《Analysis of Financial Time Series》

Ruey S. Tsay

ch 2.7-2.11 Linear Time Series Analysis and Its Applications

April 18, 2022

Outline

1. *Ch2 Linear Time Series Analysis and Its Application*
    - Unit-root Nonstationarity
    - Seasonal time Series
    - Regression Models with Time Series Errors
    - Consistent covariance matrix estimation
    - Long-memory models
    - Summary of the chapter

Unit-root Nonstationarity

Random walk

- Form: $p_t = p_{t-1} + a_t$
- If a_t has a symmetric distribution around zero, then conditional on p_{t-1}, p_t has a 50-50 chance to go up or down, implying that p_t would go up or down at random.
- Unit root? It is an AR(1) model with coefficient $\phi_1 = 1$.
- Nonstationary: Why? Because the coefficient of p_{t-1} is unity, it does not satisfy the weak stationarity condition of an AR(1) model.

- The random-walk model has widely been considered as a statistical model for the movement of logged stock prices.
- Under such a model, the stock price is not predictable or mean reverting.
  - The 1-step-ahead forecast at the forecast origin h is

    $$\hat{p}_h(1) = E(p_{h+1}|p_h, p_{h-1}, \ldots) = p_h.$$

    which is the log price of the stock at the forecast origin.
  - The 2-step-ahead forecast is

    $$\begin{aligned}\hat{p}_h(2) &= E(p_{h+2}|p_h, p_{h-1}, \ldots) = E(p_{h+1} + a_{h+2}|p_h, p_{h-1}, \ldots) \\ &= E(p_{h+1}|p_h, p_{h-1}, \ldots) = \hat{p}_h(1) = p_h.\end{aligned}$$

    Such a forecast has no practical value.
  - In fact, for any forecast horizon $l > 0$, we have $\hat{p}_h(l) = p_h$. Therefore, the process is not mean reverting.

- The MA representation of the random-walk model is

$$p_t = a_t + a_{t-1} + a_{t-2} + \cdots .$$

This representation has several important practical implications.

(1)
- The l-step-ahead forecast error is

$$e_h(l) = a_{h+l} + \cdots + a_{h+1},$$

so that $Var[e_h(l)] = l\sigma_a^2$, which diverges to infinity as $l \to \infty$.
- This result says that the usefulness of point forecast $\hat{p}_h(l)$ diminishes as l increases, which again implies that the model is not predictable.

(2)

- The unconditional variance of p_t is unbounded because $Var[p_t] = t\sigma_a^2 \to \infty$ as t increases.
- Theoretically, this means that p_t can assume any real value for a sufficiently large t.
- For the log price p_t of an individual stock, this is plausible.
- For market indexes, negative log price is very rare if it happens at all.
- The adequacy of a random-walk model for market indexes is questionable.

(3)

- From the representation, $\psi_i = 1$ for all i.
- The impact of any past shock a_{t-i} on p_t does not decay over time. The series has a strong memory as it remembers all of the past shocks.
- In economics, the shocks are said to have a permanent effect on the series.
- The strong memory of a unit-root time series can be seen from the sample ACF approaches 1 for any finite lag.

Random walk with drift

- Form: $p_t = \mu + p_{t-1} + a_t$, $\mu \neq 0$.
- Has a time trend with slope μ. Why?

$$\begin{aligned} p_t &= \mu + p_{t-1} + a_t \\ &= 2\mu + p_{t-2} + a_{t-1} + a_t \\ &= t\mu + p_0 + \sum_{i=1}^{t} a_i \end{aligned}$$

  The last equation shows that the log price consists of a time trend $t\mu$ and a pure random-walk process $\sum_{i=1}^{t} a_i$.
- Because $Var(\sum_{i=1}^{t} a_i) = t\sigma_a^2$, the conditional standard deviation of p_t is $\sqrt{t}\sigma_a$, which grows at a slower rate than the conditional expectation of p_t.
- Therefore, if we graph p_t against the time index t, we have a time trend with slope μ.

The effect of the drift parameter

- Consider the monthly log stock returns of the 3M Company from February 1946 to December 2008.
- As shown by the sample EACF in Table 2.5, the series has no significant serial correlation.
- The series thus follows the simple model

$$r_t = 0.0103 + a_t, \quad \hat{\sigma}_a = 0.0637.$$

- The mean of the monthly log returns of 3M stock is, therefore, significantly different from zero at the 1% level.

- We use the log return series to construct two log price series, namely

$$p_t = \sum_{i=1}^{l} r_i \quad \text{and} \quad p_t^* = \sum_{i=1}^{l} a_i,$$

where a_i is the mean-corrected log return (i.e., $a_t = r_t - 0.0103$).

- The $p_t = \sum_{i=1}^{l} r_i$ is the log price of 3M stock, assuming that the initial log price is zero ($p_0 = 0$).
- The $p_t^* = \sum_{i=1}^{l} a_i$ is the corresponding mean-corrected log price ($E(a_i) = 0$).
- Figure 2.10 shows the time plots of p_t and p_t^* as well as a straight line $y_t = 0.0103 \times (t - 1946)$, where t is the time sequence of the returns and 1946 is the starting year of the stock.
- From the plots,
    - The importance of the constant 0.0103 is evident.
    - It represents the slope of the upward trend of p_t.

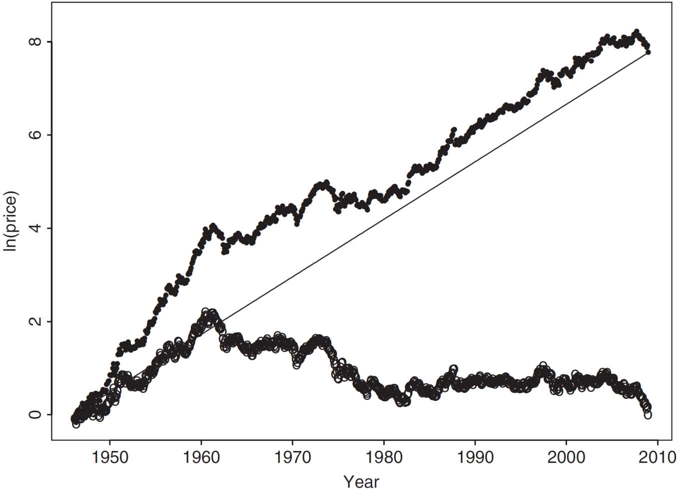


Figure 2.10 Time plots of log prices for 3M stock from February 1946 to December 2008, assuming that log price of January 1946 was zero. The "o" line is for log price without time trend. Straight line is $y_t = 0.0103 \times t + 1946$. 修改為 y_t = 0.0103 (t-1946)

Interpretation of the Constant Term

- Meaning of a constant term in a time series model:
  - For an MA(q) model, the constant term is simply the mean of the series.
  - For a stationary AR(p) model or ARMA(p, q) model, the constant term is related to the mean.
  - For a random walk with drift, the constant term becomes the time slope of the series.
- Difference between dynamic and regression models:
  - Consider an AR(1) model and a simple linear regression model

$$r_t = \phi_0 + \phi_1 r_{t-1} + a_t \quad \text{and} \quad y_t = \beta_0 + \beta_1 x_t + a_t.$$

  - For the AR(1) model to be meaningful, the coefficient ϕ_1 must satisfy $|\phi_1| \le 1$.
  - The coefficient β_1 can assume any fixed real number.

General Unit-Root Nonstationary Models

- Consider an **ARMA** model. If one extends the model by allowing the AR polynomial to have 1 as a characteristic root, then the model becomes the well-known autoregressive integrated moving-average (**ARIMA**) model.
- An ARIMA model is said to be unit-root nonstationary because its AR polynomial has a unit root.
- An ARIMA model has **strong memory** because the ψ_1 coefficients in its MA representation do not decay over time to zero.
- A conventional approach for handling unit-root nonstationarity is to use ***differencing***.

