Unit 4

NONSTATIONARY PROCESSES

It is often the case that while the processes may not have a constant level, they exhibit homogeneous behavior over time. Consider, for example, the linear trend process given in ~~Figure 5.1~~c. It can be seen that different snapshots taken in time do exhibit similar behavior except for the mean level of the process. Similarly, processes may show nonstationarity in the slope as well. We will call a time series, y_t, homogeneous nonstationary if it is not stationary but its first difference, that is, $w_t = y_t - y_{t-1} = (1 - B)\, y_t$, or higher-order differences, $w_t = (1 - B)^d\, y_t$, produce a stationary time series. We will further call y_t an **autoregressive integrated moving average** (ARIMA) process of orders p, d, and q—that is, ARIMA(p, d, q)—if its dth difference, denoted by $w_t = (1 - B)^d\, y_t$, produces a stationary ARMA(p, q) process. The term integrated is used since, for $d = 1$, for example, we can write y_t as the sum (or "integral") of the w_t process as

$y_t = \alpha + \beta_t + a_t$

$\nabla y_t = y_t - y_{t-1}$

$B y_t = y_{t-1}$

NONSTATIONARY PROCESSES

$$\begin{aligned} y_t &= w_t + y_{t-1} \\ &= w_t + w_{t-1} + y_{t-2} \\ &= w_t + w_{t-1} + \cdots + w_1 + y_0 \end{aligned} \tag{5.76}$$

Hence an ARIMA(p, d, q) can be written as

$$\Phi(B)\,(1-B)^d\,y_t = \delta + \Theta(B)\,\epsilon_t \tag{5.77}$$

NONSTATIONARY PROCESSES

Thus once the differencing is performed and a stationary time series $w_t = (1 - B)^d y_t$ is obtained, the methods provided in the previous sections can be used to obtain the full model. In most applications first differencing ($d = 1$) and occasionally second differencing ($d = 2$) would be enough to achieve stationarity. However, sometimes transformations other than differencing are useful in reducing a nonstationary time series to a stationary one. For example, in many economic time series the variability of the observations increases as the average level of the process increases; however, the percentage of change in the observations is relatively independent of level. Therefore taking the logarithm of the original series will be useful in achieving stationarity.

ARIMA models

- **A**uto-**R**egressive **I**ntegrated **M**oving **A**verage
- Are an adaptation of discrete-time filtering methods developed in 1930's-1940's by electrical engineers (Norbert Wiener et al.)
- Statisticians George Box and Gwilym Jenkins developed systematic methods for applying them to business & economic data in the 1970's (hence the name "Box-Jenkins models")

What ARIMA stands for

- A series which needs to be differenced to be made stationary is an "integrated" (**I**) series
- Lags of the stationarized series are called "auto-regressive" (**AR**) terms
- Lags of the forecast errors are called "moving average" (**MA**) terms
- *We've already studied these time series tools separately*: differencing, moving averages, lagged values of the dependent variable in regression

ARIMA models put it all together

- *Generalized random walk models* fine-tuned to eliminate all residual autocorrelation
- *Generalized exponential smoothing models* that can incorporate long-term trends and seasonality
- *Stationarized regression models* that use lags of the dependent variables and/or lags of the forecast errors as regressors
- The most general class of forecasting models for time series that can be *stationarized** by transformations such as differencing, logging, and or deflating

Construction of an ARIMA model

1. *Stationarize* the series, if necessary, by differencing (& perhaps also logging, deflating, etc.)
2. Study the pattern of *autocorrelations* and *partial autocorrelations* to determine if lags of the stationarized series and/or lags of the forecast errors should be included in the forecasting equation
3. Fit the model that is suggested and check its residual diagnostics, particularly the residual ACF and PACF plots, to see if all coefficients are significant and all of the pattern has been explained.
4. Patterns that remain in the ACF and PACF may suggest the need for additional AR or MA terms

ARIMA terminology

- A non-seasonal ARIMA model can be (almost) completely summarized by three numbers:

  p = the number of *autoregressive* terms

  d = the number of *nonseasonal differences*

  q = the number of *moving-average* terms

- This is called an "ARIMA(p,d,q)" model

- The model may also include a *constant* term (or not)

ARIMA terminology