Topic 6: Exponential Smoothing, Seasonal Model, and Regressions with Time Series Errors

Zongye Huang

International School of Economics and Management
Capital University of Economics and Business

March 28, 2017

Outline

Outline

Exponential Smoothing I

- A commonly used method in forecasting is the exponential smoothing. Consider the 1-step ahead forecast of a time series x_t. Suppose that the forecast origin is h and all past data are available.
- Under the general belief that the serial dependence of x_t decays exponentially, one can use a weighted average of the past data to predict x_{h+1} with weights decaying exponentially.

$$\begin{aligned}\hat{x}_{h+1} &\propto \omega x_h + \omega^2 x_{h-1} + \ldots = \sum_{j=1}^{\infty} \omega^j x_{h+1-j} \\ &= (1-\omega)\sum_{j=1}^{\infty} \omega^j x_{h+1-j}\end{aligned}$$

where ω is a positive real number in (0,1) referred to as the **discounting rate**.

Exponential Smoothing II

- This technique to produce forecasts is called the **exponential smoothing** method.
  - It has been widely used in practice because the technique says that more recent data contribute more in predicting x_{h+1}.
  - $\sum_{j=1}^{\infty} \omega^j = \frac{1}{1-\omega}$. The term $1-\omega$ is a scale adjustment factor.
- The exponential smoothing is a special case of the ARIMA models.
  - Specifically, consider the ARIMA(0,1,1) model

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

where $\theta \in (0,1)$. Using the AR representation,

$$x_{h+1} = (1-\theta)\left[\theta x_h + \theta^2 x_{h-1} + ...\right].$$

Exponential Smoothing III

- Therefore, the 1-step ahead forecast is

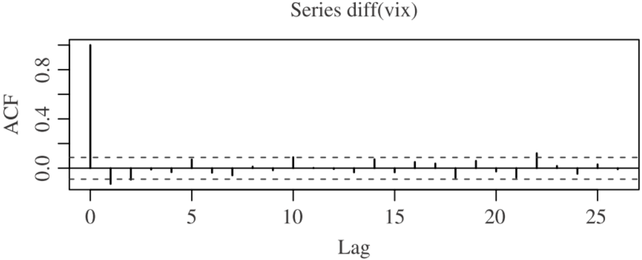
$$\hat{x}_h(1) = (1-\theta)\left[\theta x_h + \theta^2 x_{h-1} + \ldots\right],$$

which is precisely the exponential smoothing with $\theta = \omega$.

- Treating exponential smoothing as a special ARIMA(0,1,1) model has several advantages.
  - First, the discounting parameter θ can be estimated via the maximum likelihood method.
  - Second, one can identify and check the adequacy of the exponential smoothing method via the model building procedure of ARIMA models.

Example 2.5 I

- Consider the daily volatility index (VIX) of Chicago Board Options Exchange (CBOE) from May 1, 2008 to April 19, 2010. The data are obtained from the CBOE web.



Example 2.5 II

- The lag-1 ACF is significantly different from 0 at the 5% level,
  - An MA(1) model is identified for the differenced series.
  - Let $x_t = ln(VIX_t)$. The fitted model is

$$(1 - B)x_t = (1 - 0.163B)a_t$$

- Is this model adequate?

Example 2.5: Adequacy Test

- We calculate the Ljung–Box statistics

```
> Box.test(m1$residuals,lag=10,type='Ljung')
        Box-Ljung test

data:  m1$residuals
X-squared = 14.2536, df = 10, p-value = 0.1617
```

Example 2.5: Adequacy Test

- We calculate the Ljung–Box statistics

```
> Box.test(m1$residuals,lag=10,type='Ljung')
        Box-Ljung test

data:  m1$residuals
X-squared = 14.2536, df = 10, p-value = 0.1617
```

- The P-value

```
> pp=1-pchisq(14.25,9)
> pp
[1] 0.1137060
```

- We have $Q(10) = 14.25$ with p-value 0.11, based on a chi-squared distribution with 9 degrees of freedom.
  - The fitted ARIMA(0,1,1) model is adequate.
- In this particular instance, one can employ the exponential smoothing to predict the log series of daily VIX index.

