Time Series Data Analysis and Forecasting

MSDS

Module 2

Topic Covered

- Introduction to time series regression model
- Least Squares Estimation
- statistical inference
- Prediction
- Variable selection
- regression models
- Classical decomposition, X11, SEATS, STL

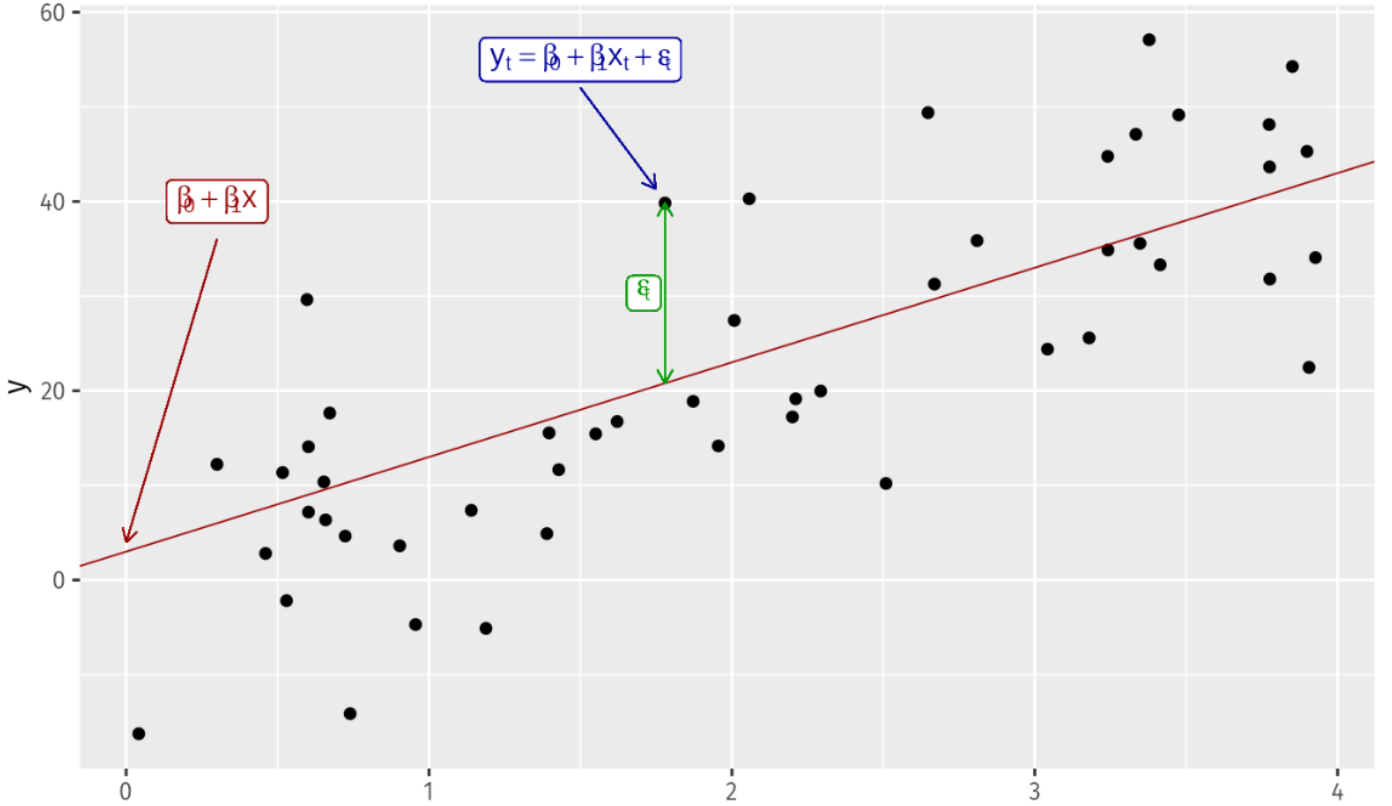
Introduction to Regression Model

- For regression analysis to be performed, data has to be stationary.
- Equation has to be rewritten in such a form that indicates a relationship among stationary variables.
- If a series is stationary then we can model it via an equation with fixed coefficents estimated from the past data
- concept is that forecast the time series of interest y assuming that it has a linear relationship with other time series x .
- The **forecast variable** y is sometimes also called the regressand, dependent or explained variable. The **predictor variables** x are sometimes also called the regressors, independent or explanatory variables.

