CHAPTER 5 MODELS FOR NONSTATIONARY TIME SERIES

CHAPTER 5 MODELS FOR NONSTATIONARY TIME SERIES

Backgournd

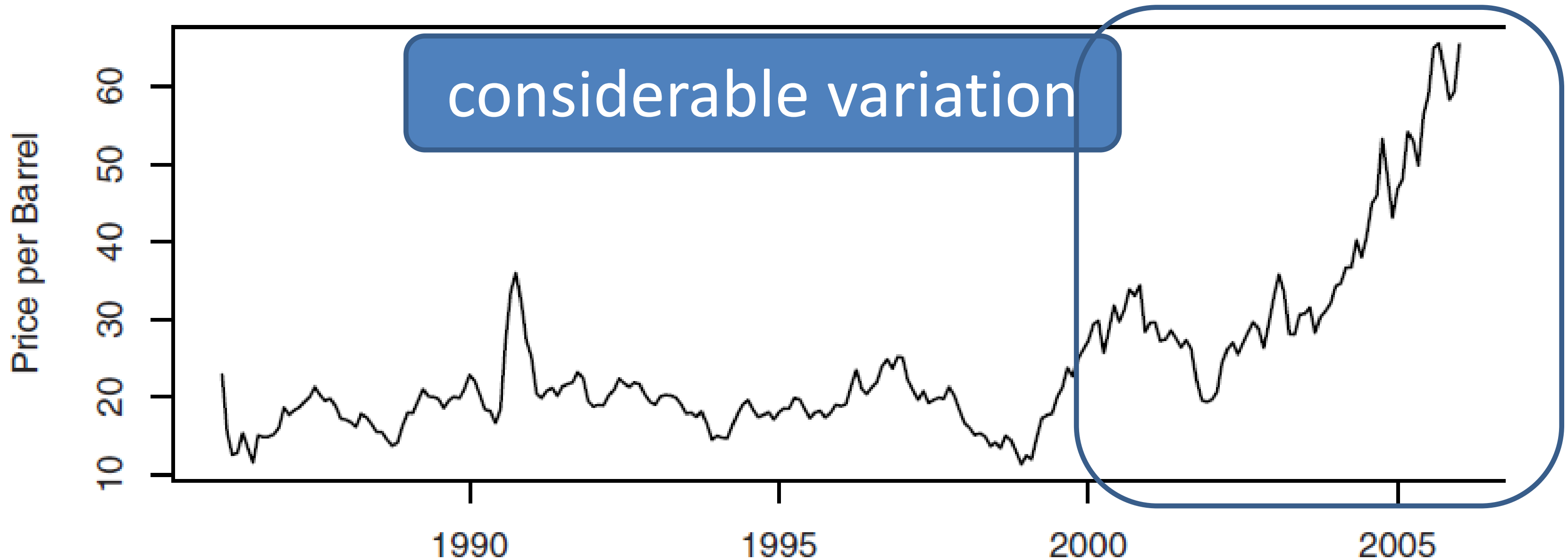
Any time series without a constant mean over time is nonstationary.

Deterministic trend ($Y_t = \mu_t + X_t$) were considered in Chapter 3.

Frequently in applications, we **cannot** legitimately assume a deterministic trend (e.g., random walk).

CHAPTER 5 MODELS FOR NONSTATIONARY TIME SERIES

Exhibit 5.1 Monthly Price of Oil: January 1986–January 2006



no deterministic trend model works well

```
> win.graph(width=4.875,height=3,pointsize=8)
> data(oil.price)
> plot(oil.price, ylab='Price per Barrel',type='l')
```

nonstationary models seem reasonable

5.1 Stationarity Through Differencing

Consider again the AR(1) model

$$Y_t = \phi Y_{t-1} + e_t$$

Stationarity $\Longleftrightarrow$ $|\phi| < 1$

What can we say about if $|\phi| \geq 1$?

5.1 Stationarity Through Differencing

Consider in particular the equation

$$Y_t = 3Y_{t-1} + e_t$$

$$Y_t = e_t + 3e_{t-1} + 3^2 e_{t-2} + \cdots + 3^{t-1} e_1 + 3^t Y_0$$

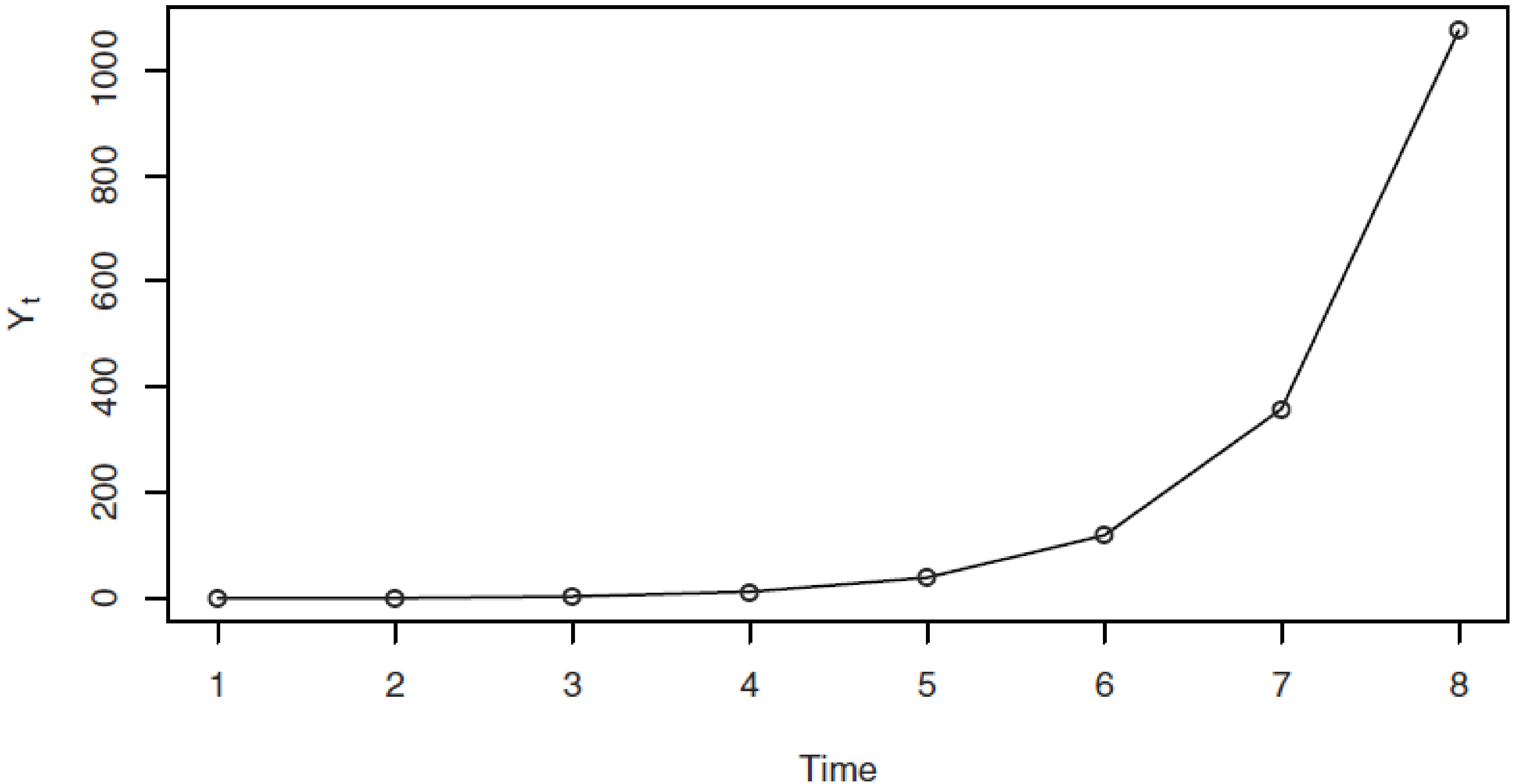
the influence of distant past values of Y_t and e_t does not die out·
the weights applied to Y_0 and e_1 grow exponentially large.