Outline

Seasonal Time Series I

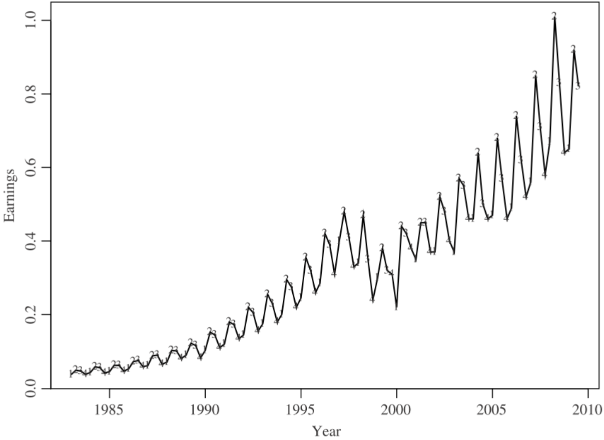


Figure 2.2. Quarterly earnings per share of Coca-Cola Company from the first quarter of 1983 to the third quarter of 2009.

Seasonal Time Series II

- Seasonal time series:
  - Some financial time series such as quarterly earnings per share of a company exhibits certain cyclical or periodic behavior.
  - The seasonal pattern repeats itself every year.
    - The periodicity of quaterly date is 4.
    - If monthly data are considered, then the periodicity is 12.
  - Seasonal time series is also highly relevant in empirical studies of the prices of weather-related derivatives and energy futures.
    - Most environmental time series exhibits strong seasonal behavior.
  - In forecasting, seasonality is as important as other characteristics of the data and must be handled accordingly.

Seasonal Time Series III

- Seasonal adjustment
  - The procedure to remove seasonality from a time series.
    - 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.
    - Most economic data published by the US government are seasonally adjusted (e.g., the growth rate of GDP and the unemployment rate).

Outline

Log Transformation I

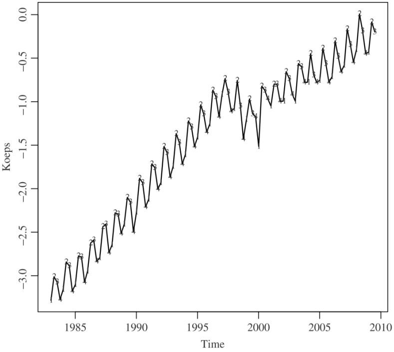


Figure 2.20. Time plot of quarterly earnings per share of the Coca-Cola from 1983.I to 2009.III: log earnings.

Log earnings per share of the Coca-Cola Company

Log Transformation II

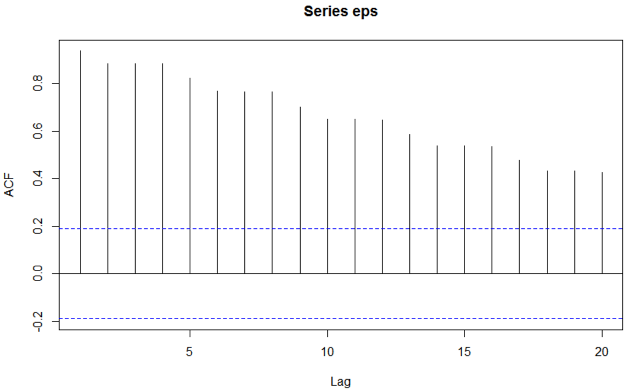
- We took the log transformation for two reasons.
    - First, it is used to handle the exponential growth of the series.
        - Indeed, the plot shows that the log earnings increased linearly and the linear trend continued even after the 1998 disturbance, albeit with a different rate.
    - Second, the transformation is used to stabilize the variability of the series.
        - The increasing pattern in variability of quarterly earnings disappears after the transformation.

Log Transformation III

- Log transformation is commonly used in analysis of financial and economic time series.
  - In some cases, one may need to add a positive constant to every data point before taking the transformation.
  - In this particular instance, all earnings are positive so that no adjustment is needed before taking the transformation.

