时间序列分析

Time Series Analysis

Lecture 3

Review

1. Examples of Time Series Models
2. Measure of Time Series Dependence
3. Stationarity, Ergodicity, LLN and CLT
4. Estimation of Covariances
5. Estimation of Long-run Covariances
6. Nonstationarity
7. Classical Regression with TS Data
8. Time Regressions

Today's Topics

1. Linear Time Series
2. Simple AR Models
3. Simple MA Models
4. Simple ARMA Models

1. Linear Time Series Models

A time series r_t is said to be linear if it can be written as

$$r_t = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i}, \qquad \psi_0 = 1,$$

where a_t is a white noise series.

$$E(r_t) = \mu, \qquad Var(r_t) = \sigma_a^2 \sum_{i=0}^{\infty} \psi_i^2.$$

Autocovariance and Autocorrelation

$$
\begin{aligned}
\gamma_\ell = Cov(r_t, r_{t-\ell}) &= E\left[\left(\sum_{i=0}^{\infty}\psi_i a_{t-i}\right)\left(\sum_{j=0}^{\infty}\psi_j a_{t-\ell-j}\right)\right] \\
&= E\left(\sum_{i,j=0}^{\infty}\psi_i\psi_j a_{t-i}a_{t-\ell-j}\right) \\
&= \sum_{j=0}^{\infty}\psi_{\ell+j}\psi_j E(a_{t-\ell-j}^2) = \sigma_a^2\sum_{j=0}^{\infty}\psi_{\ell+j}\psi_j. \\
\rho_\ell = \frac{\gamma_\ell}{\gamma_0} &= \frac{\sum_{i=0}^{\infty}\psi_{\ell+i}\psi_i}{1+\sum_{i=1}^{\infty}\psi_i^2}, \quad \ell \geq 0.
\end{aligned}
$$

Stationarity Condition

$$\sum_{i=0}^{\infty} \psi_i^2 \quad < \infty,$$

$$\sum_{i=0}^{\infty} \psi_{\ell+i}\psi_i \quad < \infty.$$

This implies that

$$\psi \to 0 \quad as \quad i \to \infty$$

ρ_ℓ converges to zero as ℓ increases.

2. Simple AR Models

- AR(1)

$$\begin{aligned} r_t &= \phi_0 + \phi_1 r_{t-1} + a_t, \\ E(r_t | r_{t-1}) &= \phi_0 + \phi_1 r_{t-1}, \\ Var(r_t | r_{t-1}) &= Var(a_t) = \sigma_a^2. \end{aligned}$$

- AR(p)

$$r_t = \phi_0 + \phi_1 r_{t-1} + \cdots + \phi_p r_{t-p} + a_t.$$

Properties of AR(1)

$$
\begin{aligned}
r_t &= \phi_0 + \phi_1 r_{t-1} + a_t, \\
E(r_t) &= \phi_0 + \phi_1 E(r_{t-1}).
\end{aligned}
$$

Under the stationary condition, $E(r_t) = E(r_{t-1}) = \mu$ and hence

$$
\mu = \phi_0 + \phi_1 \mu \qquad E(r_t) = \mu = \frac{\phi_0}{1 - \phi_1}.
$$

Using $\phi_0 = (1 - \phi_1)\mu$, the AR(1) model can be rewritten as

$$
\begin{aligned}
r_t - \mu &= \phi_1(r_{t-1} - \mu) + a_t, \\
r_t - \mu &= a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} + \cdots \\
&= \sum_{i=0}^{\infty} \phi_1^i a_{t-i}. \\
E[(r_t - \mu)a_{t+1}] &= 0, \\
Var(r_t) &= \phi_1^2 Var(r_{t-1}) + \sigma_a^2.
\end{aligned}
$$

Under the stationarity assumption,

$$
Var(r_t) = \frac{\sigma_a^2}{1 - \phi_1^2}, \quad \textit{procided that } \phi_1^2 < 1.
$$

ACF of AR(1)

