Lecture Notes 9

1. Time Series Analysis

1. Examine data collected over time.

Models are simple
No independent variables, i.e. no x's
Models are mechanical

Time series is both
(i) Data
(ii) A process

Observations are related to each other over time, i.e. covariance
Each observation has
*Mean or expected value, i.e. called a first moment
Also includes expected products – related to covariance
*A variance, i.e. called a second moment

Thus, time series analysis only examines data for patterns and then we use those patterns to forecast.

2. Sample Autocorrelation Function (ACF)

Usually we plot the ACF
A plot to analyze the data

You have a time series, $X_1, \cdots X_n$
Time series has n observations

(i) You calculate the sample mean

$$\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$$

(ii) Then calculate the autocovariance function

$$\hat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-h}\left(X_{t-h} - \bar{X}\right)\left(X_t - \bar{X}\right)$$

h is the number of lags

If h =1 , then $\hat{\gamma}(1) = \frac{1}{n}\sum_{t=1}^{n-1}(X_{t-1} - \overline{X})(X_t - \overline{X})$

Thus, this is almost a covariance

i.e. COV(X_t, X_{t-1}) of X_t with past values

If h = 0 , then $\hat{\gamma}(0) = \frac{1}{n}\sum_{t=1}^{n}(X_t - \overline{X})^2$

Thus, this is almost a variance

(iii) Then calculate an ACF

ACF is very similar to the standard Pearson correlation

$$\hat{p}(h) = \begin{Bmatrix} \frac{\hat{\gamma}(0)}{\hat{\gamma}(0)} = 1 & h = 0 & no\,lag \\ \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} & h = 1 & lag\,1 \\ \frac{\hat{\gamma}(2)}{\hat{\gamma}(0)} & h = 2 & lag\,2 \end{Bmatrix}$$

Usually this is in a bar graph with 95% Confidence Intervals

Standard noise is distributed ~N(0, 1)

Form a 95% confidence interval

Standard error $S.E. = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{1}{n}}$

n is number of observations. The more observations, the lower is the S.E.

Form a 95% confidence interval

$$0 \pm 1.96\frac{1}{\sqrt{n}}$$

The average for noise is zero, while the 1.96 is the z critical value from a normal distribution.

3. First Order Moving Average (MA(1) Process)

$$X_t = Z_t + \theta \cdot Z_{t-1}$$
$$Z_t \sim white\ noise$$
$$Z_t \sim (0, \sigma^2)$$

Expected value of MA(1)

$$E(X_t) = E(Z_t + \theta \cdot Z_{t-1}) = E(Z_t) + E(\theta \cdot Z_{t-1}) = E(Z_t) + \theta \cdot E(Z_{t-1}) = 0 + \theta \cdot 0 = 0$$

Variance of the MA(1)

$$VAR(X_t) = E(X_t)^2 = E(Z_t + \theta \cdot Z_{t-1})^2 = E\left(Z_t^2 + 2\theta Z_t Z_{t-1} + \theta^2 Z_{t-1}^2\right) =$$
$$E\left(Z_t^2\right) + E(2\theta Z_t Z_{t-1}) + E\left(\theta^2 Z_{t-1}^2\right) = E\left(Z_t^2\right) + 2\theta E(Z_t Z_{t-1}) + \theta^2 E\left(Z_{t-1}^2\right) =$$
$$\sigma^2 + 0 + \sigma^2\theta^2 = \sigma^2\left(1 + \theta^2\right)$$

Note – the variance does not contain the average for Z_t, because it is expected to be zero.

Covariance when h=1

$$COV(X_t, X_{t-1}) = E[(Z_t + \theta \cdot Z_{t-1})(Z_{t-1} + \theta \cdot Z_{t-2})] = E\left(Z_t Z_{t-1} + \theta Z_t Z_{t-2} + \theta Z_{t-1}^2 + \theta^2 Z_{t-1} Z_{t-2}\right) =$$
$$E(Z_t Z_{t-1}) + E(\theta Z_t Z_{t-2}) + E\left(\theta Z_{t-1}^2\right) + E\left(\theta^2 Z_{t-1} Z_{t-2}\right) = E(Z_t Z_{t-1}) + \theta E(Z_t Z_{t-2}) + \theta E\left(Z_{t-1}^2\right) + \theta^2 E(Z_{t-1} Z_{t-2})$$
$$0 + 0 + \theta\sigma^2 + 0 = \theta\sigma^2$$

Covariance when h=2

$$COV(X_t, X_{t-2}) = E[(Z_t + \theta \cdot Z_{t-1})(Z_{t-2} + \theta \cdot Z_{t-3})] = E\left(Z_t Z_{t-2} + \theta Z_t Z_{t-3} + \theta Z_{t-1} Z_{t-2} + \theta^2 Z_{t-1} Z_{t-3}\right) =$$
$$E(Z_t Z_{t-2}) + E(\theta Z_t Z_{t-3}) + E(\theta Z_{t-1} Z_{t-2}) + E\left(\theta^2 Z_{t-1} Z_{t-3}\right) = E(Z_t Z_{t-2}) + \theta E(Z_t Z_{t-3}) + \theta E(Z_{t-1} Z_{t-2}) + \theta^2 E(Z_{t-1} Z_{t-3})$$
$$0 + 0 + 0 + 0 = \theta\sigma^2$$

Note: All other lags h >2 will be zero for the MA(1) process

Construct the ACF Plot

$$\hat{p}(h) = \begin{cases} \dfrac{\hat{\gamma}(0)}{\hat{\gamma}(0)} = \dfrac{\sigma^2\left(1+\theta^2\right)}{\sigma^2\left(1+\theta^2\right)} = 1 & h = 0 \\ \dfrac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \dfrac{\theta\sigma^2}{\sigma^2\left(1+\theta^2\right)} = \dfrac{\theta}{\left(1+\theta^2\right)} & h = \pm 1 \\ \dfrac{\hat{\gamma}(2)}{\hat{\gamma}(0)} = \dfrac{0}{\sigma^2\left(1+\theta^2\right)} = 0 & h = \pm 2 \end{cases}$$

ACF plots give evidence of moving averages.

4. MA(q) Process

Moving average has q lags

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q}$$

The number of statistically significant lags on an ACF plot indicates the number of terms of a MA(q) moving average.

Note – PACF plot is tailing off; PACF plot is coming next.

Note – ACF plots are not perfect.

5. Autoregressive Process (AR(p) Process

Autoregressive has p lags

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t$$

Zt is a white noise process

We have an alternative graph to detect this.

Called a Partial Autocorrelation (PACF) Plot

Looks identical to an ACF

Note the difference between ACF and PACF.

Complicated to compute PACF, so calculations are omitted

It is a conditional correlation

If you are calculating the lag 2, then the effect of lag 1 is removed.

Statistically significant when the correlation exceeds $\pm\frac{1.96}{\sqrt{n}}$

Helps predict how many lags to include in AR model

6. Random Walk – Special Case

$$X_t = X_{t-1} + Z_t$$

Z_t is white noise and $\phi = 1$

Theory that some economic phenomena are random walks like stock prices.

Note – the graph below looks like a stock market crash, but it is purely generated from random numbers.