LECTURE 2

Partial autocorrelation functions (PACF)

Yule-Walker equations

Solving Yule-Walker equations—Cramer's rule and Durbin-Levinson algorithm

ACF and MA(q) processes

The maximum lag of the non-zero sample autocorrelation is a good indicator of the MA(q) processes.

- The ACF of MA(q) processes, $Y_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q}$, cut off after lag q.
- $\rho_k = \begin{cases} \dfrac{\overbrace{\theta_k + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + \cdots + \theta_{q-k}\theta_q}^{\gamma(k)}}{\underbrace{1 + \theta_1^2 + \cdots + \theta_q^2}_{\gamma(0)}} & , k = 1, \dots, q \\ 0 & , k > q \end{cases}$
- How about ACF of the $AR(p)$ processes?

Partial autocorrelation function (PACF)

The correlation between X_t and X_{t+k} after mutual linear dependency on the intervening variables, X_{t+1}, X_{t+2}, …, and X_{t+k-1} has been removed.

- The conditional correlation $\phi_{kk} = corr(X_t, X_{t+k} | X_{t+1}, \dots, X_{t+k-1})$ is usually referred to as the partial autocorrelation functions in time series analysis.
- PACF between X_t and X_{t+k} can be obtained as the regression coefficient associated with X_t when regressing X_{t+k} on its k lagged variables X_{t+k-1}, X_{t+k-2}, …, and X_t.

More PACF

The partial autocorrelation function at lag k can also be defined as the correlation between two prediction errors; that is,

- $\phi_{kk} =$ $Corr(X_t - \beta_1 X_{t-1} - \cdots \beta_{k-1} X_{t-2}, X_{t-k} - \beta_1 X_{t-k+2} - \cdots \beta_{k-1} X_{t-1})$, where β's are chosen to minimize the mean square error of the prediction.
- Example: The best linear prediction of X_t based on X_{t-1} alone is $\rho_1 X_{t-1}$. Thus, $\phi_{22} = Corr(X_t - \rho_1 X_{t-1}, X_{t-2} - \rho_1 X_{t-1})$.
- We will see in the next slide that $\phi_{kk} = 0 \; \forall \; k \geq p$ for $AR(p)$ processes.

NEXT PAGE

$\{X_t\}$ is causal stationary.

(1) $X_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$, $a_t \sim WN(0, \sigma^2)$

(2) $X_t = \phi_{k1} X_{t-1} + \cdots + \phi_{kk} X_{t-k} + a_t$

(use past observation to forecast next observation)

Yule-Walker equations

A general method for finding the partial autocorrelation function for any stationary process with autocorrelation function ρ_k is as follows.

Method:

For a given lag k, it can be shown that the ϕ_{kk} satisfy the Yule-Walker equations. $\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \phi_{k3}\rho_{j-3} + \cdots + \phi_{kk}\rho_{j-k}$, for $j = 1,2, \ldots, k$. That is, we regard $\rho_1, \ldots, \rho_k$ as given and wish to solve for ϕ_{kk}.

Remarks: If the process follows an $AR(p)$ model, then $\phi_{pp} = \phi_p$. In addition, we have shown that $\phi_{kk} = 0$ for $k > p$.

Derivation of Y-W equation

(1). $X_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$, $a_t \sim WN(0, \sigma^2)$

(2) $X_t = \phi_{k1} X_{t-1} + \phi_{k2} X_{t-2} + \cdots + \phi_{kk} X_{t-k} + e_t$

Multiply the above eqn. by $X_t, X_{t-1}, \cdots, X_{t-k}$

(3)
$$X_t^2 = \phi_{k1} X_t X_{t-1} + \phi_{k2} X_t X_{t-2} + \cdots + \phi_{kk} X_t X_{t-k} + X_t e_t$$
$$X_{t-1} X_t = \phi_{k1} X_{t-1}^2 + \phi_{k2} X_{t-1} X_{t-2} + \cdots + \phi_{kk} X_{t-1} X_{t-k} + X_{t-1} e_t$$
$$\cdots\cdots\cdots$$
$$X_{t-k} X_t = \phi_{k1} X_{t-k} X_{t-1} + \phi_{k2} X_{t-k} X_{t-2} + \cdots + \phi_{kk} X_{t-k} X_{t-k} + X_{t-k} e_t$$