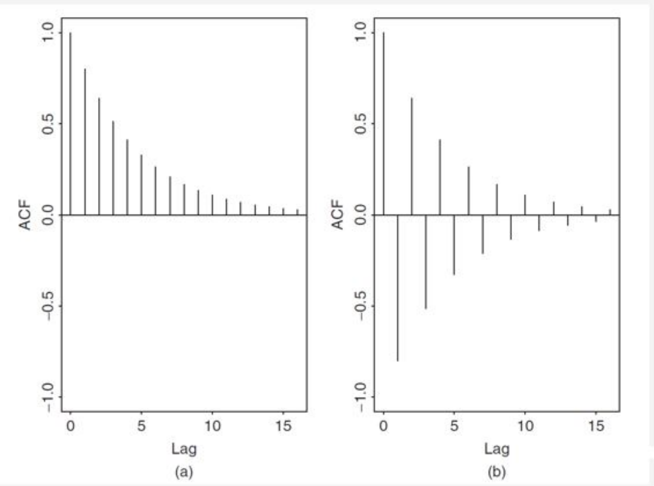
We can easily derive

$$\begin{aligned} Var(r_t) &= \gamma_0 = \frac{\sigma_a^2}{1-\phi_1^2}, \\ \gamma_\ell &= \phi\gamma_{\ell-1}, \quad \textit{for } \ell > 0. \end{aligned}$$

Thus we have

$$\begin{aligned} \rho_\ell &= \phi_1\rho_{\ell-1}, \quad \textit{for } \ell > 0. \\ \rho_0 &= 1. \end{aligned}$$

Autocorrelation function of an AR(1) model



(a) for $\phi_1 = 0.8$ and (b) for $\phi_1 = -0.8$.

Properties of AR(2)

An AR(2) writes

$$r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + a_t.$$

Similarly to AR(1) case, we obtain

$$E(r_t) = \mu = \frac{\phi_0}{1 - \phi_1 - \phi_2}, \quad \textit{provided that } \phi_1 + \phi_2 \neq 1.$$

The model can be rewritten as

$$\begin{aligned} r_t - \mu &= \phi_1(r_{t-1} - \mu) + \phi_2(r_{t-2} - \mu) + a_t, \\ \rho_\ell &= \phi_1 \rho_{\ell-1} + \phi_2 \rho_{\ell-2}, \quad \textit{for} \quad \ell > 0. \end{aligned}$$

ACF of AR(2)

We have

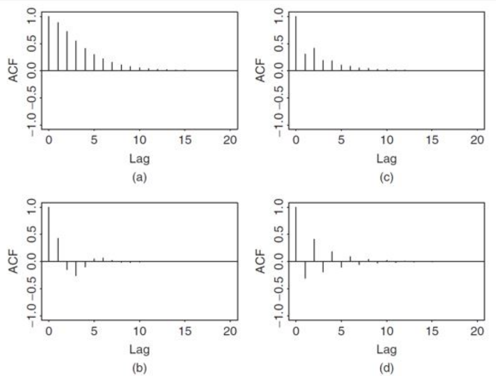
$$\begin{aligned}\rho_1 &= \frac{\phi_1}{1-\phi_2}, \\ \rho_\ell &= \phi_1\rho_{\ell-1} + \phi_2\rho_{\ell-2}, \qquad \ell \geq 2.\end{aligned}$$

Stationarity requires that roots of

$$1 - \phi_1 x - \phi_2 x^2 = 0,$$

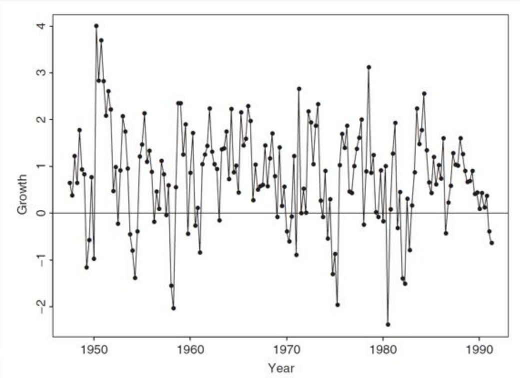
lie outside of unit circle.

