Lecture 1: Some Basic Concepts

1. *Time Series*: A sequence of random variables measuring certain quantity of interest over time.

   Convention:

   - In application, a time series is a record of values of certain quantity of interest taken at different time points.
   - Usually, data are observed at equally spaced time intervals, resulting in a discrete-time time series.
   - If treated as a stochastic process over time (continuous time), then we have a continuous-time time series.
   - Notation: X_t or Y_t or Z_t for a discrete-time time series and $X(t)$, or $Y(t)$ or $Z(t)$ for the continuous-time case.
   - X_t can be a continuous random variable or a discrete random variable, e.g. counts.

2. Basic objective of time series analysis

   The objective of (univariate) time series analysis is to find the dynamic dependence of X_t on its past values $\{X_{t-1}, X_{t-2}, \cdots\}$. Linear model means X_t depends linearly on its past values. To describe the dynamic dependence effectively, it pays to introduce the following operator.

3. Backshift (or lag) operator:

   We define the backshift operator "B" (or the lag operator "L") by

$$BX_t = X_{t-1}.$$

   In other words, BX_t is the value of the time series at time $t-1$.

   We can define a Lag (or backshift) polynomial as

$$\phi(B) = \phi_0 - \phi_1 B - \cdots - \phi_p B^p = \phi_0 - \sum_{i=1}^{p} \phi_i B^i$$

   where $\phi_0 = 1$ and p is a non-negative integer referred to as the "order" of $\phi(B)$.

   Applying this operator to the X_t sequence, we obtain

$$\phi(B)X_t = X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = X_t - \sum_{i=1}^{p} \phi_i X_{t-i}.$$

   This equation is used in time series analysis to describe the dynamic dependence of X_t on its past values.

The equation

$$\phi(B)X_t = c, \tag{1}$$

where c is a constant, is called a "difference equation" of order p. If $c = 0$, the equation is a homogeneous equation. The variable X_t, which satisfies the difference equation in (1), is a solution of that equation. In practice, different $\phi(B)$ can give rise to different dynamic behavior of X_t. We shall use such a difference equation to describe the dynamic pattern of a linear time series.

A variety of dynamic dependence patterns of X_t can be generated by considering the "rational" lag polynomial $\pi(B) = \phi(B)/\theta(B)$.

A simple example of the rational polynomial is

$$\pi(B) = \frac{1}{1-\theta B}.$$

Using long division, we have

$$\frac{1}{1-\theta B} = 1 + \theta B + \theta^2 B^2 + \cdots.$$

Therefore,

$$\frac{1}{1-\theta B} X_t = \sum_{i=0}^{\infty} \theta^i X_{t-i}.$$

If the $\{X_t\}$ sequence is bounded, then we might want the resulting sequence to be bounded as well. This is achieved by requiring $|\theta| < 1$. This special rational polynomial shows that X_t is an infinite-order moving average of its past values, $\{X_{t-1}, X_{t-2}, \cdots\}$, with weights decaying exponentially.

If $|\theta| > 1$, then it might be reasonable to define

$$\frac{1}{1-\theta B} = \frac{-(\theta B)^{-1}}{1-(\theta B)^{-1}} = \frac{-1}{\theta B}[1 + \frac{1}{\theta}B^{-1} + \frac{1}{\theta^2}B^{-2} + \cdots]$$

$$= -\sum_{i=1}^{\infty} (1/\theta)^i X_{t+i}.$$

This is a forward-looking moving average that relates X_t to its future values. Note that sometimes we write $B^{-1} = F$ such that $FX_t = X_{t+1}$ and refer to F as the forward operator.

4. First-order Difference Equations

A difference equation is a deterministic relationship between the current value X_t and its past values X_{t-i} with $i > 0$. In some cases, it may also contain the current and past values of a "forcing" or "driving" variable. A first order difference equation involves only one lagged variable:

$$X_t = \phi X_{t-1} + b a_t + c,$$

where a_t is a forcing variable, which follows a well-defined probability distribution. Using the backshift operator, we can write the model as

$$(1 - \phi B)X_t = c + ba_t.$$

The "solution" to this difference equation expresses the current value X_t as a function of time and the forcing variable a_t.

$$X_t = \frac{c}{1 - \phi B} + \frac{b}{1 - \phi B} a_t + \gamma \phi^t.$$

The term $\gamma\phi^t$ is included because this is the only function such that $(1 - \phi B)f_t = 0$. Note that $\gamma\phi^t$ is the solution to the homogeneous equation $(1 - \phi B)X_t = 0$. As in the theory of differential equations, the solution of a difference equation consists of the solution to the homogeneous part plus a particular solution to the inhomogeneous equation.