Differencing

- A time series y_t is said to be an **ARIMA**$(p, 1, q)$ process if the change series $c_t = (1 - B)y_t$ follows a stationary and invertible **ARMA**(p, q) model.
- In finance, price series are commonly believed to be nonstationary, but the log return series,

$$r_t = \ln(P_t) - \ln(P_{t-1}),$$

  is stationary. In this case, the log price series is unit-root nonstationary and hence can be treated as an ARIMA process.
- The idea of transforming a nonstationary series into a stationary one by considering its change series is called **differencing** in the time series literature.

- 1st difference: $r_t = p_t - p_{t-1}$
  - If p_t is the log price, then the 1st difference is simply the log return.
  - Typically, 1st difference means the "change" or "increment" of the original series.
- Seasonal difference: $y_t = p_t - p_{t-s}$, where s is the periodicity.
  - e.g. $s = 4$ for quarterly series and $s = 12$ for monthly series.
  - If p_t denotes quarterly earnings, then y_t is the change in earning from the same quarter one year before.

Unit-root test

- Let p_t be the log price of an asset.
- To test whether the log price p_t of an asset follows a random walk or a random walk with drift, we employ the models

$$p_t = \phi_1 p_{t-1} + e_t$$
$$p_t = \phi_0 + \phi_1 p_{t-1} + e_t.$$

where e_t denotes the error term.

- Consider the hypothesis

$$H_0 : \phi_1 = 1 \quad H_a : \phi_1 < 1.$$

This is the well-known unit-root testing problem; see Dickey and Fuller (1979).

- Dickey-Fuller test is the usual t-ratio of the OLS estimate of ϕ_1 being 1.

$$
\begin{aligned}
\frac{\hat{\phi}_1 - 1}{\mathrm{std}(\hat{\phi}_1)} &= \frac{\sum_{t=1}^{T} p_{t-1}(p_t - p_{t-1})}{\mathrm{std}(\hat{\phi}_1)\sum_{t=1}^{T} p_{t-1}^2} \\
&= \frac{\sum_{t=1}^{T} p_{t-1} e_t}{\hat{\sigma}_e \sqrt{\sum_{t=1}^{T} p_{t-1}^2}}
\end{aligned}
$$

where

$$
\hat{\phi}_1 = \frac{\sum_{t=1}^{T} p_{t-1} p_t}{\sum_{t=1}^{T} p_{t-1}^2}, \quad \hat{\sigma}_e^2 = \frac{\sum_{t=1}^{T} (p_t - \hat{\phi}_1 p_{t-1})^2}{T-1}.
$$

- The t-ratio, however, has a nonstandard limiting distribution.
  - 檢定統計量的極限分佈為 skew 分佈，當 $\phi_0 = 0$ 。
  - 檢定統計量的極限分佈為常態分佈，當 $\phi_0 \neq 0$ 。

- To verify the existence of a unit root in an AR(p) process, one may perform the test $H_0 : \beta = 1$ vs. $H_a : \beta < 1$ using the regression

$$p_t = c_t + \beta p_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta p_{t-i} + e_t$$

where c_t is a deterministic function of the time index t and $\Delta p_t = p_t - p_{t-1}$.

- The above equation can also be rewritten as

$$\Delta p_t = c_t + \beta_c p_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta p_{t-i} + e_t$$

where $\beta_c = \beta - 1$. And perform the test

$$H_0 : \beta_c = 0 \quad vs. \quad H_a : \beta_c < 0$$

- Then, the augmented (增強) Dickey–Fuller (ADF) unit-root test for an AR(p) model:

$$\text{ADF-test} = \frac{\hat{\beta} - 1}{\text{std}(\hat{\beta})},$$

where $\hat{\beta}$ denotes the least-squares estimate of β.

- Again, the statistic has a non-standard limiting distribution.

Example 1

- Consider the log series of U.S. quarterly GDP (Gross Domastic Product 國內生產毛額) from 1947.I to 2008.IV. The series is shown in Figure 2.11.
- The series exhibits an upward trend, showing the growth of the U.S. economy, and has high sample serial correlations.
- The first differenced series seems to vary around a fixed mean level, even though the variability appears to be smaller in recent years.
- To confirm the observed phenomenon, we apply the ADF unit-root test to the log series.
  - Based on the sample PACF of the differenced series, we choose $p = 10$. Other values of p are also used, but they do not alter the conclusion of the test.
  - With $p = 10$, the ADF test statistic is -1.6109 with a p value 0.4569. $\Rightarrow$ the unit-root hypothesis cannot be rejected.

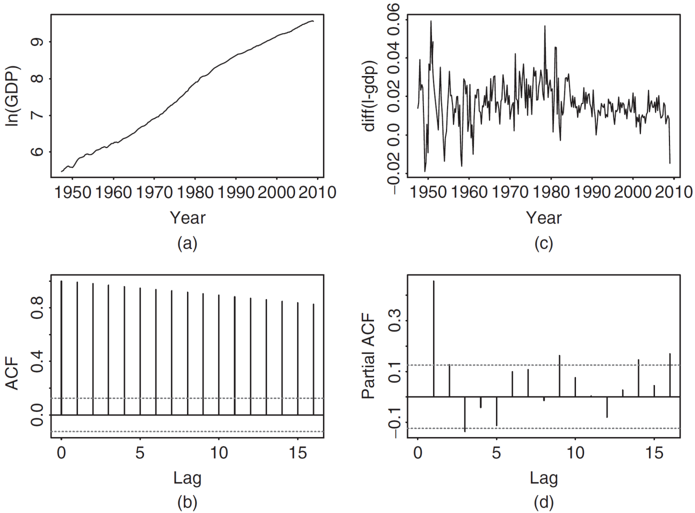


Figure 2.11 Log series of U.S. quarterly GDP from 1947.I to 2008.IV: (a) time plot of logged GDP series, (b) sample ACF of log GDP data, (c) time plot of first differenced series, and (d) sample PACF of differenced series.

R Demonstration

```
> library(fUnitRoots)
> da=read.table("q-gdp4708.txt",header=T)
> gdp=log(da[,4])
> m1=ar(diff(gdp),method='mle')
> m1$order
[1] 10
> adfTest(gdp,lags=10,type=c("c"))
Title:
 Augmented Dickey-Fuller Test

Test Results:
 PARAMETER:
  Lag Order: 10
 STATISTIC:
  Dickey-Fuller: -1.6109
 P VALUE:  0.4569
```

ADF test R

- adfTest(x, lags = 1, type = c("nc", "c", "ct"), title = NULL, description = NULL)
- **x:** a numeric vector or time series object.
  lags: the maximum number of lags used for error term correction.
  title: a character string which allows for a project title.
  type: a character string describing the type of the unit root regression. Valid choices are "nc" for a regression with no intercept (constant) nor time trend, and "c" for a regression with an intercept (constant) but no time trend, "ct" for a regression with an intercept (constant) and a time trend. The default is "c".
  description: a character string which allows for a brief description.

Example 2

- Consider the log series of the S&P 500 index from January 3, 1950, to April 16, 2008, for 14,462 observations.
- The series is shown in Figure 2.12.
- Testing for a unit root in the index is relevant if one wishes to verify empirically that the Index follows a random walk with drift.
    - Use $c_t = \omega_0 + \omega_1 t$ in applying the ADF test.
    - Choose $p = 15$ based on the sample PACF of the first differenced series.
    - The resulting test statistic is -1.9946 with a p value 0.5807.
    - Thus, the unit-root hypothesis cannot be rejected at any reasonable significance level.