ACF of AR(2)



(a) $\phi_1 = 1.2$ and $\phi_2 = -0.35$, (b) $\phi_1 = 0.6$ and $\phi_2 = -0.4$,
(c) $\phi_1 = 0.2$ and $\phi_2 = 0.35$, and (d) $\phi_1 = -0.2$ and $\phi_2 = 0.35$.

Example 3.1: quarterly growth rate of U.S. real GNP, seasonally adjusted, 1947.II-1991.I



$$r_t = 0.0047 + 0.348 r_{t-1} + 0.179 r_{t-2} - 0.142 r_{t-3} + a_t, \qquad \hat{\sigma}_a = 0.0097.$$

```
> gnp=scan(file='dgnp82.txt') % Load data
 % To create a time-series object
> gnp1=ts(gnp,frequency=4,start=c(1947,2))
> plot(gnp1)
> points(gnp1,pch='*')

> m1=ar(gnp,method=''mle'') % Find the AR order
> m1$order    % An AR(3) is selected based on AIC
[1] 3
> m2=arima(gnp,order=c(3,0,0)) % Estimation
> m2
Call:
arima(x = gnp, order = c(3, 0, 0))

Coefficients:
         ar1     ar2      ar3   intercept
      0.3480  0.1793  -0.1423      0.0077
s.e.  0.0745  0.0778   0.0745      0.0012

sigma^2 estimated as 9.427e-05: log likelihood=565.84,
   aic=-1121.68
```

```
% In R, ''intercept'' denotes the mean of the series.
% Therefore, the constant term is obtained below:
> (1-.348-.1793+.1423)*0.0077
[1] 0.0047355
> sqrt(m2$sigma2)  % Residual standard error
[1] 0.009709322

> p1=c(1,-m2$coef[1:3]) % Characteristic equation
> roots=polyroot(p1) % Find solutions
> roots
[1] 1.590253+1.063882i -1.920152+0.000000i 1.590253-1.063882i
> Mod(roots) % Compute the absolute values of the solutions
[1] 1.913308 1.920152 1.913308
```

AR(p)

Mean of a stationary AR(p) is

$$E(r_t) = \frac{\phi_0}{1 - \phi_1 - \cdots - \phi_p}$$

Stationarity requires that roots of

$$1 - \phi_1 x - \phi_2 x^2 - \cdots - \phi_p x^p = 0$$

lie outside of the unit circle.

ACF could be obtained in a similar fashion.

Identifying AR Model

Two methods

(1) Partial Autocorrelation Function (PACF)

(2) Information Criteria

Partial Autocorrelation Function

To introduce PACF, consider

$$
\begin{aligned}
r_t &= \phi_{0,1} + \phi_{1,1} r_{t-1} + e_{1t}, \\
r_t &= \phi_{0,2} + \phi_{1,2} r_{t-1} + \phi_{2,2} r_{t-2} + e_{2t}, \\
r_t &= \phi_{0,3} + \phi_{1,3} r_{t-1} + \phi_{2,3} r_{t-2} + \phi_{3,3} r_{t-3} + e_{3t}, \\
r_t &= \phi_{0,4} + \phi_{1,4} r_{t-1} + \phi_{2,4} r_{t-2} + \phi_{3,4} r_{t-3} + \phi_{4,4} r_{t-4} + e_{4t}, \\
&\vdots
\end{aligned}
$$

which could be estimated by least squares.

Sample PACF: $\hat{\phi}_{1,1}, \hat{\phi}_{2,2}, \hat{\phi}_{3,3}, \cdots$

Properties of PACF of AR(p)

For a stationary Gaussian AR(p) model, it can be shown that the sample PACF has the following properties:

- $\hat{\phi}_{p,p}$ converges to ϕ_p as the sample size T goes to infinity.
- $\hat{\phi}_{\ell,\ell}$ converges to zero for all $\ell > p$.
- The asymptotic variance of $\hat{\phi}_{\ell,\ell}$ is $1/T$ for $\ell > p$.