Seasonal Differencing I

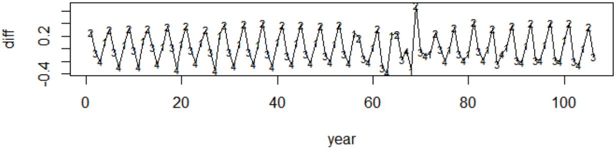
- Sample ACF

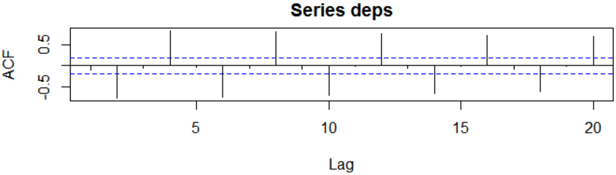


Seasonal Differencing II

- To handle such strong serial correlations is to consider the first differenced series, which induce stationarity

$$\Delta x_t = x_t - x_{t-1} = (1 - B)x_t$$





Seasonal Differencing III

- The top panel shows the time plot of Δx_t.
    - The differencing successfully removes the upward trend of the data
    - The series now shows a very strong seasonal pattern.
- The lower panel gives the sample ACF of Δx_t .
    - The autocorrelations are large at lags, which are multiples of the periodicity 4.

- We then consider a seasonal difference of Δx_t to handle the strong seasonal pattern

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

- The operation $\Delta_4 = (1 - B^4)$ is called a seasonal differencing.

Seasonal Differencing IV

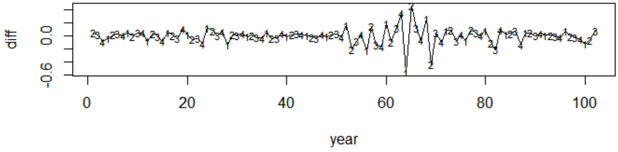
- In general, for a seasonal time series y_t with periodicity s, **seasonal differencing** means

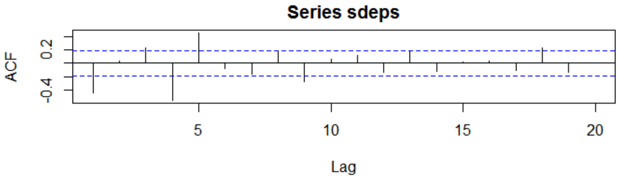
$$\Delta_s y_t = y_t - y_{t-s} = (1 - B^s) y_t.$$

It is commonly used in business and finance. For example, in reporting quarterly earnings of a company, news media often compare the earnings with that of the same quarter one year earlier.

  - The conventional difference $\Delta y_t = y_t - y_{t-1} = (1 - B) y_t$. is referred to as the *regular differencing*.

Seasonal Differencing V





- The ACFs are negative and statistically significant at lags 1 and 4.
- The ACF is positive and statistically significant at lag 5.
- The ACF is positive and marginally significant at lag 3.

Seasonal Differencing VI

- These observed characteristics are common among empirical seasonal time series.
- They lead to the development of multiplicative seasonal models introduced in the next section.

Outline

Multiplicative Seasonal Models I

- A statistical model whose autocorrelations possess the behavior shown by the sample ACF of $(1-B^4)(1-B)x_t$ is the multiplicative seasonal model.
- A simple multiplicative seasonal model assumes the form

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

where $|\theta| < 1$ and $|\Theta| < 1$.

  - This model is referred to as the *airline model* (Box and Jenkins 1976), used to study the airline passenger data.
  - The AR part of the model simply consists of the regular and seasonal differences.
  - The MA part involves two parameters.

Multiplicative Seasonal Models II

- Let $\omega_t = (1 - B^s)(1 - B)x_t$, $s > 1$, and focusing on the MA part

$$\begin{aligned} \omega_t &= (1 - \theta B)(1 - \Theta B^s)a_t \\ &= a_t - \theta a_{t-1} - \Theta a_{t-s} + \theta\Theta a_{t-s-1} \end{aligned}$$

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

Multiplicative Seasonal Models III

Consequently, the ACF of the w_t series is given by

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

and $\rho_\ell = 0$ for $\ell > 0$ and $\ell \neq 1, s-1, s, s+1$. For example, if w_t is a quarterly time series, then $s = 4$ and for $\ell > 0$, the ACF ρ_ℓ is nonzero at lags 1, 3, 4, and 5 only. This is indeed the case for the log quarterly earnings of the Coca-Cola Company.

It is interesting to compare the prior ACF with those of the MA(1) model $y_t = (1-\theta B)a_t$ and the MA(s) model $z_t = (1-\Theta B^s)a_t$. The ACF of y_t and z_t series are

$$\rho_1(y) = \frac{-\theta}{1+\theta^2}, \quad \text{and} \quad \rho_\ell(y) = 0, \quad \ell > 1,$$

$$\rho_s(z) = \frac{-\Theta}{1+\Theta^2}, \quad \text{and} \quad \rho_\ell(z) = 0, \quad \ell > 0, \quad \neq s.$$

We see that (i) $\rho_1 = \rho_1(y)$, (ii) $\rho_s = \rho_s(z)$, and (iii) $\rho_{s-1} = \rho_{s+1} = \rho_1(y) \times \rho_s(z)$.