Simple Linear Regression

- Regression model allows for a linear relationship between the forecast variable y and a single predictor variable x,

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t.$$



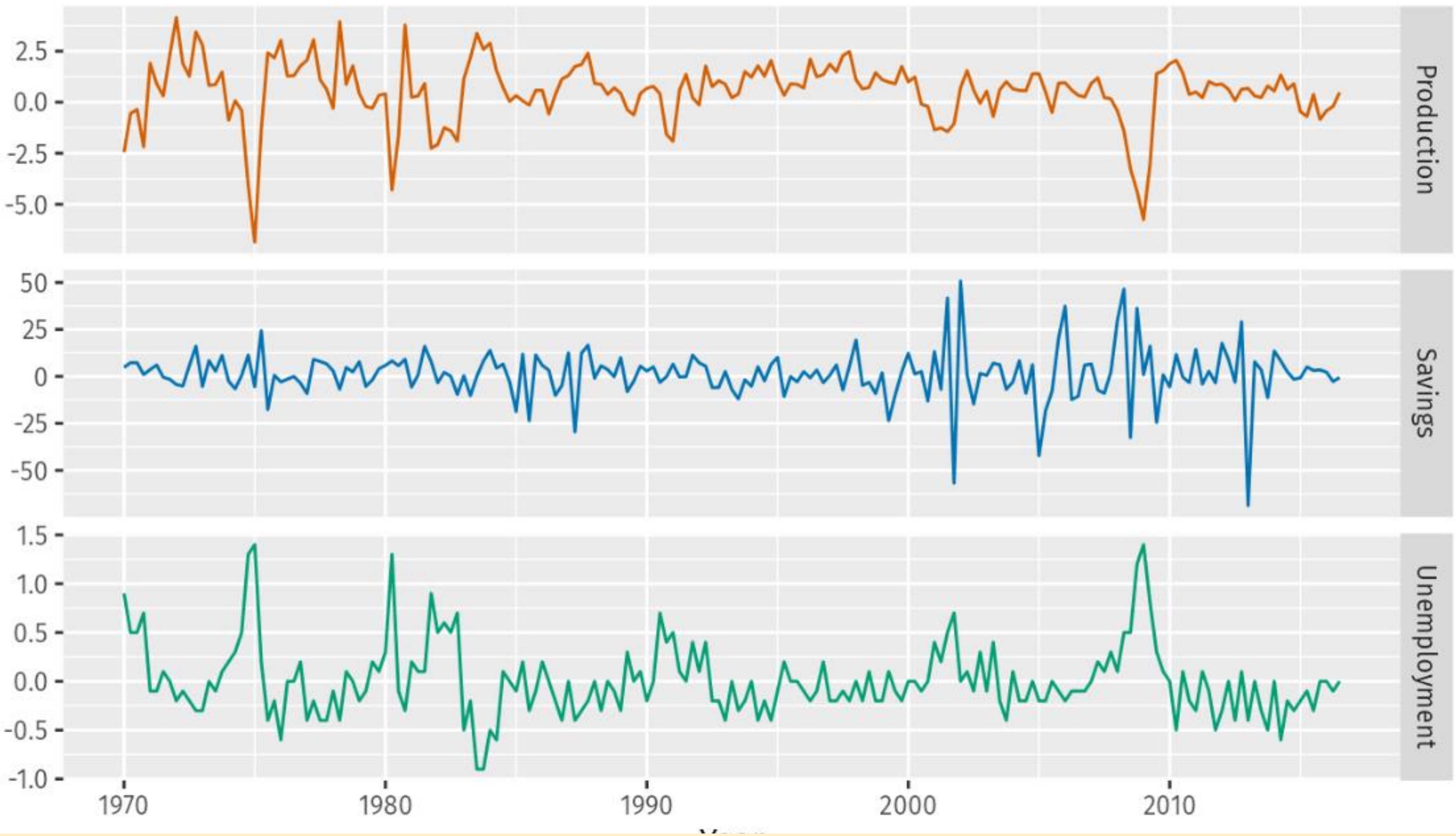
Multiple Linear Regression

- When there are two or more predictor variables, the model is called a **multiple regression model**. The general form of a multiple regression model is

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_k x_{k,t} + \varepsilon_t,$$

- Where y is predictor variable to be forecast and $x_1, x_2 \ldots,$ x_k are the k predictor variable.
- Each of predictor variable must be numerical.
- $x_1, x_2 \ldots,$ x_k coefficient measure the effect of each predictor after taking intro account the effects of all other predictors in model.
- Coefficient measure the marginal effects of predictor variable.

Example: US consumption expenditure



Quarterly percentage changes in industrial production and personal savings and quarterly changes in the unemployment rate for the US over the period 1970Q1-2016Q3

Least Square estimation(1)

- We have a collection of observation but the values of the coefficient β_0, β_1,… β_k are not known.
- least squares principle provides a way of choosing the coefficients effectively by minimising the sum of the squared errors.

$$\sum_{t=1}^{T} \varepsilon_t^2 = \sum_{t=1}^{T} (y_t - \beta_0 - \beta_1 x_{1,t} - \beta_2 x_{2,t} - \cdots - \beta_k x_{k,t})^2.$$

- This is called **least squares** estimation because it gives the least value for the sum of squared errors.
- Finding the best estimates of the coefficients is often called “fitting” the model to the data, or sometimes “learning” or “training” the model.
- The estimated coefficients are given by notation $\widehat{\beta_0}, \widehat{\beta_1},$, … $\widehat{\beta_k},$

Example: US consumption expenditure

A multiple linear regression model for US consumption is

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \beta_3 x_{3,t} + \beta_4 x_{4,t} + \varepsilon_t,$$