These results say that, for an AR(p) series, the sample PACF cuts off at lag p.

Example 3.2

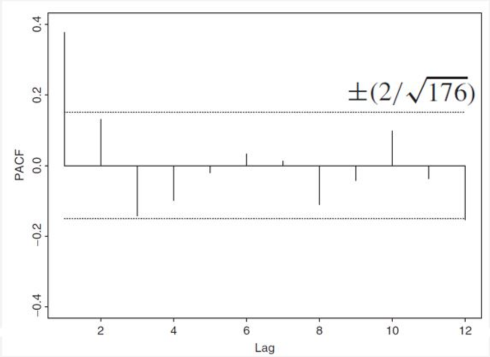
Sample Partial Autocorrelation Function and Some Information Criteria for the Monthly Simple Returns of CRSP Value-Weighted Index, Jan 1926-Dec 2008

p	1	2	3	4	5	6
PACF	0.115	−0.030	−0.102	0.033	0.062	−0.050
AIC	−5.838	−5.837	−5.846	−5.845	−5.847	−5.847
BIC	−5.833	−5.827	−5.831	−5.825	−5.822	−5.818
p	7	8	9	10	11	12
PACF	0.031	0.052	0.063	0.005	−0.005	0.011
AIC	−5.846	−5.847	−5.849	−5.847	−5.845	−5.843
BIC	−5.812	−5.807	−5.805	−5.798	−5.791	−5.784

With $T = 996$, the asymptotic standard error of the sample PACF is approximately 0.032. Therefore, using the 5% significant level, we identify an AR(3) or AR(9) model for the data (i.e., $p = 3$ or 9). If the 1% significant level is used, we specify an AR(3) model.

Example 3.1

Sample partial autocorrelation function of U.S. quarterly real GNP growth rate from 1947.II to 1991.I. Dotted lines give approximate pointwise 95% confidence interval.



Information Criteria

- Akaike information criterion (AIC) (Akaike, 1973) for Gaussian AR(p) is

$$AIC(\ell) = \ln(\tilde{\sigma}_\ell^2) + \frac{2\ell}{T}$$

- Schwarz – Bayesian information criterion (BIC) for Gaussian AR(p) is

$$BIC(\ell) = \ln(\tilde{\sigma}_\ell^2) + \frac{\ell \ln(T)}{T}$$

Selection Rule

- To use AIC to select an AR model in practice, one computes $AIC(\ell)$ for $\ell = 0, \cdots, P$, where P is a prespecified positive integer and selects the order k that has the minimum AIC value. The same rule applies to BIC.
- There is no evidence to suggest that one approach outperforms the other in a real application.

Example 3.2 Revisited

p	1	2	3	4	5	6
PACF	0.115	−0.030	−0.102	0.033	0.062	−0.050
AIC	−5.838	−5.837	−5.846	−5.845	−5.847	−5.847
BIC	−5.833	−5.827	−5.831	−5.825	−5.822	−5.818
p	7	8	9	10	11	12
PACF	0.031	0.052	0.063	0.005	−0.005	0.011
AIC	−5.846	−5.847	−5.849	−5.847	−5.845	−5.843
BIC	−5.812	−5.807	−5.805	−5.798	−5.791	−5.784

The AIC values are close to each other with minimum -5.849 occurring at $p = 9$, suggesting that an AR(9) model is preferred by the criterion. The BIC, on the other hand, attains its minimum value -5.833 at $p = 1$ with -5.831 as a close second at $p = 3$. Thus, the BIC selects an AR(1) model for the value-weighted return series.

