Recitation Eight and Nine
Time Series
Intro to Econometrics

Nathaniel Mark

November 29, 2018

1 What are Time Series

Our usual data that we have worked with has been what is called "Cross-sectional" data:

$$\{Y_i, X_i\}$$

Generally, we have assumed that each draw of these cross-sectional data came from a population distribution. We used these draws to make infer about the structure of the population distribution. ie we tried to determine how the world works by looking at data from the world. In time series, our data takes the form

$$\{Y_t, X_t\}$$

That is, our data are draws from the past. Here, we are trying to figure out how the world will look in the future, given how it has looked in the past – this is a bit of a different exercise.

What is the difference between these two kinds of data:

1. $$\{Y_t, X_t\}$$
   are not i.i.d. The value Y_4 is likely highly related to the value Y_5.
2. In order to make any inference about the future, we need to believe that the relationships we see in our data will remain the same in the future. Mathematically, we say that we need our data to be *stationary* if we are to use it to estimate time series models.
3. We care about forecasting – with is a special case of prediction.

2 Describing a Time Series

The population *autocovariance* of a time series $\{X_t\}$ is $Cov(X_t, X_{t-j})$. The estimator we use to estimate the autocovariance is $\frac{1}{T}\sum_{t=j+1}^{T}(Y_t - \bar{Y}_{j+1:T})(Y_{t-j} - \bar{Y}_{1:T-j})$.

The population *autocorrelation* of a time series $\{X_t\}$, if stationary, is $\rho_j \equiv Corr(X_t, X_{t-j}) = \frac{Cov(X_t,X_{t-j})}{Var(X_t)}$. The estimator we use to estimate the autocorrelation is $\frac{\frac{1}{T}\sum_{t=j+1}^{T}(Y_t-\bar{Y}_{j+1:T})(Y_{t-j}-\bar{Y}_{1:T-j})}{\frac{1}{T}\sum_{t=1}^{T}(Y_t-\bar{Y}_{1:T})^2}$.

These summary stats tell us the probabalistic relationship between past observations and present observations. That is, if $\rho_j > 0$, then a high X in last period means we are likely to have a high X this period.

3 Time Series Regression Models

Time series regression models have the form

$$Y_t = \beta_0+\beta_1 Y_{t-1}+\beta_2 Y_{t-2}+...+\beta_p Y_{t-p}+\delta_1 X_{t-1}^1+\delta_2 X_{t-2}^1+...+\delta_q X_{t-q}^1+\delta_1^2 X_{t-1}^2+\delta_2^2 X_{t-2}^2+...+\delta_q^2 X_{t-q}^2+...+\epsilon_t$$

Under Four assumptions, OLS estimates of the parameters in this model area "good" (unbiased, consistent, and asymptotically normal):

1. Exogeneity: the expected value of errors must not depend on the past values of Y or X. That is, for example, past Ys and Xs must not be correlated with anything that affects current Y and is not in the regression.
2. The time series Y, X must be stationary (We will talk about this later) and have non-permanent effects.
3. No large outliers
4. no perfect multicollinearity

We learned 2 names for special cases of time series regression models:

1. Autoregressive Models of order p (AR(p))

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + ... + \beta_p Y_{t-p} + \epsilon_t$$

2. Autoregressive Destributed Lag Models of order p,q

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + ... + \beta_p Y_{t-p} + \delta_1 X_{t-1}^1 + \delta_2 X_{t-2}^1 + ... + \delta_q X_{t-q}^1 + \epsilon_t$$

The same assumptions apply for these special cases.

4 Stationary

A time series $\{X_t\}$ is stationary if its joint distribution of $..., X_{t-1}, X_t, X_{t+1}, ...$ does not change over time. That is a time series is stationary if its probabalistic relationship does not change over time. If this relationship changes over time, how could be say anything about the future!

Here are a few of the common types of non-stationarity: (draw examples for students)

1. Deterministic Trend: $E[X_t]$ changes over time (e.g. increasing observations).
2. Heteroskedastic Data: $Var(X_t)$ changes over time (e.g. increasing variance).
3. Stochastic Trends/ Unit roots:
   These are harder to see by plotting the data. In fact a time series with the form $Y_t = .99999Y_{t-1} + \epsilon_t$ is stationary, while the time series with the form $Y_t = Y_{t-1} + \epsilon_t$ is non-stationary. They look almost the same! Why is $Y_t = Y_{t-1} + \epsilon_t$ non-stationary?

   $$Y_t = Y_{t-1} + \epsilon_t \Rightarrow Var(Y_t) = Var(Y_{t-1}) + \sigma_\epsilon^2$$

   For $Var(Y_t) = Var(Y_{t-1})$ to be true, σ_ϵ^2 must equal zero, which we know is not true.

These non-stationarities can lead to estimated of coefficients that are biased and inconsistent, and can make it impossible to estimate variances of estimates.

5 Forecasts

$$\text{Forecast Error} = Y_{T+1} - \hat{Y}_{T+1|T}$$

i.e. forecast error equals the Y that *will occur* next period minus our best estimate of the the Y next period.

$$\text{Expected Forecast Error} = E_T[Y_{T+1} - \hat{Y}_{T+1|T}]$$

If our model is absolutely correct, then expected forecast error is 0 as $E[Y_{T+1}] = \hat{Y}_{T+1|T}$. But, our estimates are never absolutely correct, unless we have infinite data (which we do not). Therefore, $E[Y_{T+1}] \neq \hat{Y}_{T+1|T}$.

However, we don't just care about the mean of our forecasts. We also care about how likely our forecast is to be correct. To analyze this, we use Mean Square Forecasting Error:

$$\begin{aligned}\text{MSFE} &= E_T[(Y_{T+1} - \hat{Y}_{T+1|T})^2] \\ &= E_T[(E[Y_{T+1}] + \epsilon_t - \hat{Y}_{T+1|T})^2] \\ &= Var[(E[Y_{T+1}] - \hat{Y}_{T+1|T})^2] + \sigma_\epsilon^2\end{aligned}$$

That is, MSFE is caused by squared error of our estimated model and a error caused by the inherent uncertainty in the world about the future.
As T goes to infinity, our estimates will get better and better, decreasing the first phrase to 0. The second phrase will never disappear however – MSFE will always be positive even with the best models because the world is uncertain.