Exhibit 5.2 Simulation of the Explosive "AR(1) Model" $Y_t = 3Y_{t-1} + e_t$

t	1	2	3	4	5	6	7	8
e_t	0.63	−1.25	1.80	1.51	1.56	0.62	0.64	−0.98
Y_t	0.63	0.64	3.72	12.67	39.57	119.33	358.63	1074.91

5.1 Stationarity Through Differencing

Exhibit 5.3 An Explosive "AR(1)" Series




```
> data(explode.s)
> plot(explode.s,ylab=expression(Y[t]),type='o')
```

5.1 Stationarity Through Differencing

explosive behavior

$$Var(Y_t) = \frac{1}{8}(9^t - 1)\sigma_e^2$$

$$Cov(Y_t, Y_{t-k}) = \frac{3^k}{8}(9^{t-k} - 1)\sigma_e^2$$

$$Corr(Y_t, Y_{t-k}) = 3^k\left(\frac{9^{t-k} - 1}{9^t - 1}\right) \approx 1 \quad \text{for large } t \text{ and moderate } k$$

square root

5.1 Stationarity Through Differencing

A more reasonable type of nonstationarity

$$Y_t = Y_{t-1} + e_t$$

random walk process

$$\Rightarrow \nabla Y_t = e_t$$

stationary

where $\nabla Y_t = Y_t - Y_{t-1}$ is the **first difference** of Y_t.

5.1 Stationarity Through Differencing

Some **what different sets of assumptions** can lead to models whose **first difference** is a **stationary** process.

Suppose

$$Y_t = M_t + e_t$$

where M_t is a series that is changing only slowly over time.

If M_t is approximately constant over every two consecutive time points

$$\hat{M}_t = \frac{1}{2}(Y_t + Y_{t-1})$$

5.1 Stationarity Through Differencing

$$Y_t - \hat{M}_t = Y_t - \frac{1}{2}(Y_t + Y_{t-1}) = \frac{1}{2}(Y_t - Y_{t-1}) = \frac{1}{2}\nabla Y_t$$

Stationary: detrended

If M_t is stochastic and changes slowly over time governed by a random walk model.

Suppose

$$Y_t = M_t + e_t \qquad \text{with} \qquad M_t = M_{t-1} + \varepsilon_t$$

where $\{e_t\}$ and $\{\varepsilon_t\}$ are independent white noise series.