Example 3.1 Revisited

The AIC obtained from R also identifies an AR(3) model. Note that the AIC value of the ar command in R has been adjusted so that the minimum AIC is zero.

```
> gnp=scan(file='q-gnp4791.txt')
> ord=ar(gnp,method=''mle'')
> ord$aic
 [1] 27.847 2.742 1.603 0.000 0.323 2.243
 [7]  4.052 6.025 5.905 7.572 7.895 9.679
> ord$order
[1] 3
```

Estimation of AR(p)

For a specified AR(p) model, the conditional least-squares method, which starts with the $(p+1)$th observation, is often used to estimate the parameters.

$$r_t = \phi_0 + \phi_1 r_{t-1} + \cdots + \phi_p r_{t-p} + a_t, \quad t = p+1, \cdots, T.$$

Denote the estimate of ϕ_i by $\hat{\phi}_i$. The fitted model is

$$r_t = \hat{\phi}_0 + \hat{\phi}_1 r_{t-1} + \cdots + \hat{\phi}_p r_{t-p},$$

and the associated residual is

$$\hat{a}_t = r_t - \hat{r}_t.$$

Example 3.2 Revisted

For illustration, consider an AR(3) model for the monthly simple returns of the value-weighted index. The fitted model is

$$r_t = 0.0091 + 0.11 r_{t-1} - 0.019 r_{t-2} - 0.104 r_{t-3} + \hat{a}_t, \quad \sigma_a^2 = 0.054.$$

```
> vw=read.table('m-ibm3dx2608.txt',header=T)[,3]
> m3=arima(vw,order=c(3,0,0))
> m3
Call:
arima(x = vw, order = c(3, 0, 0))

Coefficients:
         ar1      ar2      ar3  intercept
      0.1158  -0.0187  -0.1042     0.0089
s.e.  0.0315   0.0317   0.0317     0.0017

sigma^2 estimated as 0.002875: log likelihood=1500.86,
   aic=-2991.73

> (1-.1158+.0187+.1042)*mean(vw) % Compute
   the intercept phi(0).
[1] 0.00896761
> sqrt(m3$sigma2)  % Compute standard error of residuals
[1] 0.0536189
```

Model Diagnostic Check

- A fitted model must be examined carefully to check for possible model inadequacy.
- If the model is adequate, then the residual series should behave as a white noise.
- The ACF and the Ljung – Box statistics of the residuals can be used to check the closeness of the estimated residuals to a white noise.

Ljung-Box Q(m) follows asymptotically a chi-squared distribution with $m - g$ degrees of freedom, where g denotes the number of AR coefficients used in the model.

$$Q(m) = T(T+2)\sum_{\ell=1}^{m} \frac{\rho_\ell^2}{T-\ell}.$$

where $\hat{\rho}_\ell$ denotes the residual sample ACF of order ℓ.

```
> Box.test(m3$residuals,lag=12,type='Ljung')
      Box-Ljung test

data: m3$residuals   % R uses 12 degrees of freedom

X-squared = 16.3525, df = 12, p-value = 0.1756

> pv=1-pchisq(16.35,9) % Compute p-value using 9 degrees
   of freedom
> pv
[1] 0.05992276
 % To fix the AR(2) coef to zero:
> m3=arima(vw,order=c(3,0,0),fixed=c(NA,0,NA,NA))
 % The subcommand 'fixed' is used to fix parameter values,
 % where NA denotes estimation and 0 means fixing the
    parameter to 0.
 % The ordering of the parameters can be found using m3$coef.
> m3
Call:
arima(x = vw, order = c(3, 0, 0), fixed = c(NA, 0, NA, NA))

Coefficients:
          ar1  ar2      ar3  intercept
       0.1136    0  -0.1063     0.0089
s.e.   0.0313    0   0.0315     0.0017
```

```
sigma^2 estimated as 0.002876: log likelihood=1500.69,
   aic=-2993.38
> (1-.1136+.1063)*.0089 % Compute phi(0)
[1] 0.00883503
> sqrt(m3$sigma2)  % Compute residual standard error
[1] 0.05362832

> Box.test(m3$residuals,lag=12,type='Ljung')

      Box-Ljung test

data: m3$residuals
X-squared = 16.8276, df = 12, p-value = 0.1562

> pv=1-pchisq(16.83,10)
> pv
[1] 0.0782113
```