where y is the percentage change in real personal consumption expenditure, x_1 is the percentage change in real personal disposable income, x_2 is the percentage change in industrial production, x_3 is the percentage change in personal savings and x_4 is the change in the unemployment rate.

R code for regression modeling

```
fit.consMR <- tslm(
  Consumption ~ Income + Production + Unemployment + Savings,
  data=uschange)
summary(fit.consMR)
```

Fitted Values

- Predictions of y can be obtained by using the estimated coefficients in the regression equation and setting the error term to zero.
- Note that these are predictions of the data used to estimate the model, not genuine forecasts of future values of y

$$\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_{1,t} + \hat{\beta}_2 x_{2,t} + \cdots + \hat{\beta}_k x_{k,t}.$$

Goodness of Fit

- A common way to summarise how well a linear regression model fits the data is via the **coefficient of determination**, or R^2
- This can be calculated as the square of the correlation between the observed values y and the predicted $\hat{y}$ values.
- it reflects the proportion of variation in the forecast variable that is accounted for (or explained) by the regression model.

$$R^2 = \frac{\sum(\hat{y}_t - \bar{y})^2}{\sum(y_t - \bar{y})^2},$$

Standard Error of Regression

- Another measure of how well the model has fitted the data is the standard deviation of the residuals, which is often known as the "residual standard error"
- Calculated by equation,

$$\hat{\sigma}_e = \sqrt{\frac{1}{T-k-1}\sum_{t=1}^{T} e_t^2},$$

- K is number of predictor

Evaluating the Regression model

The differences between the observed y values and the corresponding fitted $\hat{y}$ values are the training-set errors or "residuals" defined as,

$$\begin{aligned} e_t &= y_t - \hat{y}_t \\ &= y_t - \hat{\beta}_0 - \hat{\beta}_1 x_{1,t} - \hat{\beta}_2 x_{2,t} - \cdots - \hat{\beta}_k x_{k,t} \end{aligned}$$

for $t = 1, \dots, T$. Each residual is the unpredictable component of the associated observation.