Multiplicative Seasonal Models IV

- Rewriting the model as

$$\frac{1-B}{1-\theta B}\left[\left(\frac{1-B^s}{1-\Theta B^s}\right)x_t\right] = a_t$$

- Let $y_t = \frac{1-B^s}{1-\Theta B^s}x_t$,

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

- The airline model can be regarded as an exponential smoothing model on top of another exponential smoothing model.
  - One exponential smoothing is for the usual serial dependence.
  - The other one is for the seasonal dependence.

Multiplicative Seasonal Models V

- A **nonmultiplicative** seasonal MA model

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

where $|\theta| < 1$ and $|\Theta| < 1$.

  - It is easy to see that for this model $\rho_{s+1} = 0$.
  - A multiplicative model and its nonmultiplicative counterpart use the same number of parameters, but the multiplicative model has more nonzero ACFs.

Example 2.6 I

- We apply the airline model to the log series of quarterly earnings per share of Coca-Cola from 1983 to 2009.
  - On the basis of the exact likelihood method, the fitted model is

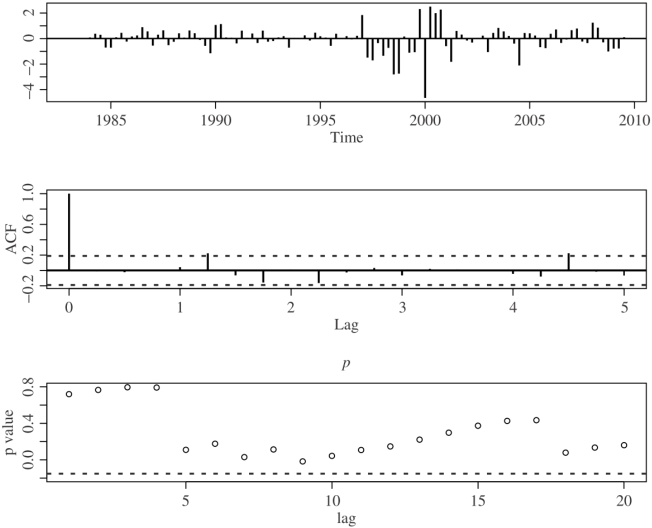
    $$(1-B)(1-B^4)x_t = (1-0.4096B)(1-0.8203B^4)a_t,$$

    where standard errors of the two MA parameters are 0.0866 and 0.0743, respectively, $\hat{\sigma}_a^2 = 0.00724$.
  - The Ljung–Box statistics of the residuals show Q(12) = 13.20 with p-value 0.35 when the degrees of freedom is 12. With adjustment to 10 degrees of freedom, the p-value becomes 0.21.

Example 2.6 II

- We can use command **tsdiag** to verify the adequacy of the fitted model.
    - The first plot shows the standardized residuals.
        - This plot can be used to examine the iid assumption of the residuals and to spot possible outliers in the data.
    - The second plot is the autocorrelations of the residuals.
        - Ideally, all the ACF of the residuals should be within the limit of two standard errors.
    - The third plot provides the p-values of the Ljung–Box statistics for several values of m.
        - If the fitted model is adequate in describing the serial dependence of the data, then all p-values should be greater than the type I error.

Example 2.6 III



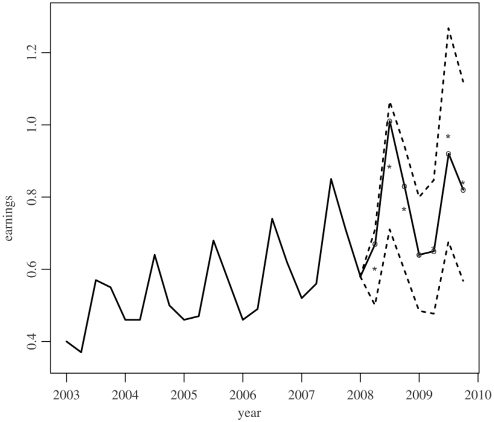
Example 2.6 IV

- Forecasting
  - To illustrate the forecasting performance of the fitted seasonal model, we reestimate the model using the first 100 observations, that is, from 1983 to 2007, and reserve the last seven data points for forecasting evaluation.
    - The fitted model becomes

$$(1-B)(1-B^4)x_t = (1-0.4209B)(1-0.8099B^4)a_t$$

      with $\sigma_a^2 = 0.00743$. We compute 1-step to 7-step ahead forecasts and their standard errors of the fitted model at the forecast origin $h = 100$.
    - An antilog transformation is taken to obtain forecasts of earnings per share using the relationship between normal and log-normal distributions.