5.1 Stationarity Through Differencing

$$\nabla Y_t = \nabla M_t + \nabla e_t$$

$$= \varepsilon_t + e_t - e_{t-1}$$

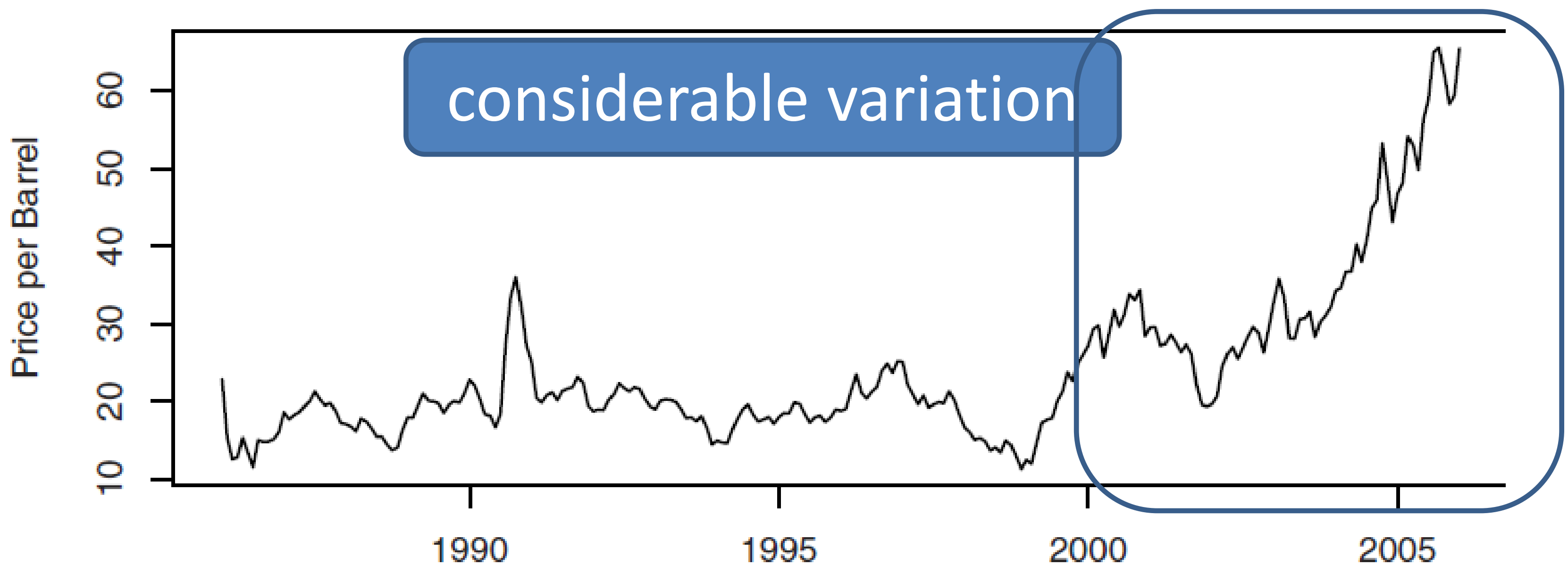
stationary

autocorrelation function of an MA(1) series with

$$\rho_1 = -\{1/[2 + (\sigma_\varepsilon^2/\sigma_e^2)]\}$$

5.1 Stationarity Through Differencing

Exhibit 5.1 Monthly Price of Oil: January 1986–January 2006



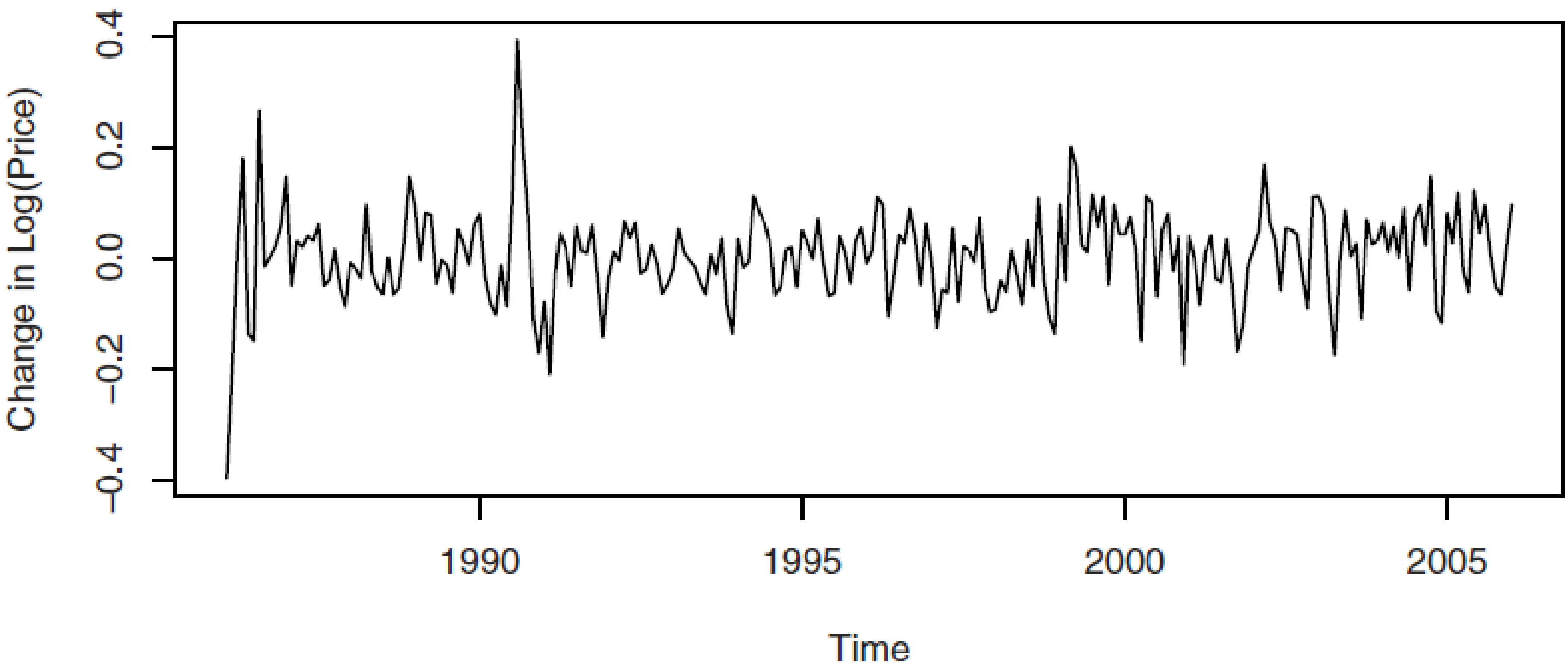
no deterministic trend model works well

```
> win.graph(width=4.875,height=3,pointsize=8)
> data(oil.price)
> plot(oil.price, ylab='Price per Barrel',type='l')
```

nonstationary models seem reasonable

5.1 Stationarity Through Differencing

Exhibit 5.4 The Difference Series of the Logs of the Oil Price Time




```
> plot(diff(log(oil.price)),ylab='Change in Log(Price)',type='l')
```

The differenced series looks much more stationary when compared with the original time series

5.1 Stationarity Through Differencing

assumptions that lead to stationary second-difference models

Suppose

$$Y_t = M_t + e_t$$

where M_t is linear in time over **three** consecutive time points.

$$\hat{M}_t = \frac{1}{3}(Y_{t+1} + Y_t + Y_{t-1})$$

5.1 Stationarity Through Differencing

the ***detrended*** series is

$$
\begin{aligned}
Y_t - \hat{M}_t &= Y_t - \left(\frac{Y_{t+1} + Y_t + Y_{t-1}}{3}\right) \\
&= \left(-\frac{1}{3}\right)(Y_{t+1} - 2Y_t + Y_{t-1}) \\
&= \left(-\frac{1}{3}\right)\nabla(\nabla Y_{t+1}) \\
&= \left(-\frac{1}{3}\right)\nabla^2(Y_{t+1})
\end{aligned}
$$

stationary

5.1 Stationarity Through Differencing

Alternatively, we might assume that

$$\boxed{Y_t = M_t + e_t, \quad M_t = M_{t-1} + W_t \qquad W_t = W_{t-1} + \varepsilon_t}$$

with $\{e_t\}$ and $\{\varepsilon_t\}$ independent white noise time series.

$$\Rightarrow \quad \nabla Y_t = \nabla M_t + \nabla e_t = W_t + \nabla e_t$$

$$\Rightarrow \quad \nabla^2 Y_t = \nabla W_t + \nabla^2 e_t$$

MA(2) process

$$= \varepsilon_t + (e_t - e_{t-1}) - (e_{t-1} - e_{t-2})$$

stationary

$$= \varepsilon_t + e_t - 2e_{t-1} + e_{t-2}$$

5.2 ARIMA Models

A time series $\{Y_t\}$ is said to follow an **integrated autoregressive moving average** model if the *d*th difference $W_t = \nabla^d Y_t$ is a stationary ARMA process. If $\{W_t\}$ follows an ARMA(*p*,*q*) model, we say that $\{Y_t\}$ is an ARIMA(*p*,*d*,*q*) process.