Forecasting AR(p)

The ℓ-step forecast at forecast origin h, $\hat{\gamma}_h(\ell)$ is chosen such that

$$E\{[r_{h+\ell} - \hat{\gamma}_h(\ell)]^2 | F_h\} \leq \min_g E\{[r_{h+\ell} - g]^2 | F_h\},$$

where g is a function of the information available at time h (inclusive), that is, a function of F_h.

1-Step-Ahead Forecast

From AR(p), we obtain

$$r_{h+1} = \phi_0 + \phi_1 r_h + \cdots + \phi_p r_{h+1-p} + a_{h+1}.$$

Under the minimum squared error loss function, the point forecast given F_h is the conditional expectation

$$\hat{r}_h(1) = E(r_{h+1}|F_h) = \phi_0 + \sum_{i=1}^{p} \phi_i r_{h+1-i}.$$

The forecast error is

$$\begin{aligned} e_h(1) &= r_{h+1} - \hat{r}_h(1) = a_{h+1}, \\ Var[e_h(1)] &= Var(a_{h+1}) = \sigma_a^2. \end{aligned}$$

2-Step-Ahead Forecast

$$
\begin{aligned}
r_{h+2} &= \phi_0 + \phi_1 r_{h+1} + \cdots + \phi_p r_{h+2-p} + a_{h+2}. \\
\hat{r}_h(2) &= E(r_{h+2}|F_h) = \phi_0 + \phi_1 \hat{r}_h(1) + \phi_2 r_h + \cdots + \phi_h r_{h+2-p}. \\
e_h(2) &= r_{h+2} - \hat{r}_h(2) = \phi_1[r_{h+1} - \hat{r}_h(1)] + a_{h+2} \\
&= a_{h+2} + \phi_1 a_{h+1}. \\
Var[e_h(2)] &= (1 + \phi_1^2)\sigma_a^2. \\
Var[e_h(2)] &\geq Var[e_h(1)]
\end{aligned}
$$

Multistep-Ahead Forecast

$$
\begin{aligned}
r_{h+\ell} &= \phi_0 + \phi_1 r_{h+\ell+1} + \cdots + \phi_p r_{h+\ell-p} + a_{h+\ell}. \\
\hat{r}_h(\ell) &= \phi_0 + \sum_{i=1}^{p} \phi_i \hat{r}_h(\ell - i) \xrightarrow{p} E(r_t) \ as \ \ell \to \infty \\
\hat{r}_h(i) &= r_{h+i} \quad if \ i \leq 0 \\
e_h(\ell) &= r_{h+\ell} - \hat{r}_h(\ell)
\end{aligned}
$$

Example 3.2 Revisited

Multistep Ahead Forecasts of an AR(3) Model for Monthly Simple Returns of CRSP Value-Weighted Index

Step	1	2	3	4	5	6
Forecast	0.0076	0.0161	0.0118	0.0099	0.0089	0.0093
Std. Error	0.0534	0.0537	0.0537	0.0540	0.0540	0.0540
Actual	−0.0623	−0.0220	−0.0105	0.0511	0.0238	−0.0786
Step	7	8	9	10	11	12
Forecast	0.0095	0.0097	0.0096	0.0096	0.0096	0.0096
Std. Error	0.0540	0.0540	0.0540	0.0540	0.0540	0.0540
Actual	−0.0132	0.0110	−0.0981	−0.1847	−0.0852	0.0215

[a] The forecast origin is $h = 984$.

Plot of 1- to 12-step-ahead out-of-sample forecasts for monthly simple returns of CRSP value-weighted index. Forecast origin is $t = 984$, which is Dec 2007. Forecasts are denoted by $\circ$ and actual observations by $\bullet$. Two dashed lines denote two standard error limits of the forecasts.