Take expectation on the system in the previous slide

(4).
$$\underbrace{E(X_t^2)}_{\gamma(0)} = \phi_{k1} \underbrace{E(X_t X_{t-1})}_{\gamma(1)} + \phi_{k2} \underbrace{E(X_t X_{t-2})}_{\gamma(2)} + \cdots + \phi_{kk} \underbrace{E(X_t X_{t-k})}_{\gamma(k)} + \underbrace{E(X_t e_t)}_{\sigma_e^2}$$

similarly

$$\underbrace{E(X_{t-1} X_t)}_{\gamma(1)} = \phi_{k1} \underbrace{E(X_{t-1}^2)}_{\gamma(0)} + \phi_{k2} \underbrace{E(X_{t-1} X_{t-2})}_{\gamma(1)} + \cdots + \phi_{kk} \underbrace{E(X_{t-1} X_{t-k})}_{\gamma(k-1)} + \underbrace{E(X_{t-1} e_t)}_{0}$$
$$\cdots\cdots\cdots$$
$$\underbrace{E(X_{t-k} X_t)}_{\gamma(k)} = \phi_{k1} \underbrace{E(X_{t-k} X_{t-1})}_{\gamma(k-1)} + \phi_{k2} \underbrace{E(X_{t-k} X_{t-2})}_{\gamma(k-2)} + \cdots + \phi_{kk} \underbrace{E(X_{t-k} X_{t-k})}_{\gamma(0)} + \underbrace{E(X_{t-k} e_t)}_{0}$$

no overlay between this different time index expected to be 0.

Y-W walker equations

$$\gamma(0) = \phi_{k1}\gamma(1) + \phi_{k2}\gamma(2) + \cdots + \phi_{kk}\gamma(k) + \sigma_e^2$$

This following equations are used to calculate PACF.

$$\gamma(1) = \phi_{k1}\gamma(0) + \phi_{k2}\gamma(1) + \cdots + \phi_{kk}\gamma(k-1)$$

$$\cdots\cdots\cdots$$

$$\gamma(k) = \phi_{k1}\gamma(k-1) + \phi_{k2}\gamma(k-2) + \cdots + \phi_{kk}\gamma(0)$$

We now use ρ's to substitute γ's

Autocorrelation functions ⟺ Autocovariance functions.

Matrix form of Y-W equations

$$\begin{bmatrix} \gamma(1) \\ \gamma(2) \\ \vdots \\ \gamma(k) \end{bmatrix} = \begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(k-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(k-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(k-1) & \gamma(k-2) & \cdots & \gamma(0) \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} \qquad \begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(k) \end{bmatrix} = \begin{bmatrix} 1 & \rho(1) & \cdots & \rho(k-1) \\ \rho(1) & 1 & \cdots & \rho(k-2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(k-1) & \rho(k-2) & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix}$$

To REMEMBER, just change the diagonal.

Estimating PACF using Y-W equations

2 possible ways to do the "switch"

Cramer's rule

Easier to understand

$$\phi_{kk} = \frac{\det\begin{pmatrix} 1 & \rho_1 & \cdots \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \cdots \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots \rho_1 & \rho_k \end{pmatrix}}{\det\begin{pmatrix} 1 & \rho_1 & \cdots \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \cdots \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots \rho_1 & 1 \end{pmatrix}}$$

Durbin-Levinson algorithm:

Easier to compute with computer

$\phi_{11} = \rho_1,$

$$\phi_{k+1,k+1} = \frac{\rho_{k+1} - \sum_{j=1}^{k} \phi_{kj}\rho_{k+1-j}}{1 - \sum_{j=1}^{k} \phi_{kj}\rho_j},$$

and

$$\phi_{k+1,j} = \phi_{kj} - \phi_{k+1,k+1}\phi_{k,k+1-j}, \quad j = 1, 2, \ldots, k.$$

$\rho(l) = \rho_l, \forall l$

Cramer's rule

For the AR(2) process, the Yule-Walker equations may be written as $\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-1}$, for $k \geq 1$.