Consider then an ARIMA$(p,1,q)$ process.

With $W_t = Y_t - Y_{t-1}$, we have

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

5.2 ARIMA Models

or, in terms of the observed series,

$$Y_t - Y_{t-1} = \phi_1(Y_{t-1} - Y_{t-2}) + \phi_2(Y_{t-2} - Y_{t-3}) + \cdots + \phi_p(Y_{t-p} - Y_{t-p-1})$$
$$+ e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

which we may rewrite as

$$Y_t = (1 + \phi_1)Y_{t-1} + (\phi_2 - \phi_1)Y_{t-2} + (\phi_3 - \phi_2)Y_{t-3} + \cdots$$
$$+ (\phi_p - \phi_{p-1})Y_{t-p} - \phi_p Y_{t-p-1} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

difference equation form

?

ARMA($p + 1, q$) process

5.2 ARIMA Models

The characteristic polynomial satisfies

$$1-(1+\phi_1)x-(\phi_2-\phi_1)x^2-(\phi_3-\phi_2)x^3-\cdots-(\phi_p-\phi_{p-1})x^p+\phi_p x^{p+1}$$

$$= \underbrace{(1-\phi_1 x-\phi_2 x^2-\cdots-\phi_p x^p)}_{\textbf{\textit{stationary}}}\underbrace{(1-x)}_{\textbf{\textit{nonstationarity}}}$$

Explicit representations of the observed series in terms of either $\{W_t\}$ or the white noise series underlying $\{W_t\}$

For convenience, we take $Y_t = 0$ for $t < -m$.

5.2 ARIMA Models

for the ARIMA(p,1,q) process

$$Y_t = \sum_{j=-m}^{t} W_j$$

for the ARIMA(p,2,q) process

$$Y_t = \sum_{j=-m}^{t} \sum_{i=-m}^{j} W_i$$

$$= \sum_{j=0}^{t+m} (j+1) W_{t-j}$$

investigate covariance properties

Definitions: *IMA(d,q), ARI(p,d)*

5.2.1 The IMA(1,1) Model

In difference equation form, the model is

$$Y_t = Y_{t-1} + e_t - \theta e_{t-1}$$

$$W_t = e_t - \theta e_{t-1}$$

$$Y_t = e_t + (1-\theta)e_{t-1} + (1-\theta)e_{t-2} + \cdots + (1-\theta)e_{-m} - \theta e_{-m-1}$$

The weights on the white noise terms *do not die out*.

nonstationarity

5.2.1 The IMA(1,1) Model

$$Var(Y_t) = [1 + \theta^2 + (1 - \theta)^2(t + m)]\sigma_e^2$$

increases

$$Corr(Y_t, Y_{t-k}) = \frac{1 - \theta + \theta^2 + (1 - \theta)^2(t + m - k)}{[Var(Y_t)Var(Y_{t-k})]^{1/2}}$$

$$\approx \sqrt{\frac{t + m - k}{t + m}}$$

$$\approx 1 \qquad \text{for large } m \text{ and moderate } k$$

strongly positive

5.2.2 The IMA(2,2) Model

In difference equation form

$$\nabla^2 Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

or

$$Y_t = 2Y_{t-1} - Y_{t-2} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

$$Y_t = e_t + \sum_{j=1}^{t+m} \psi_j e_{t-j} - [(t+m+1)\theta_1 + (t+m)\theta_2]e_{-m-1} - (t+m+1)\theta_2 e_{-m-2}$$

