma(h)$

	white noise	AR(p)	MA(q)	ARMA (p.q)	random walk
PACF	zero	spike through lag p, cut off after lag p	tails off exp decay (direct or oscillation)	decay at lag p (direct or oscillation)	zero after lag 1
ACF	zero	exp decay (direct or oscillation)	cutoff after lag q	decay at lag q (direct or oscillation)	zero

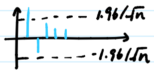
	ACF	PCF
$AR(P)_s$ / $ARIMA(0)\times(P,0,0)_s$	tails off at lag k×s	cuts off after lag P×s
$MA(Q)_s$ / $ARIMA(0)\times(0,0,Q)_s$	cuts off at lag Q×s	tails off at lag k×s
$ARMA(P,Q)_s$ / $ARIMA(0)\times(P,0,Q)_s$	tails off at lag k×s	tails off at lag k×s

Final Review Side B

2023年4月20日 15:08

- week station: $E(x_t)$ doesn't depend on t, finite variance
  auto covariance: $\gamma(h) = Cov(x_t, x_{t+h})$

- model selection: max R^2 min AIC, SSE
  $AIC = \log \hat\sigma_k^2 + \frac{n+2k}{n}$ $\quad BIC = \log \hat\sigma_k^2 + \frac{k \log n}{n}$ $\quad$ linear model
  $AICc = \log \hat\sigma_k^2 + \frac{n+k}{n-k-2}$ $\quad \hat\sigma_k^2 = SSE(k)/n$ $\quad \Rightarrow$ residue uncorrelated
  $DF = n - \#$ of β_i $\quad q = \#$ of x_i in full model
  $r = \#$ of x_i in reduced
  $F = \frac{(SSE_r - SSE_q)/(q-r)}{SSE_q / df_q}$ $\quad F \geq$ critical reject H_0

- ACF graph  Large $n \to \hat\rho(h)$ normal

- Box-Cox $\quad x^\lambda - 1/\lambda$ if $\lambda \neq 0$, $\log X$ if $\lambda = 0$
  Ljung-Box statistics: $H_0: \rho(h) = 0$ $\quad p < 0.05$: reject H_0

- Sample ACF: $\hat\rho_x(h) = \frac{\Sigma(x_t - \bar x)(x_{t+h} - \bar x)}{\Sigma(x_t - \bar x)^2}$
  Test $H_0\ \rho(1) = 0$: t-test $\frac{\hat\rho_x(1)}{1/\sqrt n}$ $\quad t >$ critical reject H_0

- Theory CCF: $\rho_{xy}(h) = \frac{\gamma_{xy}(h)}{\sqrt{\gamma_x(0)\gamma_y(0)}}$ $\quad \gamma_x(0) = V(x)$ $\quad \gamma_y(0) = V(y)$ $\quad \gamma_{xy}(h) = Cov(x,y)$
  Sample CCF: $\hat\gamma_x(0) = \frac{1}{n}\Sigma(x_t - \bar x)^2$ $\quad \hat\gamma_y(0) = \frac{1}{n}\Sigma(y_t - \bar y)^2$ $\quad \hat\gamma_{xy}(h) = \frac{1}{n}\sum^{n-h}(x_{t+h} - \bar x)(y_t - \bar y)$

- Best Linear Predictor
  given $x_1 \cdots x_n$, $x_{n+m}^n = a_0 + \Sigma_1^n a_k x_k$
  $E((x_{n+m} - x_{n+m}^n) x_k) = 0, k = 0, 1 \cdots n, x_0 = 1$ $\quad$ (estimation - true) $\perp x_k$
  ✡ $E(x_{n+m}) = E(x_{n+m}^n)$ when $k = 0$, $x_0 = 1$
  $X_{n+m} = a_0 + a_1 x_1 + \cdots + a_k x_k$

Example:
$P(a_1 Y_1 + a_2 Y_2 | Z) = a_1 P(Y_1|Z) + a_2 P(Y_2|Z)$
BLE satisfy $\quad E[\underset{x_k}{Z} \cdot (\underset{\text{true}}{Y} - \underset{\text{estimation}}{P(Y|Z)})] = 0$
where $P(Y|Z) = aZ + b$
$E(Z \cdot Y) - aE(Z^2) - bE(Z)$
$E(Z \cdot Y) - aE(Z^2) - (E(Y) - aE(Z)) \cdot E(Z)$
$\underbrace{E(Z \cdot Y) - E(Y) \cdot E(Z)}_{Cov(Y,Z)} + \underbrace{a(E(Z))^2 - aE(Z^2)}_{aV(Z)}$
got a, got b, $\to$ calculate $P(Y|Z)$
substitute Y for $a_1Y_1 + a_2Y_2$ in $P(Y|Z)$