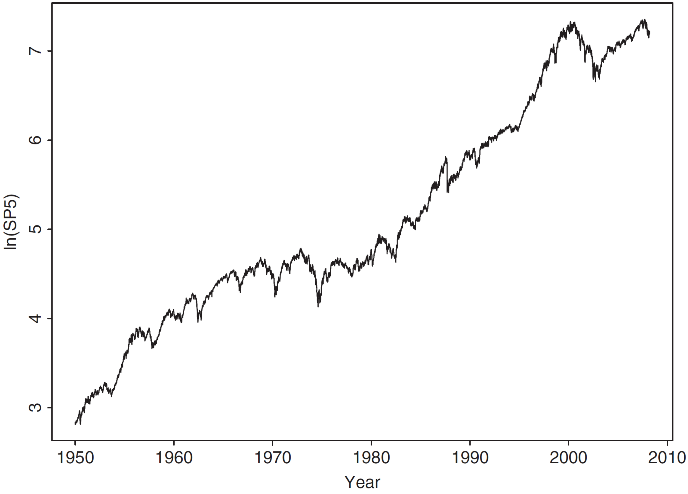


Figure 2.12 Time plot of logarithm of daily S&P 500 index from January 3, 1950, to April 16, 2008.

R Demonstration

```
> library(fUnitRoots)
> da=read.table("d-sp55008.txt",header=T)
> sp5=log(da[,7])
> m2=ar(diff(sp5),method='mle')
> m2$order
[1] 2
> adfTest(sp5,lags=2,type=("ct"))

Title:
 Augmented Dickey-Fuller Test

Test Results:
 PARAMETER:
  Lag Order: 2
 STATISTIC:
  Dickey-Fuller: -2.0179
 P VALUE: 0.5708

> adfTest(sp5,lags=15,type=("ct"))
Title:
 Augmented Dickey-Fuller Test

Test Results:
 PARAMETER:
  Lag Order: 15
 STATISTIC:
  Dickey-Fuller: -1.9946
 P VALUE:  0.5807
```

Seasonal time Series

- Some financial time series exhibits certain cyclical or periodic behavior. Such a time series is called a seasonal time series.
- TS with periodic patten, e.g. quarterly earnings.
- Useful in weather-related derivative pricing.
- Useful in analysis of transactions data (high-frequency data), e.g. U-shaped pattern in intraday data.

E.g. Demand of electricity of a manufacturing sector of U.S. from 1972 to 1993. The data are logged usage on the 15th day of each month. See Figure 1.

E.g. Quarterly earnings of Johnson & Johnson See the time plot, Figures 2.13.

E.g. Quarterly earning per share of FedEx from the fourth quarter of 1991 to the fourth quarter of 2006. Figure 4.

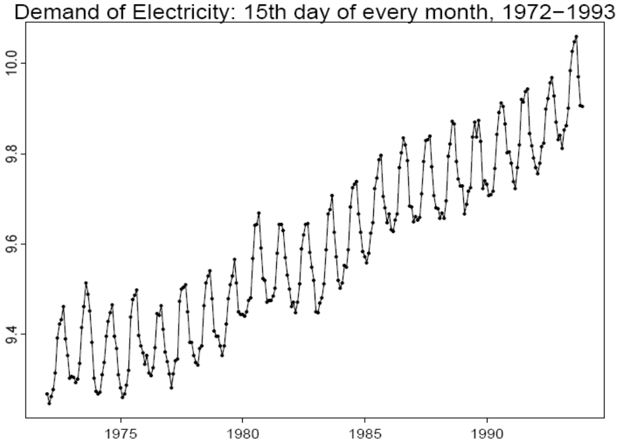


Figure 1: Time plot of electricity demand of an industrial sector: 15th day of each month from1972 to 1993.

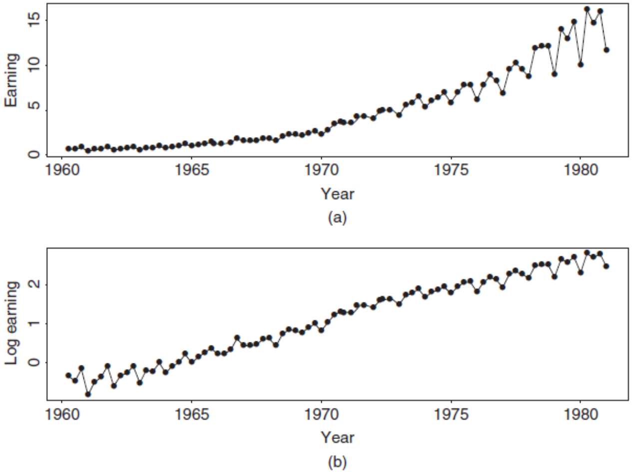


Figure 2.13 Time plots of quarterly earnings per share of Johnson & Johnson from 1960 to 1980: (a) observed earnings and (b) log earnings.

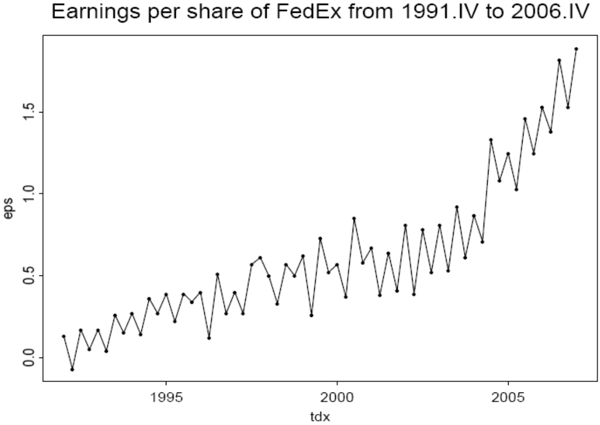


Figure 4: Time plot of quarterly earnings per share of FedEx: 1991.IV to 2006.IV

Seasonal adjustment

- Analysis of seasonal time series has a long history.
- In some applications, seasonality is of secondary importance and is removed from the data, resulting in a seasonally adjusted time series that is then used to make inference.
- The procedure to remove seasonality from a time series is referred to as *seasonal adjustment*.

Reasons for log transformation

- Figure 2.13(b) shows the time plot of log earnings per share of Johnson & Johnson.
- We took the log transformation for two reasons.
  - It is used to handle the exponential growth of the series. (The plot confirms that the growth is linear in the log scale.)
  - The transformation is used to stablize the variability of the series.

Johnson & Johnson example

- x_t = the log earnings of Johnson & Johnson.
- Figure 2.14(a) shows the sample ACF of x_t, $\Longrightarrow$ the quarterly log earnings per share has strong serial correlations.
- A conventional method to handle such strong serial correlations is to consider the first differenced series of x_t, i.e.

$$\Delta x_t = (1 - B)x_t.$$

- Figure 2.14(b) gives the sample ACF of x_t. $\Longrightarrow$ The ACF is strong when the lag is a multiple of periodicity 4.
- This is a well-documented behavior of sample ACF of a seasonal time series.

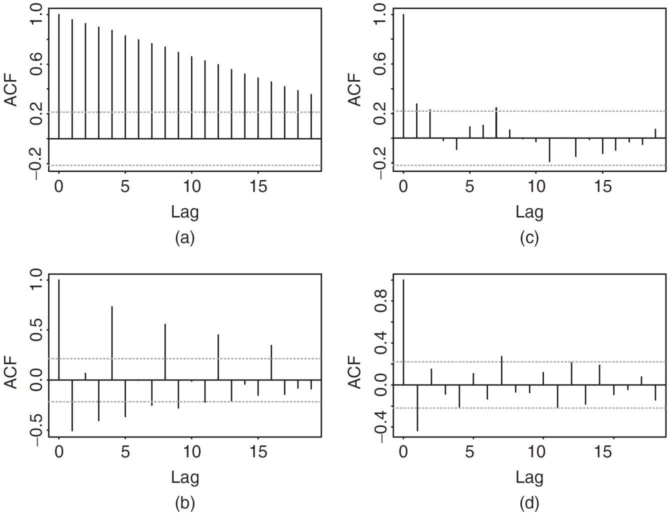


Figure 2.14 Sample ACF of log series of quarterly earnings per share of Johnson & Johnson from 1960 to 1980. (a) log earnings, (b) first differenced series, (c) seasonally differenced series, and (d) series with regular and seasonal differencing.