- Given a set of ACFs, we can solve $\phi_{11}, \phi_{22}, \phi_{33}$ based on Yule-Walker equations:

Calculate for lag 1. (plug in k=1)

$X_t = \phi_{11} X_{t-1} + a_t$

$X_{t-1} X_t = \phi_{11} X_{t-1}^2 + a_t X_{t-1}$ => causal, no overlap zero

Take expectation: $\gamma(1) = \phi_{11}\gamma(0) + 0$

$\phi_{11} = \frac{\gamma(1)}{\gamma(0)} = \rho_1$

$$\phi_{11} = \rho_1 = \frac{\phi_1}{1-\phi_2} \qquad \phi_{22} = \frac{\det\begin{bmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{bmatrix}}{\det\begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}} = \frac{\rho_2 - \rho_1^2}{1-\rho_1^2} = \phi_2$$

lag 2 (k=2)

$X_t = \phi_{21} X_{t-1} + \phi_{22} X_{t-2} + a_t$

$$\begin{cases} \rho_1 = \phi_{21}\rho_0 + \phi_{22}\rho_1 \\ \rho_2 = \phi_{21}\rho_1 + \phi_{22}\rho_0 \end{cases}$$

since $\frac{\gamma(0)}{\gamma(0)} = \rho_0 = 1$

$$\Rightarrow \begin{cases} \rho_1 = \phi_{21} + \phi_{22}\rho_1 \\ \rho_2 = \phi_{21}\rho_1 + \phi_{22} \end{cases}$$

$$\Rightarrow \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix} \begin{bmatrix} \phi_{21} \\ \phi_{22} \end{bmatrix}$$

$$\Rightarrow \phi_{22} = \frac{\det\begin{bmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{bmatrix}}{\det\begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}} = \frac{\rho_2 - \rho_1^2}{1-\rho_1^2} = \phi_2$$

Cramer's rule (cont'd)

try to derive this

$$\phi_{33} = \frac{\det\begin{bmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{bmatrix}}{\det\begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix}} = \frac{\det\begin{bmatrix} 1 & \rho_1 & \phi_1 + \phi_2\rho_1 \\ \rho_1 & 1 & \phi_1\rho_1 + \phi_2 \\ \rho_2 & \rho_1 & \phi_1\rho_2 + \phi_2\rho_1 \end{bmatrix}}{\det\begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix}} = 0 \qquad \phi_{kk} = 0,\ k \geq 3$$

This example confirm our proof that for an AR(p) model, PACF at lag k equals zero if k is greater than p, where k and p are integers.

D-L algorithm

can
(ignore this)

Example: Wei (2006)

- $X_t = \{13,8,15,4,4,12,11,7,14,12\}$
- data<-c(13,8,15,4,4,12,11,7,14,12)
- sacf<-as.vector(acf(data,plot=F)$acf)
- 1.00000000 -0.18750000 -0.20138889 0.18055556 -0.13194444 -0.32638889 0.11805556 -0.04861111 0.05555556 0.04166667

$$\phi_{11} = \rho_1,$$

$$\phi_{k+1,k+1} = \frac{\rho_{k+1} - \sum_{j=1}^{k} \phi_{kj}\rho_{k+1-j}}{1 - \sum_{j=1}^{k} \phi_{kj}\rho_j},$$

and

$$\phi_{k+1,j} = \phi_{kj} - \phi_{k+1,k+1}\phi_{k,k+1-j}, \quad j = 1, 2, \ldots, k.$$

D-L algorithm (cont'd)