Example:
$X_t \sim AR(1)$, given X_1, X_3 find BLE of X_2
BLE is $Y = aX_1 + bX_3$ satisfies $E(X_1(X_2 - Y)) = 0$
and $E(X_3(X_2 - Y)) = 0$ (+)
$E(X_t) = \phi E(X_{t-1}) \Rightarrow \mu = E(X) = 0$
$AR(2): \gamma(h) = \sigma_w^2 \frac{\phi^h}{1 - \phi^2}$
solve (+): $\frac{\phi}{1-\phi^2} - \frac{a}{1-\phi^2} - \frac{\phi^2 b}{1-\phi^2} = 0$, $\frac{\phi}{1-\phi^2} - \frac{a\phi^2}{1-\phi^2} - \frac{b}{1-\phi^2} = 0$
get $a = \phi - \phi^2 b$
substitute into 2nd: $b = \frac{\phi}{1+\phi^2}$, $a = \frac{\phi}{1+\phi^2}$
back to Y get $Y = \frac{\phi}{1+\phi^2}(X_1 + X_3)$

- Yule-Walker equation (for $AR(p)$ only)
  $\sigma_w^2 = \gamma(0) - \phi_1\gamma(1) - \cdots - \phi_p\gamma(p)$
  $\gamma(h) = \phi_1\gamma(h-1) + \cdots + \phi_p\gamma(h-p)$. $h = 1, 2 \cdots p$

$\phi = \Gamma_p^{-1}\hat\gamma_p$, $\hat\sigma_w^2 = \hat\gamma(0) - \hat\gamma_p' \Gamma_p^{-1} \hat\gamma_p$ $\quad$ or $\quad \phi = \hat R_p^{-1}\hat\rho_p$, $\hat\sigma_w^2 = \hat\gamma(0)[1 - \hat\rho_p \hat R_p^{-1} \hat\rho_p]$

$$\begin{pmatrix}\phi_1\\ \phi_2\\ \vdots\\ \phi_p\end{pmatrix} = \begin{pmatrix}\gamma(0) & \gamma(-1) & \cdots & \gamma(1-p)\\ \gamma(1) & \gamma(0) & \cdots & \gamma(2-p)\\ & \vdots & & \\ \gamma(p-1) & \gamma(p-2) & \cdots & \gamma(0)\end{pmatrix}^{-1} \cdot \begin{pmatrix}\gamma(1)\\ \gamma(2)\\ \vdots\\ \gamma(p)\end{pmatrix} \qquad \begin{pmatrix}\phi_1\\ \phi_2\\ \vdots\\ \phi_p\end{pmatrix} = \begin{pmatrix}\rho(0) & \rho(-1) & \cdots & \rho(1-p)\\ \rho(1) & \rho(0) & \cdots & \rho(2-p)\\ & \vdots & & \\ \rho(p-1) & \rho(p-2) & \cdots & \rho(0)\end{pmatrix}^{-1} \cdot \begin{pmatrix}\rho(1)\\ \rho(2)\\ \vdots\\ \rho(p)\end{pmatrix}$$

$$\hat\sigma_w^2 = \hat\gamma(0) - (\gamma(1) \cdots \gamma(p)) \cdot \begin{pmatrix}\phi_1\\ \vdots\\ \phi_p\end{pmatrix} \qquad \hat\sigma_w^2 = \hat\gamma(0) \cdot [1 - (\rho(1) \cdots \rho(p)) \cdot \begin{pmatrix}\phi_1\\ \vdots\\ \phi_p\end{pmatrix}]$$

- Large Sample Asymptotic: $V(\hat\phi) = \frac{\sigma^2 \cdot \Gamma_p^{-1}}{n} = \frac{\sigma^2 R_p^{-1}}{n \cdot \gamma(0)}$ $\quad$ 95% CI: $\hat\phi_i \pm 1.96 \cdot se(\hat\phi_i)$
  $\phi_{nn} \sim N(0, \frac{1}{n})$

Prediction Interval for x_{n+m}

$x_{n+m}^n \pm 1.28\sqrt{P_{n+m}^n}$
$\quad$ critical value
$P_{n+m}^n = \sigma_w^2 \cdot \Sigma_0^m \psi_j^2$

ex: $P_{100}^{103} = \psi_0^2 + \psi_1^2 + \psi_2^2$
$\begin{cases}\psi_k = \Sigma_1^k \phi_j \psi_{k-j} & k < p\\ \psi_k = \Sigma_1^p \phi_j \psi_{k-j} & k \geq p\end{cases}$