- Following the procedure of Box, Jenkins, and Reinsel (1994, Chapter 9), we take another difference of the data,

$$\Delta_4(\Delta x_t) = (1 - B^4)\Delta x_t = \Delta x_t - \Delta x_{t-4} = x_t - x_{t-1} - x_{t-4} + x_{t-5}.$$

- The operation $\Delta_4 = (1 - B^4)$ is called a *seasonal differencing*.
- In general, for a seasonal time series y_t with periodicity s, seasonal differencing means

$$\Delta_s y_t = y_t - y_{t-s} = (1 - B^s)y_t.$$

- The conventional difference $\Delta y_t = (1 - B)y_t$ is referred to as the *regular differencing*.
- Figure 2.14 (c)(d) also gives the plots of sample ACF of $\Delta_4 x_t$ and $\Delta_4 \Delta x_t$.

Multiplicative model

- Consider the following airline model:

$$(1 - B^s)(1 - B)x_t = (1 - \theta B)(1 - \Theta B^s)a_t,$$

where s is the periodicity of the series, a_t is a white noise series, $|\theta| < 1$, and $|\Theta| < 1$.

- The AR part of the model simply consists of the regular and seasonal differences, whereas the MA part involves two parameters.
- Define the differenced series w_t as

$$w_t = (1 - \theta B)(1 - \Theta B^s)a_t = a_t - \theta a_{t-1} - \Theta a_{t-s} + \theta\Theta a_{t-s-1},$$

where $w_t = (1 - B^s)(1 - B)x_t$ and $s > 1$.

$E(w_t) = 0$ and w_t 的 ACVF:

$$\psi(B) = (1-\theta B)(1-\Theta B^s), \quad \psi(B^{-1}) = (1-\theta B^{-1})(1-\Theta B^{-s})$$

$$\begin{aligned}
\psi(B)\psi(B^{-1}) &= (1-\theta B)(1-\Theta B^s)(1-\theta B^{-1})(1-\Theta B^{-s}) \\
&= (1-\theta B-\Theta B^s+\theta\Theta B^{s+1})(1-\theta B^{-1}-\Theta B^{-s}+\theta\Theta B^{-s-1}) \\
&= \theta\Theta B^{-s-1} - \Theta(1+\theta^2)B^{-s} + \theta\Theta B^{-s+1} - (\theta+\Theta^2\theta)B^{-1} \\
&\quad + 1+\theta^2+\Theta^2+\theta^2\Theta^2 + \\
&\quad - (\theta+\Theta^2\theta)B^1 + \theta\Theta B^{s-1} - \Theta(1+\theta^2)B^s + \theta\Theta B^{s+1}
\end{aligned}$$

Therefore,

$$\begin{aligned}
Var(w_t) &= (1+\theta^2)(1+\Theta^2)\sigma_a^2, \\
Cov(w_t, w_{t-1}) &= -\theta(1+\Theta^2)\sigma_a^2, \\
Cov(w_t, w_{t-s+1}) &= \theta\Theta\sigma_a^2, \\
Cov(w_t, w_{t-s}) &= -\Theta(1+\theta^2)\sigma_a^2, \\
Cov(w_t, w_{t-s-1}) &= \theta\Theta\sigma_a^2, \\
Cov(w_t, w_{t-l}) &= 0, \quad \text{for } l \neq 0, 1, s-1, s, s+1.
\end{aligned}$$

- Consequently, the ACF of the w_t series is given by

$$\rho_1 = \frac{-\theta}{1+\theta^2}, \quad \rho_s = \frac{-\Theta}{1+\Theta^2},$$
$$\rho_{s-1} = \rho_{s+1} = \rho_1\rho_s = \frac{\theta\Theta}{(1+\theta^2)(1+\Theta^2)},$$

and $\rho_l = 0$ for $l > 0$ and $l \neq 1, s-1, s, s+1$.

- For example, if w_t is a quarterly time series, then $s = 4$ and for $l > 0$, the ACF ρ_l is nonzero at lags 1, 3, 4, and 5 only.

- It is interesting to compare the prior ACF with those of the following models:

  MA(1) model: $y_t = (1 - \theta B)a_t$ and MA(s) model: $z_t = (1 - \Theta B^s)a_t$.

- The ACF of y_t and z_t series are

$$\rho_1(y) = \frac{-\theta}{1+\theta^2} \quad \text{and} \quad \rho_l(y) = 0,\ l > 1,$$
$$\rho_s(z) = \frac{-\Theta}{1+\Theta^2} \quad \text{and} \quad \rho_l(z) = 0,\ l > 0, \neq s.$$

  we see that

$$\rho_1 = \rho_1(y),\ \rho_s = \rho_s(z), \text{ and } \rho_{s-1} = \rho_{s+1} = \rho_1(y)\rho_s(z).$$

- Therefore, the ACF of w_t at lags $(s-1)$ and $(s+1)$ can be regarded as the interaction between lag-1 and lag-s serial dependence, and the model of w_t is called a multiplicative seasonal MA model.

- The model

$$w_t = (1 - \theta B - \Theta B^s)a_t,$$

where $|\theta| < 1$ and $|\Theta| < 1$, a nonmultiplicative seasonal MA model. It is easy to see that $\rho_{s+1} = 0$.

- A multiplicative model is more parsimonious than the corresponding nonmultiplicative model because both models use the same number of parameters, but the multiplicative model has more nonzero ACFs.

Multiplicative model for Johnson & Johnson

- We apply the airline model to the log series of quarterly earnings per share of Johnson & Johnson from 1960 to 1980.
- Based on the exact-likelihood method, the fitted model is

  $$(1-B)(1-B^4)x_t = (1-0.678B)(1-0.314B^4)a_t, \quad \hat{\sigma}_a = 0.089,$$

  where standard errors of the two MA parameters are 0.080 and 0.101, respectively.
- The Ljung–Box statistics of the residuals show $Q(12) = 10.0$ with a p value of 0.44. The model appears to be adequate.

- To illustrate the forecasting performance of the prior seasonal model, we reestimate the model using the first 76 observations and reserve the last 8 data points for forecasting evaluation.
- We compute 1-step- to 8-step-ahead forecasts and their standard errors of the fitted model at the forecast origin $h = 76$.
- Figure 2.15 shows the forecast performance of the model, where the observed data are in solid line, point forecasts are shown by dots, and the dashed lines show 95% interval forecasts.
- The forecasts show a strong seasonal pattern and are close to the observed data.

Figure 2.15 Out-of-sample point and interval forecasts for quarterly earnings of Johnson & Johnson. Forecast origin is fourth quarter of 1978. In plot, solid line shows actual observations, dots represent point forecasts, and dashed lines show 95% interval forecasts.

CRSP Decile 1 Index Example

- To demonstrate deterministic seasonal behavior, consider the monthly simple returns of the CRSP Decile 1 Index from January 1970 to December 2008 for 468 observations.
- The series is shown in Figure 2.16(a), and the time plot does not show any clear pattern of seasonality.
- The sample ACF of the return series shown in Figure 2.16(b) contains significant lags at 12, 24, and 36 as well as lag 1.
- If seasonal ARMA models are entertained, a model in the form

$$(1-\phi_1 B)(1-\phi_{12}B^{12})R_t = (1-\theta_{12}B^{12})a_t$$

is identified, where R_t denotes the monthly simple return.

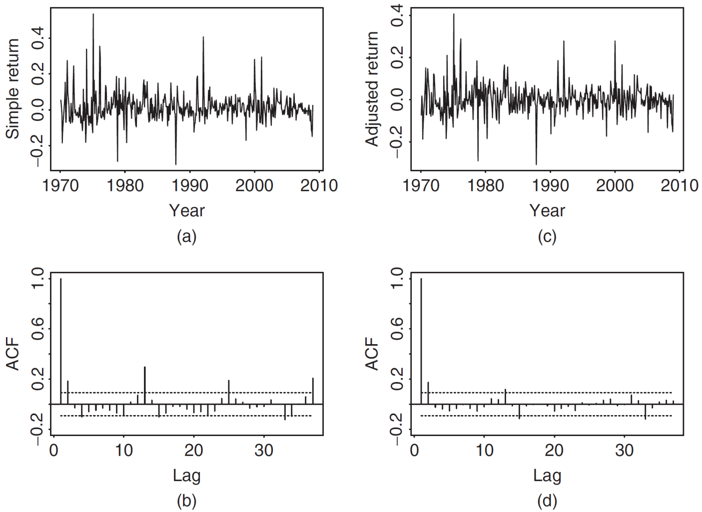


Figure 2.16 Monthly simple returns of CRSP Decile 1 index from January 1970 to December 2008: (a) time plot of the simple returns, (b) sample ACF of simple returns, (c) time plot of simple returns after adjusting for January effect, and (d) sample ACF of adjusted simple returns.