Step 1: initialization

- $\phi_{11} = \rho_1 = -0.188$

Step 2: k=1

$$\phi_{22} = \phi_{1+1,1+1} = \frac{\rho_{1+1} - \sum_{j=1}^{1} \phi_{1j}\rho_{1+1-j}}{1 - \sum_{j=1}^{1} \phi_{1j}\rho_j} = \frac{\rho_2 - \phi_{11}\rho_{1+1-1}}{1 - \phi_{11}\rho_1} = -0.245$$

$$\phi_{1+1,1} = \phi_{11} - \phi_{1+1,1+1} \cdot \phi_{1,1+1-1} = -0.234$$

2,1

D-L algorithm (cont'd)

Step 3: k=2

$$\phi_{33} = \phi_{2+1,2+1} = \frac{\rho_{2+1} - \sum_{j=1}^{2} \phi_{2j}\rho_{2+1-j}}{1 - \sum_{j=1}^{2} \phi_{2j}\rho_j} = \frac{\rho_3 - \phi_{21}\rho_{2+1-1} - \phi_{22}\rho_{2+1-2}}{1 - \phi_{21}\rho_1 - \phi_{22}\rho_2} = 0.097$$

$$\phi_{k+1,j} = \phi_{kj} - \phi_{k+1,k+1} \cdot \phi_{k,k+1-j}, \ j = 1,2$$

$$(k = 2, j = 1), \phi_{2+1,1} = \phi_{21} - \phi_{2+1,2+1}\phi_{2,1+1-1} = ?$$

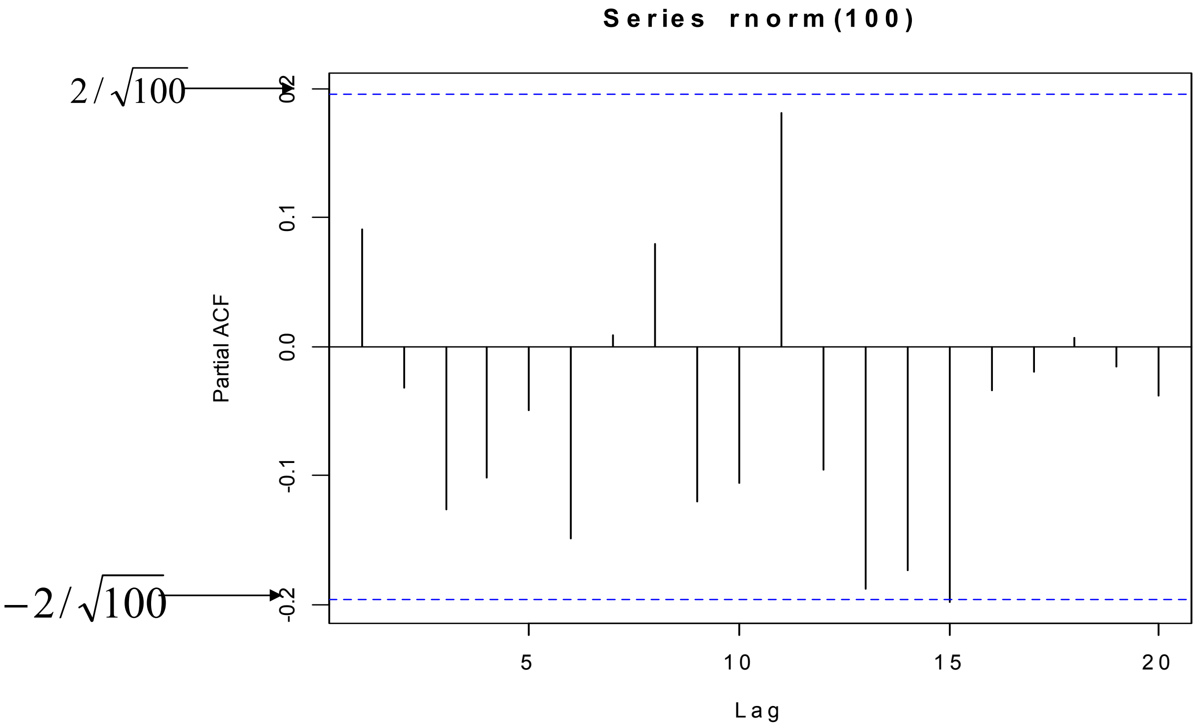
$$(k = 2, j = 2), \phi_{2+1,2} = \phi_{22} - \phi_{2+1,2+1}\phi_{2,2+1-1} = ?$$

Distribution of sample PACF

Under the hypothesis that the underlying process is white noise sequence, sample PACF are normally distributed with $var\left(\widehat{\phi_{kk}}\right) = 1/n$ asymptotically.