Example 2.6 V


1.2
1.0
0.8
0.6
0.4
earnings
2003
2004
2005
2006
2007
2008
2009
2010
year

Example 2.6 VI

- Data are represented by solid line, observations during the forecasting period are marked by "o" point forecasts are indicated by "*" and the dashed lines represent 95% interval forecasts.
- The forecasts show a strong seasonal pattern and are close to the observed data. The actual earnings are all in the interval forecasts.

Outline

Seasonal Dummy Variable I

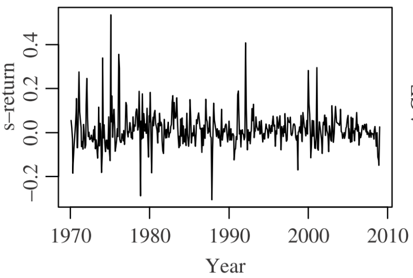
- When the seasonal pattern of a time series is stable over time (e.g., close to a deterministic function), dummy variables may be used to handle the seasonality.
  - By seasonal dummy variables, we mean the indicator variables for the seasons within a year.
  - For quarterly data, the dummy variables represent spring, summer, autumn and winter, respectively, and three of them are used in an analysis.
- Deterministic seasonality is a special case of the multiplicative seasonal model discussed earlier.
  - Specifically, if $\Theta = 1$, the above model contains a deterministic seasonal component.

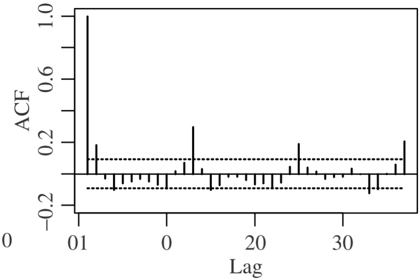
Seasonal Dummy Variable II

- Consequently, the same forecasts are obtained by using either dummy variables or a multiplicative seasonal model when the seasonal pattern is deterministic.
- The use of dummy variables can lead to inferior forecasts if the seasonal pattern is not deterministic.
- In practice, we recommend that the exact likelihood method should be used to estimate a multiplicative seasonal model, especially when the sample size is small or when there is the possibility of having a deterministic seasonal component.

Example 2.7 I

- To demonstrate the deterministic seasonal behavior, consider the monthly simple returns of the CRSP Decile 1 Index from January 1970 to December 2008 for 468 observations.





Example 2.7 II

- The time plot does not show any clear pattern of seasonality.
- The sample ACF of the return series contains significant lags at 12, 24, and 36, as well as lag 1.
- If seasonal ARMA models are entertained, a model in the form

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

where X_t denotes the monthly simple return. After removing the insignificant parameter, the fitted model becomes

$$(1-0.179B)(1-0.989B^{12})X_t = (1-0.913B^{12})a_t,\ \tilde{\sigma}_a^2 = 0.00472.$$

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

Example 2.7 III

- We define the dummy variable for January,

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

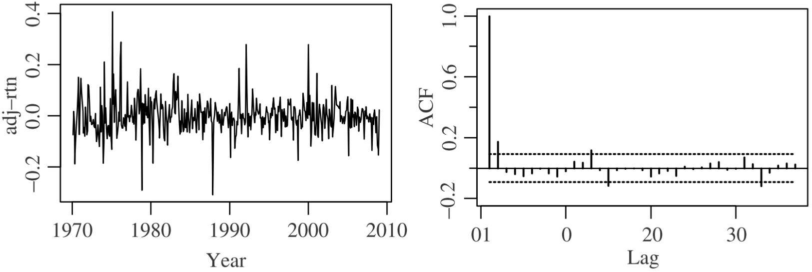
and employ the simple liner regression

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

- The fitted model is $X_t = 0.0029 + 0.1253\mathrm{Jan}_t + e_t$, where the standard errors of the estimates are 0.0033 and 0.0115, respectively.