- Using the conditional-likelihood method, the fitted model is

  $$(1-0.18B)(1-0.87B^{12})R_t = (1-0.74B^{12})a_t, \quad \tilde{\sigma}_a = 0.069.$$

  The estimates of the seasonal AR and MA coefficients are of similar magnitude.

- If the exact-likelihood method is used, we have

  $$(1-0.188B)(1-0.951B^{12})R_t = (1-0.997B^{12})a_t, \quad \tilde{\sigma}_a = 0.063.$$

  The cancellation between seasonal AR and MA factors is clearly seen.

- This highlights the usefulness of using the exact-likelihood method and, the estimation result suggests that the seasonal behavior might be deterministic.

R Demonstration

The following output has been edited and % denotes explanation:

```
> da=read.table("m-deciles08.txt",header=T)
> d1=da[,2]
> acf(d1)
> pacf(d1)
> m1=arima(d1,order=c(1,0,0),seasonal = list(order=c(1,0,1),period=12))
> m1

Call:
arima(x = d1, order = c(1, 0, 0), seasonal = list(order = c(1, 0, 1), period = 12))

Coefficients:
         ar1    sar1     sma1  intercept
      0.1769  0.9882  -0.9144     0.0118
s.e.  0.0456  0.0093   0.0335     0.0129

sigma^2 estimated as 0.004717:  log likelihood = 584.07,  aic = -1158.14
> tsdiag(m1,gof=36)
```

tsdiag(m1,gof=36): This is a generic function. It will generally plot the residuals, often standardized, the autocorrelation function of the residuals, and the p-values of the Ljung–Box version of Portmanteau test for all lags up to gof.lag.

```
> m1=arima(d1,order=c(1,0,0),seasonal = list(order=c(1,0,1),period=12),include.mean = F)
> m1

Call:
arima(x = d1, order = c(1, 0, 0), seasonal = list(order = c(1, 0, 1), period = 12),
    include.mean = F)

Coefficients:
         ar1    sar1     sma1
      0.1787  0.9886  -0.9127
s.e.  0.0456  0.0089   0.0335

sigma^2 estimated as 0.00472:  log likelihood = 583.68,  aic = -1159.36
> Box.test(m1$residuals,lag=12)

        Box-Pierce test

data:  m1$residuals
X-squared = 7.2298, df = 12, p-value = 0.8421
```

```
> require(astsa)
> m1=sarima(d1,p=1,d=0,q=0,P=1,D=0,Q=1,S=12);
> m1$ttable
      Estimate     SE  t.value p.value
ar1     0.1769 0.0456   3.8819  0.0001
sar1    0.9882 0.0093 106.6654  0.0000
sma1   -0.9144 0.0335 -27.2789  0.0000
xmean   0.0118 0.0129   0.9140  0.3612
```

- To further confirm this assertion, we define the dummy variable for January, that is,

$$\mathrm{Jan}_t = \begin{cases} 1 & \text{if } t \text{ is January,} \\ 0 & \text{otherwise,} \end{cases}$$

and employ the simple linear regression

$$R_t = \beta_0 + \beta_1 \mathrm{Jan}_t + e_t.$$

- The fitted model is

$$R_t = 0.0029 + 0.1253\mathrm{Jan}_t + e_t,$$

where the standard errors of the estimates are 0.0033 and 0.0115, respectively.

- The right panel of Figure 2.16 shows the time plot and sample ACF of the residual series of the prior simple linear regression.
- From the sample ACF, serial correlations at lags 12, 24, and 36 largely disappear, suggesting that the seasonal pattern of the Decile 1 returns has been successfully removed by the January dummy variable.
- The seasonal behavior in the monthly simple return of Decile 1 is mainly due to the *January effect*.

```
> Date_tmp=as.Date(as.character(da$date),"%Y%m%d")
> month_tmp=months(Date_tmp)
> jan=as.numeric(month_tmp==" 一月")
> m2=lm(d1~jan)
> summary(m2)

Call:
lm(formula = d1 ~ jan)

Residuals:
     Min       1Q   Median       3Q      Max
-0.30861 -0.03475 -0.00176  0.03254  0.40671

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.002864   0.003333   0.859    0.391
jan         0.125251   0.011546  10.848   <2e-16 ***
---
Signif. codes:
0  '***'  0.001  '**'  0.01  '*'  0.05  '.'  0.1  ' '  1

Residual standard error: 0.06904 on 466 degrees of freedom
Multiple R-squared:  0.2016,       Adjusted R-squared:  0.1999
F-statistic: 117.7 on 1 and 466 DF,  p-value: < 2.2e-16

> Box.test(m2$residuals,lag=12)

        Box-Pierce test

data:  m2$residuals
X-squared = 27.195, df = 12, p-value = 0.007243
```

```
> m3=arima(d1,order=c(1,0,0),seasonal = list(order=c(1,0,1),period=12),xreg = jan)
> m3

Call:
arima(x = d1, order = c(1, 0, 0), seasonal = list(order = c(1, 0, 1), period = 12),
    xreg = jan)

Coefficients:
         ar1     sar1    sma1  intercept     jan
      0.1802  -0.0831  0.2127     0.0032  0.1184
s.e.  0.0458   0.3493  0.3420     0.0044  0.0123

sigma^2 estimated as 0.004522:  log likelihood = 599.14,  aic = -1186.29
> m4=arima(d1,order=c(1,0,0),seasonal = list(order=c(0,0,1),period=12),xreg = jan,include.mean = F)
> m4

Call:
arima(x = d1, order = c(1, 0, 0), seasonal = list(order = c(0, 0, 1), period = 12),
    xreg = jan, include.mean = F)

Coefficients:
         ar1    sma1     jan
      0.1805  0.1328  0.1204
s.e.  0.0459  0.0473  0.0121

sigma^2 estimated as 0.004527:  log likelihood = 598.86,  aic = -1189.71
> Box.test(m4$residuals,lag=12)

        Box-Pierce test

data:  m4$residuals
X-squared = 4.8418, df = 12, p-value = 0.9631
```

Regression Models with Time Series Errors

- Has many applications!
  e.g. 考慮不同的到期日利率的迴歸模型；

$$r_{1,t} = \alpha + \beta r_{2,t} + e_t,$$

  $r_{1,t}$ 與 $r_{2,t}$ 為兩個 times series; e_t 為誤差項。
- Impact of serial correlations in regression is often overlooked.
- It may introduce biases in estimates of α and β and in standard errors, resulting in unreliable t-ratios.
  (亦可能造成 α 與 β 的 L.S.E 成為不一致的估計值。)

Example U.S. weekly interest rate data: 1-year and 3-year constant maturity rates. Data are shown in Figure 2.17.

$$\begin{cases} r_{1t} = 1 \text{ 年期的債券利率} \\ r_{3t} = 3 \text{ 年期的債券利率} \end{cases}$$

各有 2467 筆資料 (1962/1 ~ 2009/4)。

- 兩者差分前後均高度相關 (see p.92)。 Figure 2.18
- 建模：$\underset{}{r_{3t}} = \underset{(0.024)}{0.832} + \underset{(0.004)}{0.930}\ r_{1t} + e_t,\ \hat{\sigma}_e = 0.523$，
  $R^2 = 96.5\%$ 且係數均顯著。
- p.93 殘差的 ACF 圖顯示殘差仍然具有高度相關。 Figure 2.19
- 由於 r_{1t} 與 r_{3t} 均為 $I(1)$，故此表示 r_{1t} 與 r_{3t} 並非共整合 (cointegration)。(即表示兩者不是維持長期均衡的關係)
- 從圖 2.17 亦可發現在某些時段出現 invertible yield curve(逆向收益曲線)，即利率與到期日成反比的異常現象。

改建 $c_{1t} = r_{1t} - r_{1,t-1}$ 及 $c_{3t} = r_{3t} - r_{3,t-1}$ 的迴歸模型：

$$c_{3t} = 0.792c_{1t} + e_t,$$

$\hat{\sigma}_e = 0.0690$, $R^2 = 82.5\%$.