where $\psi_j = 1 + \theta_2 + (1 - \theta_1 - \theta_2)j$

ψ-weights do not die out

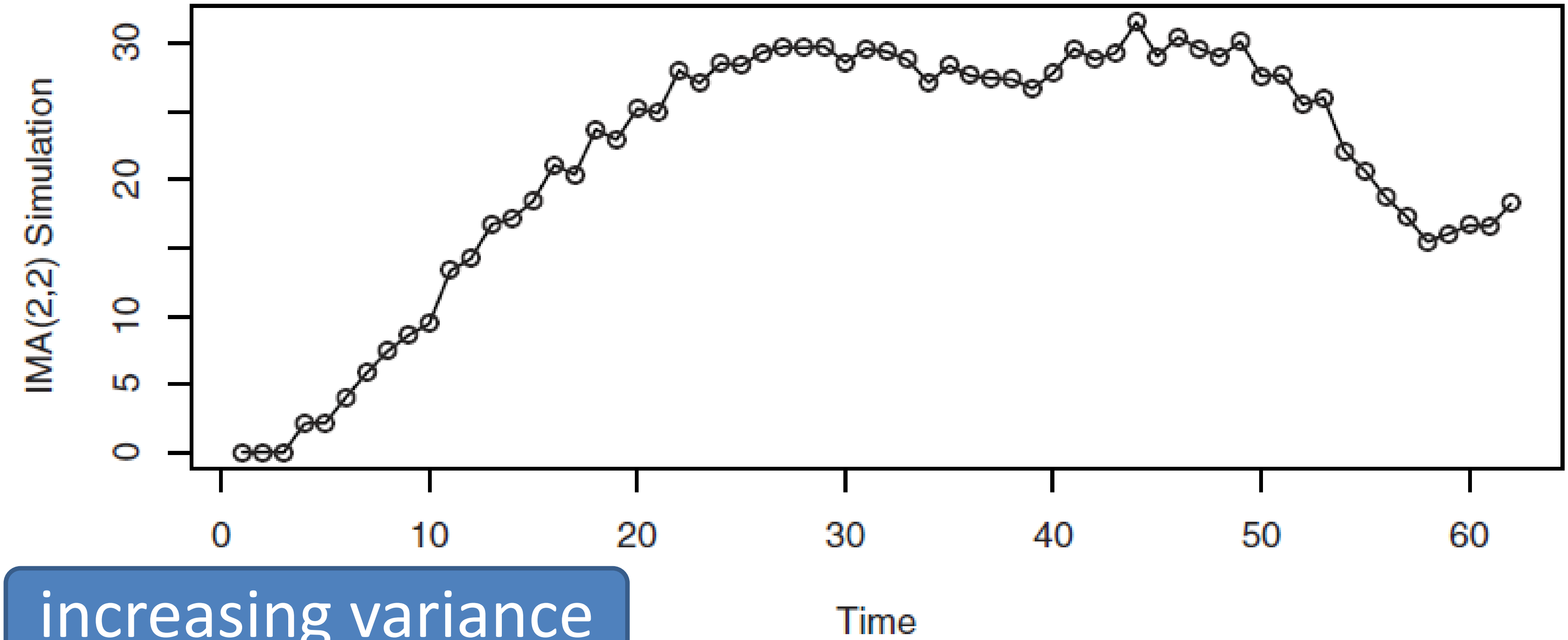
5.2.2 The IMA(2,2) Model

the variance of Y_t increases rapidly with t and again $Corr(Y_t, Y_{t-k})$ is nearly 1 for all moderate k.

5.2.2 The IMA(2,2) Model

Exhibit 5.5 Simulation of an IMA(2,2) Series with $\theta_1 = 1$ and $\theta_2 = -0.6$

smooth change



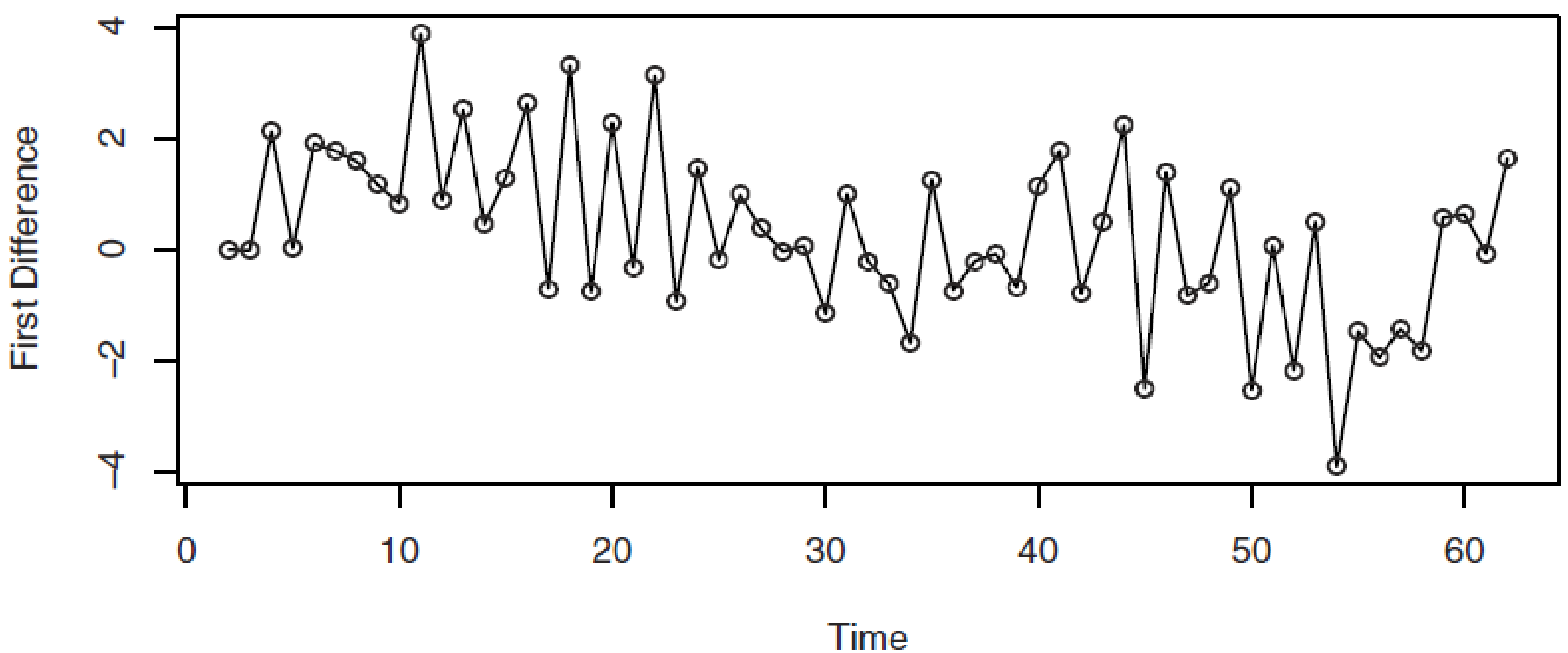
increasing variance

```
> data(ima22.s)
> plot(ima22.s,ylab='IMA(2,2) Simulation',type='o')
```

