Time Series

John Green

April 2024

Outline

Macroeconomic data

- Volatility clustering
- Seasonal adjustments
- Data revisions and sources
- Unemployment rate
- Labor force participation rate

Uses of time series econometrics

Time series analysis is a highly valuable skill in industry:

- Forecasting (eg interest rates, returns to a stock, GDP growth)
- Dynamic treatment effects: impact of a policy after 1 month, 3 months, 2 years...
- Dynamic system analysis

But, we will have to deal with problems like correlation over time.

Time series data

- Generally similar to panel data, but with only one unit
- Takes the form $\{Y_t\}_{1<=t<=T}$ with covariates $\{X_t\}_{1<=t<=T}$
- For now, we assume timing is evenly spaced and there are no missing periods

Lags

- Lags are very important in time series analysis
- X_{t-1} is often a very good predictor of X_t
- More generally, the j^{th} lag of Y_t is Y_{t-j}
- We might take a difference: $\Delta_j Y_t = Y_t - Y_{t-j}$

Log variables

- Oftentimes we will use *log* variables instead of the raw values
- Two benefits:
  - Makes exponential growth linear
  - Differences in logs are (approximately) equal to the percentage change:

$$\log(1+\alpha) \approx \alpha$$

    So long as α is "small"
  - So, $\log X_t - \log X_{t-1} = \log \frac{X_t}{X_{t-1}}$ is roughly the percentage change in X from $t-1$ to t

Autocorrelation

- Just like with panel data, we will often have correlation between lagged values
- First autocovariance of Y_t: $cov(Y_t, Y_{t-1})$
- First autocorrelation of Y_t: $corr(Y_t, Y_{t-1}) \equiv \rho_1$
- Can calculate for any lag j in the same way and get the j^{th} serial correlation coefficient

Forecasting and stationarity

- We will often be interested in making an out-of-sample (OOS) forecast of our variable
- This is similar to what we did before trying to target the oracle prediction
- We no longer care about causal interpretation of coefficients, OVB, etc.; we just want to get an accurate OOS prediction
- This requires that our data be *stationary*
  - Intuitively, our out-of-sample data needs to be similar to our sample Data
  - Technically, we need the joint distribution across time to be independent of the time period we are looking at