Example 2.7 IV

- Look at the residual series



- From the sample ACF, serial correlations at lags 12, 24, and 36 largely disappear, suggesting that the seasonal pattern has been removed by the January dummy variable.
  - Consequently, the seasonal behavior in the monthly simple return of Decile 1 is mainly due to the January effect.

Outline

Regression Models with Time Series Errors I

- In many applications, the relationship between two time series is of major interest.
  - An obvious example is the Market Model in finance that relates the excess return of an individual stock to that of a market index.
  - The term structure of interest rates is another example in which the time evolution of the relationship between interest rates with different maturities is investigated.

Regression Models with Time Series Errors II

- Consider a linear regression in the form

$$y_t = \alpha + \beta x_t + e_t,$$

  - If e_t is a white noise series, then the LS method produces consistent estimates.
  - However, it is common to see that the error term is serially correlated. In this case, we have a regression model with time series errors, and the LS estimates of α and β may not be consistent.

Two Interest Rates I

- For example, consider two US weekly interest rate series:
    1. x_{1t} : The 1-year treasury constant maturity rate,
    2. x_{3t} : The 3-year treasury constant maturity rate.
- These two interest rates are highly correlated.

Two Interest Rates II

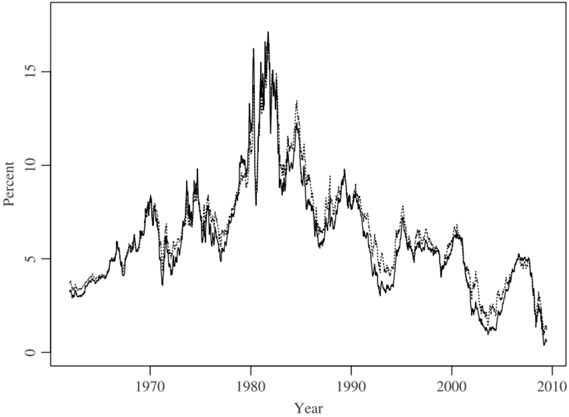


Figure 2.26. Time plots of US weekly interest rates (in percentages) from January 5, 1962 to April 10, 2009. The solid line represents the treasury 1-year constant maturity rate and the dashed line represents the treasury 3-year constant maturity rate.

Two Interest Rates III

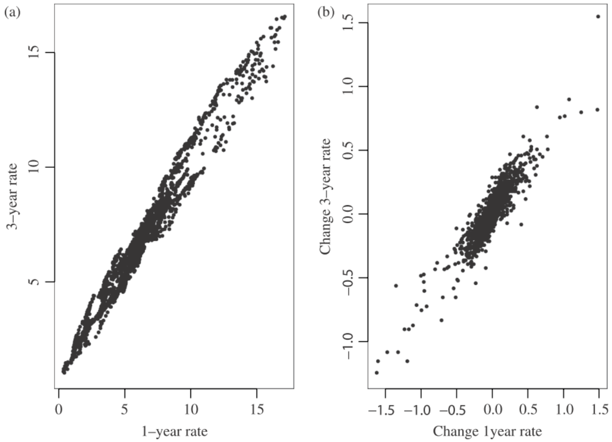


Figure 2.27. Scatter plots of US weekly interest rates from January 5, 1962 to April 10, 2009: (a) 3-year rate versus 1-year rate and (b) changes in 3-year rate versus changes in 1-year rate.

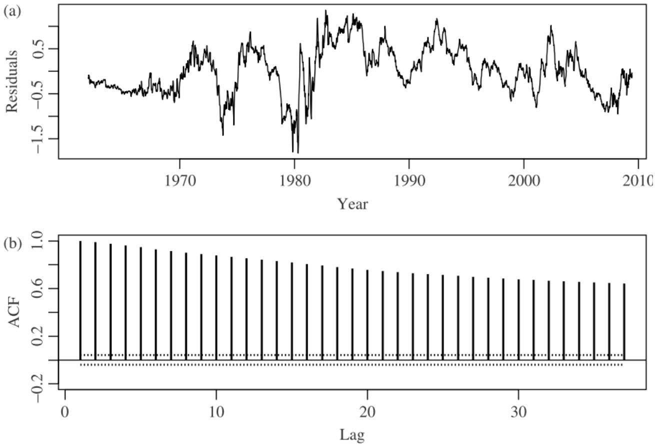
Two Interest Rates IV

- A naive way to describe the relationship between the two interest rates is to use the simple model $x_{3t} = \alpha + \beta x_{1t} + e_t$.
    - This results in a fitted model

$$x_{3t} = 0.832 + 0.930 x_t + e_t, \ \hat{\sigma}_e = 0.523.$$

        - $R^2 = 0.965$.
        - The model is seriously inadequate. In particular, the sample ACF of the residuals is highly significant and decays slowly, showing the pattern of a unit-root nonstationary time series.