strong, positive neighboring correlations

5.2.2 The IMA(2,2) Model

Exhibit 5.6 First Difference of the Simulated IMA(2,2) Series



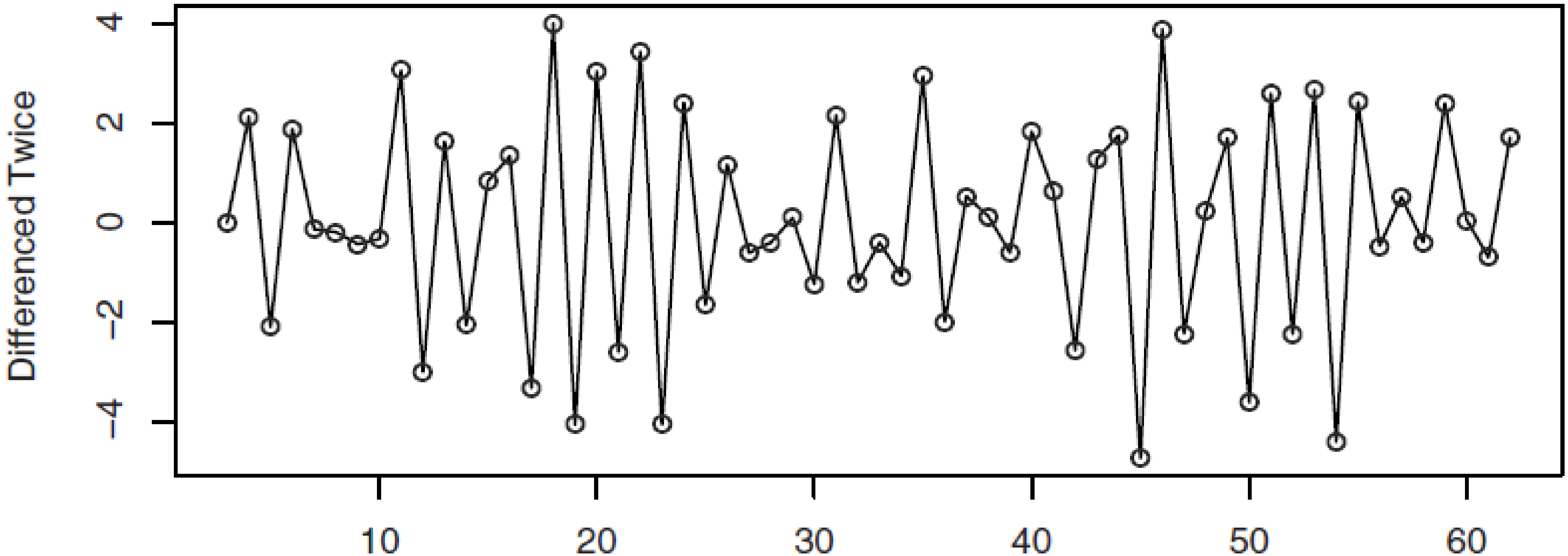

```
> plot(diff(ima22.s),ylab='First Difference',type='o')
```

IMA(1,2) model

5.2.2 The IMA(2,2) Model

Exhibit 5.7 **Second Difference of the Simulated IMA(2,2) Series**

$$\rho_1 = -0.678 \text{ and } \rho_2 = 0.254$$



stationary MA(2) model

```
> plot(diff(ima22.s,difference=2),ylab='Differenced
   Twice',type='o')
```

5.2.3 The ARI(1,1) Model

The ARI(1,1) process will satisfy $|\phi| < 1$

$$Y_t - Y_{t-1} = \phi(Y_{t-1} - Y_{t-2}) + e_t$$

or

$$Y_t = (1+\phi)Y_{t-1} - \phi Y_{t-2} + e_t$$

5.2.3 The ARI(1,1) Model

To find the ψ-weights in this case, we use a technique that will generalize to arbitrary ARIMA models. ***The ψ-weights can be obtained by equating like powers of x in the identity***:

$$(1-\phi_1 x-\phi_2 x^2-\cdots-\phi^p x^p)(1-x)^d(1+\psi_1 x+\psi_2 x^2+\psi_3 x^3+\cdots)$$

$$=(1-\theta_1 x-\theta_2 x^2-\theta_3 x^3-\cdots-\theta_q x^q)$$

$$(1-\phi x)(1-x)(1+\psi_1 x+\psi_2 x^2+\psi_3 x^3+\cdots)=1$$

or

$$[1-(1+\phi)x+\phi x^2](1+\psi_1 x+\psi_2 x^2+\psi_3 x^3+\cdots)=1$$

5.2.3 The ARI(1,1) Model

$$-(1+\phi)+\psi_1 = 0$$

$$\phi-(1+\phi)\psi_1+\psi_2 = 0$$

$$\psi_k = (1+\phi)\psi_{k-1}-\phi\psi_{k-2} \qquad \text{for } k \geq 2$$

$$\psi_k = \frac{1-\phi^{k+1}}{1-\phi} \quad \text{for } k \geq 1$$

5.3 Constant Terms in ARIMA Models

A nonzero constant mean, μ, in a stationary ARMA model $\{W_t\}$ can be accommodated in either of two ways.

(i):

$$W_t - \mu = \phi_1(W_{t-1} - \mu) + \phi_2(W_{t-2} - \mu) + \cdots + \phi_p(W_{t-p} - \mu)$$
$$+ e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

(ii):

$$W_t = \theta_0 + \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p}$$
$$+ e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

5.3 Constant Terms in ARIMA Models

$$\mu = \theta_0 + (\phi_1 + \phi_2 + \cdots + \phi_p)\mu$$

$$\mu = \frac{\theta_0}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}$$

$$\theta_0 = \mu(1 - \phi_1 - \phi_2 - \cdots - \phi_p)$$

Since the alternative representations are equivalent, we shall use whichever parameterization is convenient.

5.3 Constant Terms in ARIMA Models

What will be the effect of a nonzero mean for $\{W_t\}$ on the undifferenced series Y_t?

Consider the IMA(1,1) case with a constant term.

$$Y_t = Y_{t-1} + \theta_0 + e_t - \theta e_{t-1}$$

$$W_t = \theta_0 + e_t - \theta e_{t-1}$$

$$Y_t = e_t + (1-\theta)e_{t-1} + (1-\theta)e_{t-2} + \cdots + (1-\theta)e_{-m} - \theta e_{-m-1} \boxed{+ (t+m+1)\theta_0}$$

linear deterministic time with slope θ0

5.3 Constant Terms in ARIMA Models

An equivalent representation of the process would be

$$Y_t = Y_t' + \beta_0 + \beta_1 t$$

where Y_t' is an IMA(1,1) series with $E(\nabla Y_t') = 0$ and $E(\nabla Y_t) = \beta_1$.

For a general ARIMA(*p*,*d*,*q*) model where $E(\nabla^d Y_t) \neq 0$, it can be argued that $Y_t = Y_t' + \mu_t$, where μ_t is a deterministic polynomial of degree *d* and Y_t' is ARIMA(*p*,*d*,*q*) with $EY_t' = 0$. With *d* = 2 and $\theta_0 \neq 0$, a quadratic trend would be implied.

5.4 Other Transformations

Consider series where increased dispersion seems to be associated with higher levels of the series—***the higher the level of the series, the more variation*** there is around that level and conversely.

