STAT 485/685
Fitting Trends

Richard Lockhart

Simon Fraser University

STAT 485/685 — Fall 2017

Purposes of These Notes

- Show how to use Ordinary Least Squares to fit a *trend*
- Discuss some specific trends: seasonal, linear, cosine, quadratic.
- Use `R` to fit some trends, examine residuals.
- Discuss OLS SEs and impact of correlated errors.
- Sections 3.3 to 3.6 in text.

Fitting Trends

- Studying $Y_t = \mu_t + X_t$ where X_t has mean 0 and is stationary.
- Use data to get *fitted values* for μ_t, denoted $\hat{\mu}_t$.
- Method uses *Ordinary Least Squares*.
- Example 1: Linear trend: for all t

$$\mu_t = \beta_0 + \beta_1 t.$$

- Find *estimates* $\hat{\beta}_0$ and $\hat{\beta}_1$ by minimizing the *Error Sum of Squares*:

$$ESS = \sum_{t=1}^{T} (y_t - \beta_0 - \beta_1 t)^2$$

Regression: linear and multiple

- Can use calculus to find minimum; not part of this course
- Taking derivatives gives two equations to solve:

$$\sum_{t=1}^{T}(y_t - \beta_0 - \beta_1 t) = 0$$

and

$$\sum_{t=1}^{T} t(y_t - \beta_0 - \beta_1 t) = 0$$

- System of two linear equations in two unknowns
- Solution is (using $\bar{t} = \sum_{t=1}^{T} t/T = (T+1)/2$)

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{t} \text{ and } \hat{\beta}_1 = \frac{\sum(t-\bar{t})y_t}{\sum_t (t-\bar{t})^2}$$

Regression: linear and multiple

- Key feature: equation for Y_t has form

$$Y_t = \text{const}\beta_0 + \text{const}\beta_1 + \text{error}$$

- Other trend equations similar.
- Seasonal: $\mu_t = \mu_{t+S}$ where S is 12 for monthly, 4 for quarterly data.
- Quarterly data: four values of μ: $\mu_{Q1}, \ldots, \mu_{Q4}$:

$$Y_1 = 1 \cdot \mu_{Q1} + 0 \cdot \mu_{Q2} + 0 \cdot \mu_{Q3} + 0 \cdot \mu_{Q4} + \text{error}_1$$
$$Y_2 = 0 \cdot \mu_{Q1} + 1 \cdot \mu_{Q2} + 0 \cdot \mu_{Q3} + 0 \cdot \mu_{Q4} + \text{error}_2$$

and so on.

Linear models

- Linear trend, seasonal trend are examples of *linear* models:

$$Y_t = \beta_0 + d_{t1}\beta_1 + \cdots + d_{tp}\beta_p + \text{error}_t$$

- This are often written in the form

$$\mathbf{Y} = \mathbf{D}\boldsymbol{\beta} + \text{error}$$

- In this formula $\mathbf{Y}$ and 'error' are *column* vectors of T entries.
- $\mathbf{D}$ is a matrix with T rows and $p+1$ columns.
- $\boldsymbol{\beta}$ is a column vector with $p+1$ entries which are $\beta_0, \ldots, \beta_p$.
- This is a *linear* model but with *correlated* errors, usually.

Software

- We use `lm` ir `R` to estimate the coefficients in these models.
- I will do examples in class.
- We use a 'hat' on top of a letter to indicate an 'estimate'.
- So `lm` produces $\hat{\beta}_j$ for $j = 0, \ldots, p$.
- Now some specific formulas.

Periodic trends

- Two slightly different ways to write the model
- One mean for each month; no intercept term
- An intercept term (the January mean by default) and 11 monthly corrections to the January mean.
- OLS estimates for mean in February: average all February values!

Output from `lm`

- R code will be available in link from class slides.
- `fit = lm(y ~time(y))` fits linear trend to y.
- Output has columns Estimate, Standard Error, t-Statistic, P value, and some stars.
- Estimate column is correct – contains $\hat{\beta}_j$.
- Other columns wrong unless X_t, the error process, is white noise.

Residual Analysis

- Want to know if we have formula for trend right.
- So 'estimate' X by $\hat{X} = Y - D\hat{\beta}$.
- This is a vector.
- Use `residual(fit)` in `R`.
- Then convert residuals back to time series and plot as time series.
- Look for stationary process – constant mean, constant variability, etc.

Residual Analysis

- Can also assess normality of residuals.
- Text suggests $Q - Q$ plot (will show in class).
- And Shapiro-Wilk test; not clear P value is valid here.
- Look for clustering of large values in plot of $|\hat{X}_{t+1}|$ vs $|\hat{X}_t|$.
- Might show up as correlation.

Sample Autocorrelation Function

- Quick introduction now.
- Covariance between Y_t and Y_{t+k} same for all t: γ_k.
- Estimate this covariance using

$$\hat{\gamma}_k = \frac{1}{T} \sum_{t=1}^{T-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})$$

- Some writers prefer $T - 1$.
- Not important because we care about ρ_k estimated by

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{T-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^{T} (Y_t - \bar{Y})^2}.$$

Sample Autocorrelation Function in `R`

- Plotted in `R` by `acf`.
- Notice different k gives different numbers of terms.
- Not exactly averages.
- But consensus has formed in favour of these formulas.