Figure 2.20 shows time plots of the two change series. Figure 2.20

殘差仍然具有相關性 (Figure 2.21)，但可以簡單的 MA(1) 模型來配適：

$$e_t = a_t - \theta_1 a_{t-1}$$

將模型重新以聯合估計的方式，估計各係數 (p.84)：

$$c_{3t} = 0.792c_{1t} + e_t, \quad e_t = a_t + 0.1823a_{t-1}, \hat{\sigma}_a = 0.0678$$

$R^2 = 83.1\%$，此時的殘差 a_t 不再具有序列相關。

1. r_{1t} 與 r_{3t} 的 R^2 值雖高，並未必表示模型配的好 (稱為偽迴歸 spurious regression)。
2. 建立迴歸模型後，必須檢查其殘差是否具序列相關。
3. 加入與不加入 MA 模型的 R^2 沒有太大的差異，主因是此例中的 MA(1) 係數值並不大。

Remark **Parameterization in R**. With additional explanatory variable X in ARIMA model, **R** use the model

$$W_t = \phi_1 W_{t-1} + \cdots + \phi_p W_{t-p} + a_t + \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q},$$

where $W_t = Y_t - \beta_0 - \beta_1 X_t$. This is the proper way to handle regression model with time series errors, because W_{t-1} is not subject to the effect of X_{t-1}.

- It is different from the model

$$Y_t = \beta_0^* + \beta_1^* X_t + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + a_t + \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q},$$

for which the Y_{t-1} contains the effect of X_{t-1}.

R Demonstration

The following output has been edited.

```
> r1=read.table("w-gs1yr.txt",header=T)[,4]
> r3=read.table("w-gs3yr.txt",header=T)[,4]
> m1=lm(r3 r1)
> summary(m1)
Call:
lm(formula = r3 ~ r1)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept ) 0.83214    0.02417   34.43  <2e-16 ***
r1           0.92955    0.00357  260.40  <2e-16 ***
---
Residual standard error: 0.5228 on 2465 degrees of freedom
Multiple R-squared: 0.9649,    Adjusted R-squared: 0.9649
```

```
> plot(m1$residuals,type='l')
> acf(m1$residuals,lag=36)
> c1=diff(r1)
> c3=diff(r3)
> m2=lm(c3 -1+c1)
> summary(m2)
Call:
lm(formula = c3 ~ -1 + c1)
Coefficients:
   Estimate Std. Error t value  Pr(>|t|)
c1  0.791935   0.007337   107.9   <2e-16 ***
---
Residual standard error: 0.06896 on 2465 degrees of freedom
Multiple R-squared: 0.8253,   Adjusted R-squared: 0.8253
```

```
> acf(m2$residuals,lag=36)

> m3=arima(c3,order=c(0,0,1),xreg=c1,include.mean=F)
> m3
Call:
arima(x = c3, order = c(0, 0, 1), xreg = c1, include.mean = F)
Coefficients:
         ma1      c1
      0.1823  0.7936
s.e.  0.0196  0.0075

sigma^2 estimated as 0.0046: log likelihood=3136.62,
   aic=-6267.23
>
> rsq=(sum(c3^2)-sum(m3$residuals^2))/sum(c3^2)
> rsq
[1] 0.8310077
```

Summary

We outline a general procedure for analyzing linear regression models with time series errors:

(1) Fit the linear regression model and check serial correlations of the residuals.

(2) If the residual series is unit-root nonstationary, take the first difference of both the dependent and explanatory variables. Go to step 1. If the residual series appears to be stationary, identify an ARMA model for the residuals and modify the linear regression model accordingly.

(3) Perform a joint estimation via the maximum-likelihood method and check the fitted model for further improvement.

Remark

- For a residual series e_t with T observations, the Durbin–Watson statistic is

$$\text{DW} = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}.$$

  Straightforward calculation shows that DW$\approx 2(1-\hat{\rho}_1)$, which $\hat{\rho}_1$ is the lag-1 ACF of $\{e_t\}$.
- To check the serial correlations of residuals, we recommend that the Ljung–Box statistics be used instead of the Durbin–Watson (DW) statistic because the latter only considers the lag-1 serial correlation.

Consistent covariance matrix estimation

- Consider again the regression model in Eq. (2.43).

$$y_t = \alpha + \beta x_t + e_t, \tag{2.43}$$

  where y_t and x_t are two time series and e_t denotes the error term.
  - There may exist situations in which the error term e_t has serial correlations and/or conditional heteroscedasticity, but the main objective of the analysis is to make inference concerning the regression coefficients α and β.
  - See Chapter 3 for discussion of conditional heteroscedasticity.

- In situations under which the OLS estimates of the coefficients remain consistent, methods are available to provide consistent estimate of the covariance matrix of the coefficient estimates.
- Two such methods are widely used.
  - The first method is called the heteroscedasticity consistent (HC) estimator; see Eicker (1967) and White (1980).
  - The second method is called the heteroscedasticity and autocorrelation consistent (HAC) estimator; see Newey and West (1987).

- For ease in discussion, we shall rewrite the regression model as

$$y_t = \mathbf{x}_t' \boldsymbol{\beta} + e_t, \quad t = 1, \ldots, T, \tag{2.48}$$

  - y_t is the dependent variable.
  - $\mathbf{x}_t = (x_{1t}, \ldots, x_{kt})'$ is a k-dimensional vector of explanatory variables including constant.
  - $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)'$ is the parameter vector.
  - $\mathbf{c}'$ denotes the transpose of the vector $\mathbf{c}$.

- The LS estimate of $\boldsymbol{\beta}$ and the associate covariance matrix are

$$\hat{\boldsymbol{\beta}} = \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'\right]^{-1} \sum_{t=1}^{T} \mathbf{x}_t y_t, \quad \mathsf{Cov}(\hat{\boldsymbol{\beta}}) = \sigma_\epsilon^2 \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'\right]^{-1},$$

  where σ_ϵ^2 is the variance of e_t and is estimated by the variance of the residuals of the regression.
- In the presence of serial correlations or conditional heteroscedasticity, the prior covariance matrix estimator is inconsistent, often resulting in inflating the t ratios of $\hat{\boldsymbol{\beta}}$.

- The estimator of White (1980) is

$$\text{Cov}(\hat{\boldsymbol{\beta}})_{\text{HC}} = \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'\right]^{-1} \left[\sum_{t=1}^{T} \hat{e}_t^2 \mathbf{x}_t \mathbf{x}_t'\right] \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'\right]^{-1}, \qquad (2.49)$$

where $\hat{e}_t = y_t - \mathbf{x}_t'\hat{\boldsymbol{\beta}}$ is the residual at time t.

- The estimator of Newey and West (1987) is

$$\text{Cov}(\hat{\boldsymbol{\beta}})_{\text{HAC}} = \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'\right]^{-1} \hat{\mathbf{C}}_{\text{HAC}} \left[\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'\right]^{-1}, \qquad (2.50)$$

where

$$\hat{\mathbf{C}}_{\text{HAC}} = \sum_{t=1}^{T} \hat{e}_t^2 \mathbf{x}_t \mathbf{x}_t' + \sum_{j=1}^{l} w_j \sum_{t=j+1}^{T} (x_t \hat{e}_t \hat{e}_{t-j} \mathbf{x}_{t-j}' + \mathbf{x}_{t-j} \hat{e}_{t-j} \hat{e}_t \mathbf{x}_t'),$$

where l is a truncation parameter and w_j is a weight function such as the Bartlett weight function defined by

$$w_j = 1 - \frac{j}{l+1}.$$

- Other weight functions can also be used.
- Newey and West (1987) suggest choosing l to be the integer part of $4(T/100)^{2/9}$. This estimator essentially uses a nonparametric method to estimate the covariance matrix of $\{\sum_{t=1}^{T} \hat{e}_t \mathbf{x}_t\}$.
- For illustration, we employ the first differenced interest rate series in Eq. (2.45).
- The t ratio of the coefficient of c_{1t} is 107.91 if both serial correlation and heteroscedasticity in the residuals are ignored, it becomes 48.44 when the HC estimator is used, and it reduces to 45.33(Use Andrews weights) when the HAC estimator is used.