Two Interest Rates V



Two Interest Rates VI

- The unit-root behavior of both interest rate series and the residuals leads to the consideration of the change series of interest rates

  1. $c_{1t} = x_{1t} - x_{1,t-1} = (1-B)x_{1t}$ for $t \geq 2$: changes in the 1-year interest rate;
  2. $c_{3t} = x_{3t} - x_{3,t-1} = (1-B)x_{3t}$ for $t \geq 2$: changes in the 3-year interest rate;

  - Consider a regression model $c_{3t} = \beta c_{1t} + e_t$

$$c_{3t} = 0.782 c_{1t} + e_t$$

$R^2 = 0.825.$

Two Interest Rates VII

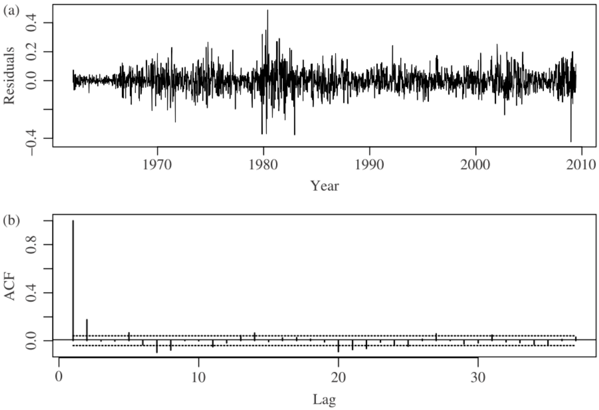


Figure 2.30. Residual series of the linear regression (Eq 2.47) for two change series of US weekly interest rates: (a) time plot and (b) sample ACF.

Two Interest Rates VIII

- The ACF shows some significant serial correlations in the residuals, but magnitudes of the correlations are much smaller.
- This weak serial dependence in the residuals can be modeled by using the simple time series models.
- We have a linear regression with time series errors.

A Simple Approach I

- To build a linear regression model with time series errors.
  - We employ a simple time series model for the residual series and estimate the whole model jointly.
- From the sample ACF of the residuals, we specify an MA(1) model for the residuals and modify the linear regression model to

$$c_{3t} = \beta c_{1t} + e_t, \;\; e_t = a_t - \theta_1 a_{t-1}$$

  - We simply use an MA(1) model, without the constant term, to capture the serial dependence in the error term.
  - The resulting model is a simple example of linear regression with time series errors.

A Simple Approach II

- If the time series model used is stationary and invertible, then one can estimate the model jointly via the maximum likelihood method. This is the approach we take using the command *arima* in R.
- The fitted model is

$$c_{3t} = 0.794c_{1t} + e_t, \; e_t = a_t - 0.1823a_{t-1}.$$

  - $R^2 = 0.831$.
  - The model no longer has a significant lag-1 residual ACF, even though some minor residual serial correlations remain at lags 4, 6, and 7.
  - The incremental improvement of adding additional MA parameters at lags 4, 6, and 7 to the residual equation is small.

A Simple Approach III

- Comparing the three models, we make the following observations.
  - First, the high R^2 96.5% and coefficient 0.930 of simple linear model are misleading because the residuals of the model show strong serial correlations.
  - Second, for the change series, R^2 and the coefficient of the second and third models are close.
    - In this particular instance, adding the MA(1) model to the change series provides only a marginal improvement.
    - This is not surprising because the estimated MA coefficient is small numerically, even though it is statistically highly significant.
  - Third, the analysis demonstrates that it is important to check residual serial dependence in linear regression analysis.

A Simple Approach IV

- Overall, the two weekly interest rate series are related as

$$x_{3t} = x_{3,t-1} + 0.794(x_{1t} - x_{1,t-1}) + a_t - 0.1823a_{t-1}.$$

  - The interest rates are concurrently and serially correlated.

Summary

- We outline a general procedure for analyzing linear regression models with time series errors as follows:
  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.
- 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.