AR(2) prediction
$x_{n+1}^n = \phi_{n1}x_n + \phi_{n2}x_{n-1} \cdots \phi_{nn}x_1$
$X_2^1 = \phi_{11}x_1$
$X_3^2 = \phi_{21}x_1 + \phi_{22}x_2$
✡ $AR(2)$: $x_{n+1}^n = \phi_1 x_n + \phi_2 x_{n-1}$

- (Spectral Density - $\Sigma_{h=-\infty}^{\infty} \gamma(h)\exp(-2\pi i w h)$ $\quad -\infty < w < \infty$
  Normalized SD - $\Sigma_{h=-\infty}^{\infty} \rho(h)\exp(-2\pi i w h)$

BLE: $E(X_{n+m} | x_1 \cdots x_n)$
$E(w_t | x_n \cdots x_1) = \begin{cases}0 & t > n\\ w_t & t \leq n\end{cases}$
Example: given $AR(1)$ $x_t = c + \phi x_{t-1} + w_t$
show $E(x_{t+h} | x_t \cdots x_1) \to E(x_t) = \frac{c}{1-\phi}$
Let $E(x_t) = E(x_{t-1}) \cdots = \mu$
$E(x_t) = c + \phi E(x_{t-1}) + E(w_t) \Rightarrow \mu = \frac{c}{1-\phi}$
$x_{t+h} = c + \phi x_{t+h-1} + w_{t+h}$
$= c + \phi(c + \phi x_{t+h-2} + w_{t+h-1}) + w_{t+h}$
$\vdots$
$= c + \phi c + \phi^2 c + \cdots + \phi^h x_t + w_{t+h} + \phi w_{t+h-1} \cdots \phi^h w_t$
$E(x_{t+h}) = c + \phi c + \cdots + \phi^h x_t + 0 + \cdots + \phi^h w_t$
$|\phi| < 1$ $\quad |\phi^h| \to 0$ $\quad E(x_{t+h} | x_1 \cdots x_n) = c(1 + \phi + \phi^2 \cdots) = \frac{c}{1-\phi}$

- Dickey-Fuller Test for stationary
  $H_0: \delta = 0, (\phi = 1)$ $\quad (\Delta x_t = x_t - x_{t-1})$ $\quad p \leq 0.05$ reject H_0
  3 types: $\Delta x_t = \delta x_{t-1} + w_t$ $\quad$ no drift / trend
  $\Delta x_t = a_0 + \delta x_{t-1} + w_t$ $\quad$ drift only
  $\Delta x_t = a_0 + a_1 t + \delta x_{t-1} + w_t$ $\quad$ drift & trend

- Residue Analysis
  $H_0: \rho_e(h) = 0$, $H_1: \rho_e(h) \neq 0$ $\quad$ not white noise
  $\hat\rho(h) = \Sigma_{h+1}^n e_t e_{t-h} / \Sigma_1^n e_t^2$

(Box-Pierce Portmanteau - $Q_m = n\Sigma_{h=1}^m \hat\rho^2(h) \sim \chi^2_{m-p-q}$
Ljung-Box Portmanteau - $\tilde Q_m = n(n+2)\Sigma_{h=1}^m \hat\rho^2(h)/(n-h) \sim \chi^2_{m-p-q}$
reject H_0 if $Q \geq \chi^2$

- Durbin-Levinson Algorithm for PACF
  $\phi_{nn} = \frac{\rho(n) - \Sigma_{k=1}^{n-1}\phi_{n-1,k}\rho(n-k)}{1 - \Sigma_1^{n-1}\phi_{n-1,k}\rho(k)}$ $\quad n \geq 1$
  $\phi_{nk} = \phi_{n-1,k} - \phi_{nn}\phi_{n-1,n-k}$ $\quad k = 1 \cdots n-1$, $n \geq 2$
  $P_{n+1}^n = \gamma(0)\prod_{j=1}^n(1 - \phi_{jj}^2)$

Shapiro test $\to$ normality $\quad H_0$: normal. $p \leq 0.05$ reject

- Durbin-Watson Test for autocorrelation
  H_0: not autocorrelated or $x_t = w_t$
  $d = \frac{\Sigma_{t=2}^n (e_t - e_{t-1})^2}{\Sigma_{t=1}^n e_t^2}$ $\quad d < d_{L\alpha}$: reject
  $d > d_{U\alpha}$: do not reject
  otherwise: inconclusive