R Demonstration

The following output has been edited and % denotes explanation:

```
> require(sandwich)
> require(lmtest)
> r1=read.table("w-gs1yr.txt",header=T)
> r3=read.table("w-gs3yr.txt",header=T)
> c1=diff(r1$rate)
> c3=diff(r3$rate)
> reg.fit=lm(c3~c1);
> summary(reg.fit)

Call:
lm(formula = c3 ~ c1)

Residuals:
     Min       1Q   Median       3Q      Max
-0.42459 -0.03578 -0.00117  0.03467  0.48921

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.0001051  0.0013890  -0.076     0.94
c1           0.7919323  0.0073391 107.906   <2e-16 ***
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.06897 on 2464 degrees of freedom
Multiple R-squared:  0.8253,      Adjusted R-squared:  0.8253
F-statistic: 1.164e+04 on 1 and 2464 DF,  p-value: < 2.2e-16
```

```
> coeftest(reg.fit,vcov=vcovHC(reg.fit,type="HC"))

t test of coefficients:

                Estimate  Std. Error t value Pr(>|t|)
(Intercept) -0.00010515  0.00138797 -0.0758   0.9396
c1           0.79193231  0.01634193 48.4601   <2e-16 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> coeftest(reg.fit,vcov=vcovHAC(reg.fit))

t test of coefficients:

                Estimate  Std. Error t value Pr(>|t|)
(Intercept) -0.00010515  0.00158285 -0.0664    0.947
c1           0.79193231  0.01746723 45.3382   <2e-16 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- The R demonstration below also uses a regression that includes lagged values $c_{1,t-1}$ and $c_{3,t-1}$ as regressors to take care of serial correlations in the residuals.

$$c_{3,t} = \beta_0 + \beta_1 c_{1,t} + \beta_2 c_{3,t-1} + \beta_3 c_{1,t-1} + e_t$$

- One can also apply the HC or HAC estimator to the fitted model to refine the t ratios of the coefficient estimates.

R Demonstration

The following output has been edited and % denotes explanation:

```
> #Below, fit a regression model with time series errors
> reg.ts=lm(c3[-1]~c1[-1]+c3[-length(c3)]+c1[-length(c1)])
> summary(reg.ts)

Call:
lm(formula = c3[-1] ~ c1[-1] + c3[-length(c3)] + c1[-length(c1)])

Residuals:
     Min       1Q   Median       3Q      Max
-0.44812 -0.03551 -0.00075  0.03408  0.45821

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)     -8.686e-05  1.367e-03  -0.064    0.949
c1[-1]           7.971e-01  7.692e-03 103.632   <2e-16 ***
c3[-length(c3)]  1.766e-01  1.983e-02   8.906   <2e-16 ***
c1[-length(c1)] -1.580e-01  1.745e-02  -9.058   <2e-16 ***
---
Signif. codes:  0  ‘***’  0.001  ‘**’  0.01  ‘*’  0.05  ‘.’  0.1  ‘ ’  1

Residual standard error: 0.06785 on 2461 degrees of freedom
Multiple R-squared:  0.8312,        Adjusted R-squared:  0.831
F-statistic:  4039 on 3 and 2461 DF,  p-value: < 2.2e-16
```

```
> coeftest(reg.ts,vcov=vcovHC(reg.ts,type="HC"))

t test of coefficients:

                  Estimate  Std. Error t value  Pr(>|t|)
(Intercept)     -8.6859e-05  1.3649e-03 -0.0636    0.9493
c1[-1]           7.9714e-01  1.6416e-02 48.5591 < 2.2e-16 ***
c3[-length(c3)]  1.7662e-01  2.9199e-02  6.0487 1.683e-09 ***
c1[-length(c1)] -1.5802e-01  3.0570e-02 -5.1693 2.539e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> coeftest(reg.ts,vcov=vcovHAC(reg.ts))

t test of coefficients:

                  Estimate  Std. Error t value  Pr(>|t|)
(Intercept)     -8.6859e-05  1.3758e-03 -0.0631    0.9497
c1[-1]           7.9714e-01  1.6840e-02 47.3372 < 2.2e-16 ***
c3[-length(c3)]  1.7662e-01  2.8105e-02  6.2842 3.884e-10 ***
c1[-length(c1)] -1.5802e-01  2.8745e-02 -5.4975 4.250e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

- Let $\hat{\beta}_j$ be the jth element of $\hat{\boldsymbol{\beta}}$.
- When $k > 1$, the HC variance of $\hat{\beta}_j$ in Eq. (2.49) can be obtained by using an auxiliary regression.
- Let $\mathbf{x}_{-j,t}$ be the $(k-1)$-dimensional vector obtained by removing the element x_{jt} from $\mathbf{x}_t$.
- Consider the auxiliary regression

$$x_{jt} = \mathbf{x}'_{-j,t}\boldsymbol{\gamma} + v_t, \quad t = 1, \ldots, T. \tag{2.51}$$

- Let $\hat{v}_t$ be the least-squares residual of this auxiliary regression. It can be shown that

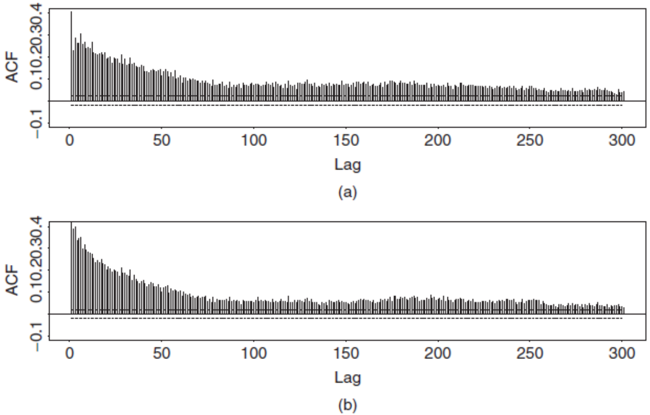
$$\mathrm{Var}(\hat{\beta}_j)_{\mathrm{HC}} = \frac{\sum_{t=1}^{T} \hat{e}_t^2 \hat{v}_t^2}{\left(\sum_{t=1}^{T} \hat{v}_t^2\right)^2},$$

  where $\hat{e}_t$ is the residual of original regression in Eq. (2.48).
- The auxiliary regression is simply a step taken to achieve orthogonality between $\hat{v}_t$ and the rest of the regressors so that the formula in Eq. (2.49) can be simplified.

Long-memory models

- Meaning? ACF decays to zero very slowly!
- Example: ACF of squared or absolute log returns ACFs are small, but decay very slowly.(see p.103 Fig. 2.22 ▸ Figure 2.22)
- How to model long memory? Use "fractional" difference: namely, $(1-B)^d x_t$, where $-0.5 < d < 0.5$.
- Importance? In theory, Yes. In practice, yet to be determined.

Figure: Sample acf of absolute series of daily simple returns for CRSP value- and equal-weighted indexes: (a) value-weighted index return and (b) equal-weighted index return. 1970.01.02 $\sim$ 2008.12.31.