Specifically, suppose that $Y_t > 0$ for all t and that

$$E(Y_t) = \mu_t \qquad \text{and} \qquad \sqrt{Var(Y_t)} = \mu_t\sigma$$

$$\log(Y_t) \approx \log(\mu_t) + \frac{Y_t - \mu_t}{\mu_t}$$

$$E[\log(Y_t)] \approx \log(\mu_t) \qquad \text{and} \qquad Var(\log(Y_t)) \approx \sigma^2$$

5.4.1 Percentage Changes and Logarithms

Suppose *Yt* tends to have relatively stable percentage changes from one time period to the next. Specifically, assume that

$$Y_t = (1 + X_t)Y_{t-1}$$

where $100X_t$ is the percentage change (possibly negative) from Y_{t-1} to Y_t.

$$\log(Y_t) - \log(Y_{t-1}) = \log\left(\frac{Y_t}{Y_{t-1}}\right)$$

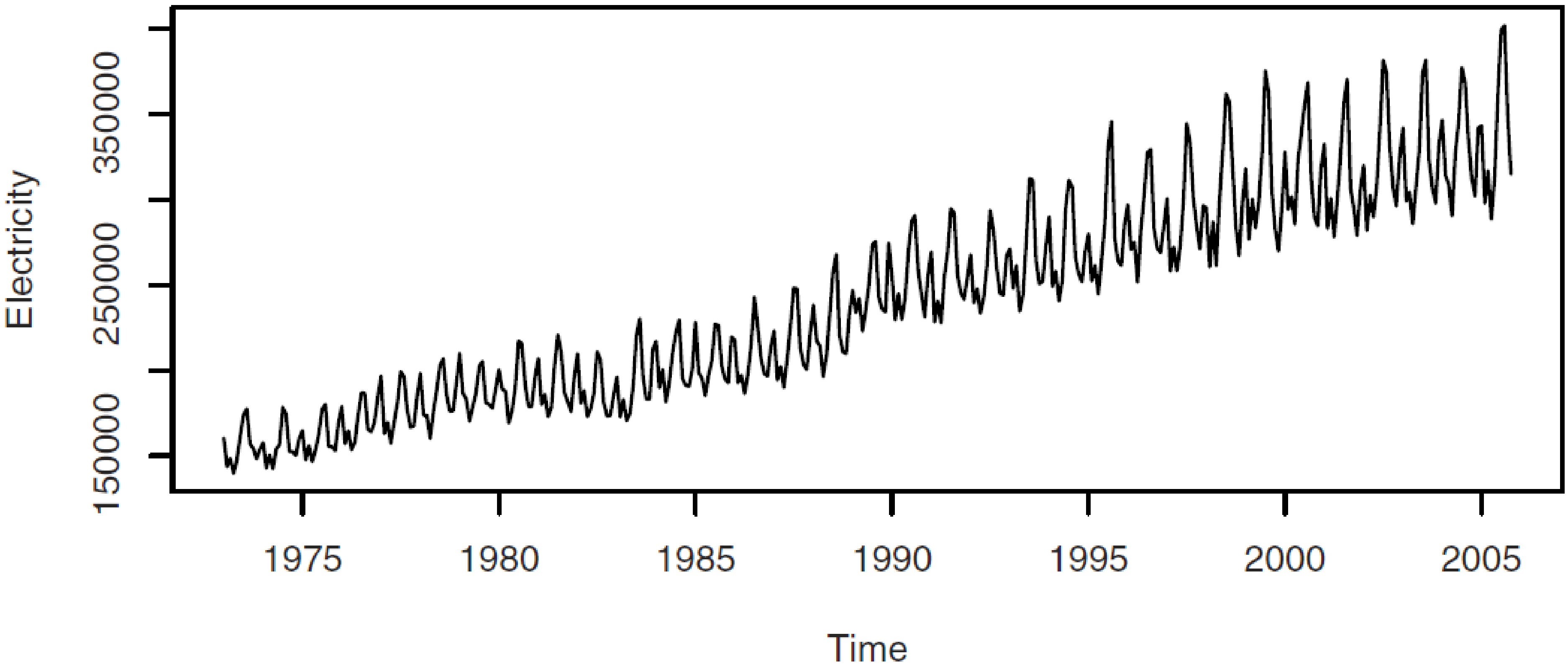
stationary

$$= \log(1 + X_t)$$

$$\nabla[\log(Y_t)] \approx X_t \qquad \log(1+X_t) \approx X_t$$

5.4.1 Percentage Changes and Logarithms

Exhibit 5.8 U.S. Electricity Generated by Month

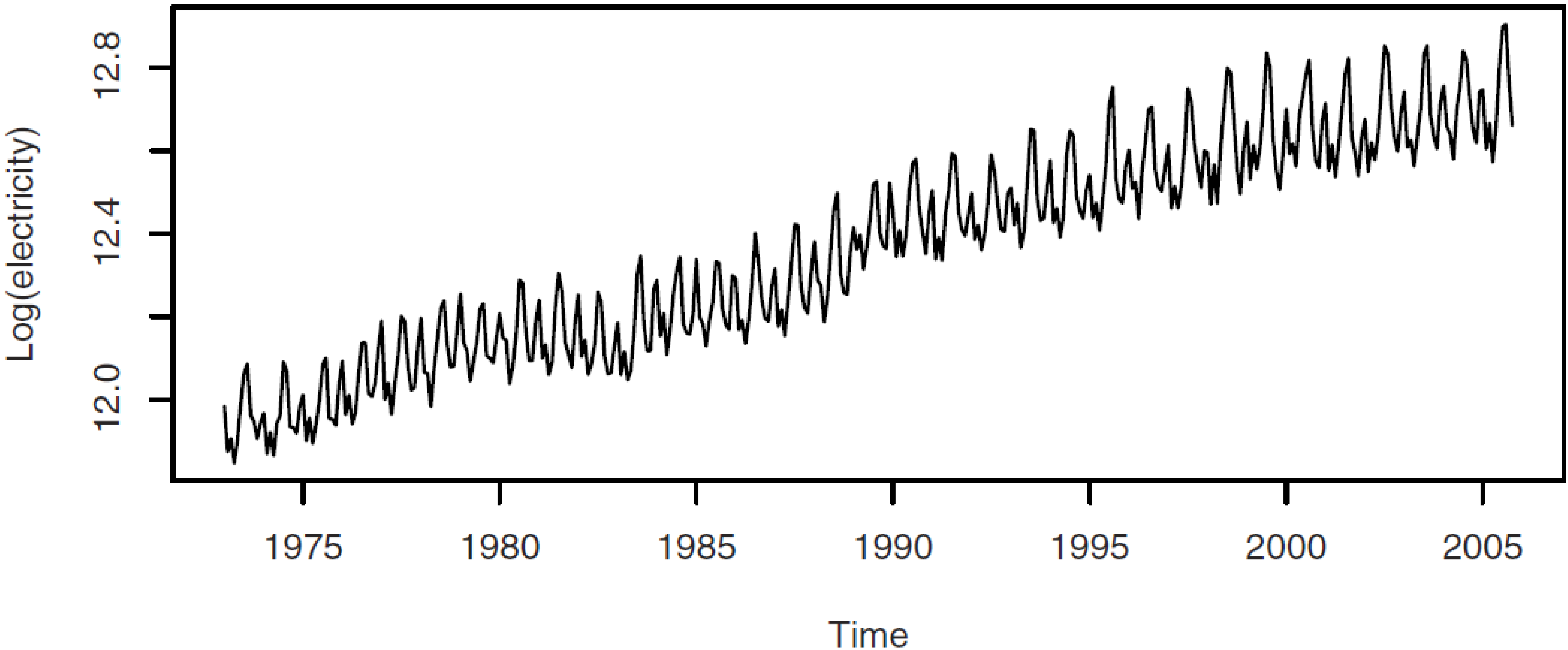



```
> data(electricity); plot(electricity)
```