The solution, $X_t = c/(1 - \phi) + b \sum_{i=0}^{\infty} \phi^i a_{t-i} + \gamma\phi^t$, is fully determined by knowledge of the initial conditions of the sequence. If X_0 is known, then

$$X_t = \frac{c(1 - \phi^t)}{1 - \phi} + b \sum_{i=0}^{t-1} \phi^i a_{t-i} + \phi^t X_0.$$

If $|\phi| < 1$ and a_t is bounded with mean zero, then $E(X_t)$ approaches $c/(1 - \phi)$ from any starting point.

Obviously, the value of ϕ determines the qualitative behavior of the equation $X_t = \phi^t X_0$: There are three types of solution: (a) smooth damped, (b) Oscillatory damped, (c) Explosive.

5. Second-order Difference Equations:

   The same idea can be extended to higher order difference equations. Solutions to higher order difference equations can exhibit more interesting sinusoidal patterns.

   Consider the general second-order difference equation:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \delta + ba_t$$

   or

$$(1 - \phi_1 B - \phi_2 B^2)X_t = \delta + ba_t.$$

   Solutions of this equation can be computed by factoring the backshift polynomial as, $(1 - \phi_1 B - \phi_2 B^2) = (1 - \lambda_1 B)(1 - \lambda_2 B)$. Considering just the homogeneous equation, we find a solution of the form

$$X_t = c_1(\lambda_1)^t + c_2(\lambda_2)^t \quad \text{(why?)}$$

   Note that $1/\lambda_1, 1/\lambda_2$ are the zeros of the polynomial $1 - \phi_1 B - \phi_2 B^2$. If the homogeneous solution is to remain bounded, we would require $|\lambda_i| < 1$ for $i = 1$, 2, or equivalently that the zeros of the polynomial $1 - \phi_1 B - \phi_2 B^2$ lie outside the unit circle (modulus > 1). [Note: Zeros of $1 - \phi_1 B - \phi_2 B^2$ are roots of the equation $1 - \phi_1 B - \phi_2 B^2 = 0$.]

For a second-order equation, we have three possibilities: (a) distinct real roots, (b) equal real roots, and (c) complex roots. The quadratic formula gives the roots as

$$\lambda_i = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2}.$$

We have complex roots if $\phi_1^2 + 4\phi_2 < 0$. The roots are $a \pm bi$ where i is $\sqrt{-1}$. Note that the complex roots come in a conjugate pair.

If we write the roots using polar form, we can see how oscillatory solutions are possible.

$$a \pm bi = r(\cos\theta \pm \sin\theta)$$

where $r = \sqrt{a^2+b^2}$ and $\cos\theta = a/r = \phi_1/(2\sqrt{-\phi_2})$, or $\theta = \cos^{-1}(\phi_1/(2\sqrt{-\phi_2})$.

Using DeMoive's formula, namely $\cos\theta + i\sin\theta = e^{i\theta}$, we can write

$$\begin{aligned} X_t &= c_1(re^{i\theta})^t + c_2(re^{-i\theta})^t \\ &= r^t(c_1e^{it\theta} + c_2e^{-it\theta}) \\ &= r^t[(c_1+c_2)\cos(t\theta) + i(c_1-c_2)\sin(t\theta)]. \end{aligned}$$

Since X_t is real, c_1+c_2 must be real while c_1-c_2 must be imaginary. Thus, c_1 and c_2 are complex conjugates. Using DeMoive's formula again and some identifies from trigonometrics, we obtain

$$X_t = kr^t\cos(t\theta + \omega).$$

Equal roots case:

$$(1-\lambda B)^2 X_t = 0$$

The solution of which is

$$X_t = c_1\lambda^t + c_2 t\lambda^t.$$

6. General Case:

   The above results can readily be extended to the general higher order difference equations.

7. Stochastic Difference Equations: When the "forcing" factor is stochastic, we have a general difference equation. In particular, the case in which the forcing variable is a sequence of independent and identically distributed normal random variable $\{a_t\}$ plays an important role in time series analysis. Here the solution X_t is usually correlated and follows certain statistical distribution.

8. Let a_t and X_t be input and output at time t, respectively. Consider the linear system

   $$X_t = \psi_0 a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots = \psi(B)a_t,$$

   where $\psi(B) = \psi_0 + \psi_1 B + \psi_2 B^2 + \cdots$ and $\psi_0 = 1$. Consider the relationship between ψ_i and coefficients $\pi(B)$ discussed earlier.