- x_t is an I(d) process, if $(1-B)^d x_t = a_t$.
- For $d < 0.5$, $\Rightarrow x_t$ 為 weakly stationary 且具有 MA(∞) 表式

$$x_t = a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j},$$

where

$$\psi_k = \frac{d(1+d)\cdots(k-1+d)}{k!} = \frac{(k+d-1)!}{k!(d-1)!} = \binom{k+d-1}{k} = (-1)^k \binom{-d}{k}$$

- For $d > -0.5$, $\Rightarrow x_t$ 為 invertible 且具有 AR(∞) 表式

$$x_t = \sum_{i=1}^{\infty} \pi_i x_{t-i} + a_t,$$

where

$$\pi_k = \frac{-d(1-d)\cdots(k-1-d)}{k!} = \frac{(k-d-1)!}{k!(-d-1)!} = \binom{k-d-1}{k} = (-1)^k \binom{d}{k}$$

- For $-0.5 < d < 0.5$, x_t 的 ACF 為

$$\rho_k = \frac{d(1+d)\cdots(k-1+d)}{(1-d)(2-d)\cdots(k-d)}, \quad k = 1, 2, \ldots$$

In particular, $\rho_1 = d/(1-d)$ and

$$\rho_k \approx \frac{(-d)!}{(d-1)!} k^{2d-1} \quad \text{as } k \to \infty.$$

- For $-0.5 < d < 0.5$, x_t 的 PACF 為

$$\phi_{kk} = \frac{d}{k-d}, \quad k = 1, 2, \ldots.$$

- For $-0.5 < d < 0.5$, x_t 的 spectral density function

$$\begin{aligned} f(\omega) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} r_k e^{-i\omega k} \\ &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} r_k \cos(\omega k), \quad \omega \in [0, 2\pi] \\ &\approx \omega^{-2d} \quad \text{as } \omega \to 0. \end{aligned}$$

- Of particular interest here is the behavior of ACF of x_t when $d < 0.5$.
- The property says that $\rho_k \sim k^{2d-1}$, which decays at a polynomial, instead of exponential, rate. For this reason, such an x_t process is called a long-memory time series.
- If the fractionally differenced series

$$(1-B)^d x_t$$

follows an ARMA(p, q) model, then x_t is called an ARFIMA(p, d, q) process, which is a generalized ARIMA model by allowing for noninteger d.

Summary of the chapter

- Sample ACF $\Rightarrow$ MA order.
- Sample PACF $\Rightarrow$ AR order.
- Some packages have "automatic" procedure to select a simple model for "conditional mean" of a FTS, e.g., R uses "ar" for AR models.
- Check a fitted model before forecasting, e.g. residual ACF and hetroscedasticity (chapter 3)

- Interpretation of a model, e.g. constant term
- For an AR(1) with coefficient ϕ_1, the speed of mean reverting as measured by half-life (find k such that $\phi_1^k = 0.5$) is

$$k = \frac{\ln(0.5)}{\ln(\phi_1)}.$$

- For an MA(q) model, forecasts revert to the mean in $q+1$ steps.
- Make proper use of regression models with time series errors, e.g. regression with AR(1) residuals Perform a joint estimation instead of using any two-step procedure, e.g. Cochrane-Orcutt (1949).

Figure 2.17: Time plot of US weekly interest rates from 1962/1/5 *to* 2009/4/10*. The solid line is the Treasury* 1*-year constant maturity rate and the dashed line the Treasury* 3 *year constant maturity rate.*

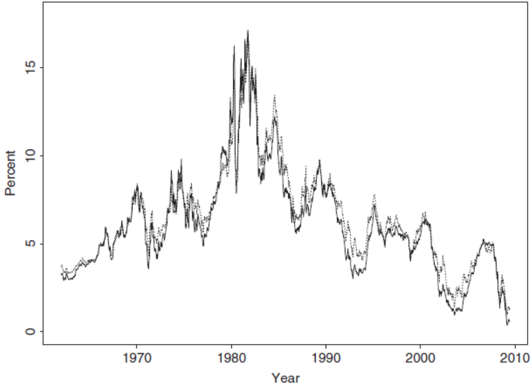
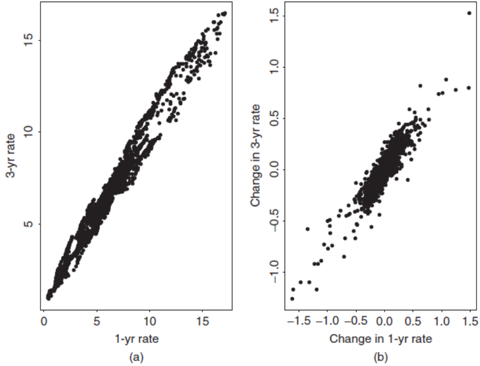


Figure 2.18: Scatterplots of US weekly interest rates from 1962/1/5 *to* 2009/4/10*. (a)* 3*-year rate versus* 1*-year rate and (b) changes in* 3*-year rate versus changes in* 1*-year rate.*



Back

Figure 2.19: Residual series of linear regression for two US weekly interest rates: (a) time plot and (b) sample acf.

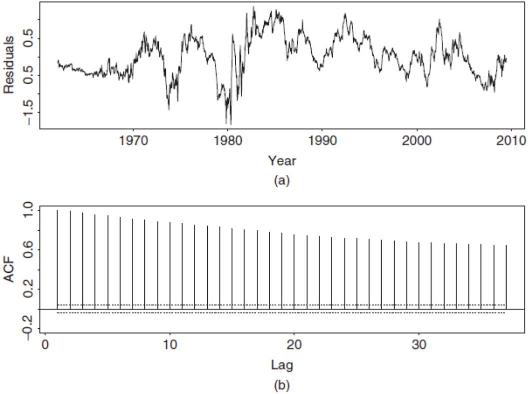
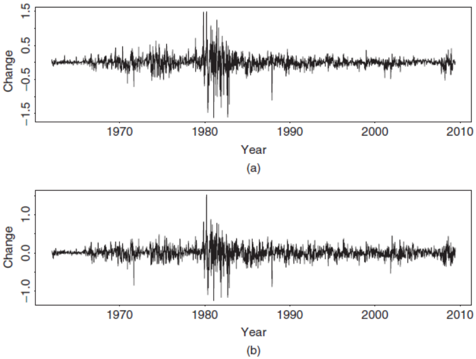
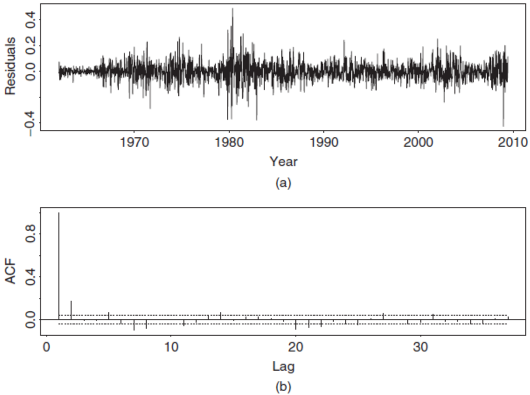


Figure 2.20: Time plot of the change series of US weekly interest rates from 1962/1/12 to 2009/4/10. (a) changes in the Treasury 1-year constant maturity rate and (b) changes in the Treasury 3-year constant maturity rate.



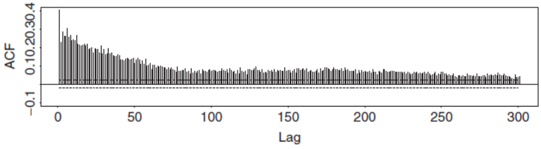
Back

Figure 2.21: Residual series of the linear regression for two change series of US weekly interest rates (a) time plot and (b) sample acf.

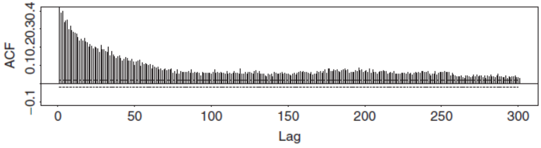


Back

Figure 2.22: Sample autocorrelation function of absolute series of daily simple returns for CRSP value- and equal-weighted indexes: (a) value-weighted index return and (b) equal-weighted index return. Sample period is from January 2, 1970, to December 31, 2008. Back



(a)



(b)