The higher values display considerably more variation than the lower values.

5.4.1 Percentage Changes and Logarithms

Exhibit 5.9 Time Series Plot of Logarithms of Electricity Values

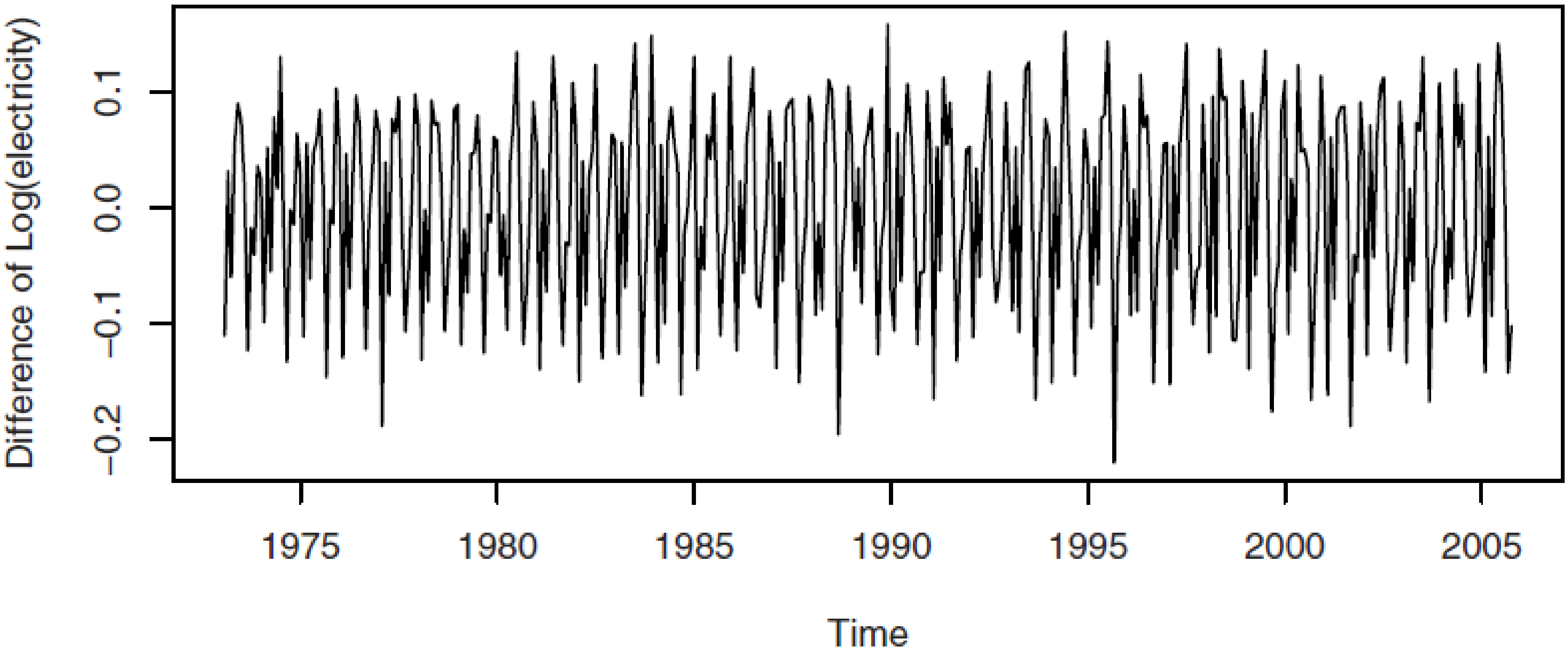



```
> plot(log(electricity),ylab='Log(electricity)')
```

The amount of variation around the upward trend is much more uniform across high and low values of the series.

5.4.1 Percentage Changes and Logarithms

Exhibit 5.10 Difference of Logarithms for Electricity Time Series




```
> plot(diff(log(electricity)),
   ylab='Difference of Log(electricity)')
```

A stationary model may be appropriate

5.4.2 Power Transformations

For a given value of the parameter λ, the power transformation is defined by

$$g(x) = \begin{cases} \dfrac{x^{\lambda} - 1}{\lambda} & \text{for } \lambda \neq 0 \\ \log x & \text{for } \lambda = 0 \end{cases}$$

5.5 Summary

This chapter introduced the concept of differencing to ***induce stationarity on certain nonstationary processes***. This led to the important integrated autoregressive moving average models (ARIMA). The properties of these models were then thoroughly explored. Other transformations, namely percentage changes and logarithms, were then considered. More generally, power transformations were introduced as useful transformations to stationarity and often normality.

作业

自己写程序完成5.10题

5.11，5.12，5.13题