The residuals have some useful properties including the following two:

$$\sum_{t=1}^{T} e_t = 0 \quad \text{and} \quad \sum_{t=1}^{T} x_{k,t} e_t = 0 \qquad \text{for all } k.$$

Forecasting with Regression(1)

Recall that predictions of y can be obtained using

$$\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_{1,t} + \hat{\beta}_2 x_{2,t} + \cdots + \hat{\beta}_k x_{k,t},$$

which comprises the estimated coefficients and ignores the error in the regression equation. Plugging in the values of the predictor variables $x_{1,t}, \ldots, x_{k,t}$ for $t = 1, \ldots, T$ returned the fitted (training-sample) values of y. What we are interested in here, however, is forecasting *future* values of y.

Forecasting with Regression(2)

There are two types-

- **Ex-ante Forecast**

-Those forecast that are made using only the information that is available in advance.

-For example, the percentage change in US consumption for quarters following the end of the sample, should only use information that was available *up to and including* 2016 Q3.

-These are made in advance using whatever information is available at the time. Model requires forecasts of the predictors.

- **ex-post forecasts**

-Those that are made using later information on the predictors.

-For example, ex-post forecasts of consumption may use the actual observations of the predictors, once these have been observed.

-These are not genuine forecasts, but are useful for studying the behaviour of forecasting models.

Scenario Based forecasting

- In this setting, the forecaster assumes possible scenarios for the predictor variables that are of interest.
- For example, a US policy maker may be interested in comparing the predicted change in consumption when there is a constant growth of 1%
- 0.5% respectively for income and savings with no change in the employment rate, versus a respective decline of 1% and 0.5%, for each of the four quarters following the end of the sample.
- prediction intervals for scenario based forecasts do not include the uncertainty associated with the future values of the predictor variables.
- They assume that the values of the predictors are known in advance.

Building a predictive regression model

- The great advantage of regression models is that they can be used to capture important relationships between the forecast variable of interest and predictor variables.
- A major challenge is that in order to generate ex-ante forecasts, the model requires future values of each predictor.
- If scenario based forecasting is of interest then these models are extremely useful.

Time Series Decomposition

Time Series Decomposition

- Time series data can exhibit a variety of patterns, and it is often helpful to **split a time series** into several components, each representing
- While decomposing a time series into components, combine the trend and cycle into a single **trend-cycle** component (sometimes called the **trend** for simplicity).
- A time series as comprising three components: a trend-cycle component, a seasonal component, and a remainder component

Time Series Components

If we assume an additive decomposition, then we can write

$$y_t = S_t + T_t + R_t,$$

where y_t is the data, S_t is the seasonal component, T_t is the trend-cycle component, and R_t is the remainder component, all at period t. Alternatively, a multiplicative decomposition would be written as

$$y_t = S_t \times T_t \times R_t.$$

Time Series Components

- **additive decomposition** is most appropriate if the magnitude of the seasonal fluctuations, or the variation around the trend-cycle, does not vary with the level of the time series.
- When variation in the seasonal pattern, or variation around the trend-cycle, appears to be proportional to the level of the time series, then a **multiplicative decomposition** is more appropriate.
- Multiplicative decompositions are common with economic time series.
- When a log transformation has been used, this is equivalent to using a multiplicative decomposition because

$$y_t = S_t \times T_t \times R_t \quad \text{is equivalent to} \quad \log y_t = \log S_t + \log T_t + \log R_t.$$

Review Question

- Discuss Linear Regression Model?
- What is coefficient of determination?