- Hence, $\pm\, {}^{2}/_{\sqrt{n}}$ can be used as critical limits (95% confidence level) to test for the hypothesis of a white noise process.

Simulation example and Sample autocorrelation function



In midterm

Solving ACF of autoregressive process using Y-W equations

$$X_t = \phi \cdot X_{t-1} + a_t, \qquad (*)$$

k=1: multiply X_{t-1} on both sides of (*) and take expectation on both sides of the equation

Application of Y-W:

① use data & ACF ⇒ PACF

② solve ACF

③ ...?

$$X_t X_{t-1} = \phi X_{t-1}^2 + X_{t-1} a_t$$

Take Expecation:

$$E(X_t X_{t-1}) = \phi \cdot Var(X_t)$$

$$\Rightarrow \gamma(1) = \phi \cdot \gamma(0)$$

Solving ACF of autoregressive process using Y-W equations

$$X_t X_{t-1} = \phi X_{t-1}^2 + X_{t-1} a_t$$

Take Expecation:

$$\boxed{E(X_t X_{t-1})} = \boxed{\phi \cdot Var(X_t)}$$

$$\Rightarrow \gamma(1) = \phi \cdot \gamma(0)$$

$$X_{t-1} = \sum_{j=0}^{\infty} \phi^j \cdot a_{t-1-j}$$

$$\operatorname{cov}(a_t, X_{t-1}) = \operatorname{cov}(a_t, \sum_{j=0} \phi^j a_{t-1-j}) = 0$$

Solving ACF of autoregressive process using Y-W equations

$$X_t = \phi \cdot X_{t-1} + a_t, \qquad (*)$$

k=2: multiply X_{t-2} on both sides of (*) and take expectation on both sides of the equation

$$X_t X_{t-2} = \phi X_{t-1} X_{t-2} + X_{t-2} a_t$$

Take Expecation:

$$E(X_t X_{t-2}) = \phi \cdot E(X_{t-1} X_{t-2})$$

$$\Rightarrow \gamma(2) = \phi \cdot \gamma(1)$$

Using the result that $\gamma(1) = \phi \cdot \gamma(0)$

$$\Rightarrow \gamma(2) = \phi \cdot \gamma(1) = \phi^2 \cdot \gamma(0)$$

Solving ACF of autoregressive process using Y-W equations

For **k** ≥**3**, similarly we have

$$X_t X_{t-k} = \phi X_{t-1} X_{t-k} + X_{t-k} a_t$$

Take Expecation:

$$E(X_t X_{t-k}) = \phi \cdot E(X_{t-1} X_{t-k})$$

$$\Rightarrow \gamma(k) = \phi \cdot \gamma(k-1)$$

$$\Rightarrow \cdots\cdots$$

$$\Rightarrow \boxed{\gamma(k) = \phi^k \gamma(0)}$$

Revisit stationary AR(p) processes

The Yule-Walker equations of an AR(p) model $\rho(k) = \phi_1\rho(k-1) + \cdots + \phi_p(k-p)$, for all $k \geq 0$.

It is a set of difference equations and may have the general solution $\rho(k) = A_1(\alpha_1)^k + A_2(\alpha_2)^k + \cdots + A_p\left(\alpha_p\right)^k$, where $\{\alpha_i\}$ are the roots of the characteristic equation $z^p - \phi_1 y^{p-1} - \cdots - \phi_p = 0$.

The constants $\{A_i\}$ are chosen to satisfy some initial conditions, such as $|\rho(0)| = 1$.

From the general form of the solution, $\rho(k)$ tends to zero as k increases provided $|\alpha_i| < 1$ for all i, and this is a necessary and sufficient condition for the process to be stationary.