Chapter 9 Stochastic trends and seasonality

1. In previous chapters, no special attention was paid to trends and seasonality. But, the modeling procedures work well to obtain adequate models for various data. However, we can see that

   $\lambda \to 1$, when there is trend. e.g., p171 IBM data

   Key to trend/seasonality analysis:

   μ and λ close to their stability boundary.

   real roots --> trend

   complex roots --> seasonality

2. Stochastic trends (constant, growth or decay trends in the Green's Function)

(1) Constant trend : A tendency in the data to remain at the "same" level, arising from a first order system.

$$G_j = \phi^j = \lambda^j \rightarrow G_j \equiv 1 \quad \text{as } \phi = \lambda \rightarrow 1$$

$$G(t) = e^{-\alpha_0 t} = e^{\mu_1 t} \rightarrow G(t) \equiv 1 \quad \text{as } \alpha_0 = \mu_1 \rightarrow 0$$

These two limiting cases give a random walk.

$$(1-B)X_t = a_t \quad \text{or} \quad \nabla X_t = a_t$$

From Fig. 9.1, although the series tends to remain at the same level, this level changes randomly due to the <u>stochastic</u> nature of the trend.

(2) Linear trend: Linear growth or decay trends may be present in the data when two roots are close to one for the discrete or close to zero for the continuous case.

$$X_t = \frac{(1-\theta_1 B)}{(1-\lambda B)(1-\lambda B)} a_t = [1+\lambda B+(\lambda B)^2+\ldots][1+\lambda B+(\lambda B)^2+\ldots](1-\theta_1 B)\, a_t$$

$$= [1+2(\lambda B)+3(\lambda B)^2+4(\lambda B)^3+\ldots](1-\theta_1 B)\, a_t$$

$$= \sum_{j=0}^{\infty} (j+1)\lambda^j a_{t-j} - \sum_{j=0}^{\infty} \theta_1 j\lambda^{j-1} a_{t-j} = \sum_{j=0}^{\infty} [(j+1)\lambda^j - \theta_1 j\lambda^{j-1}]\, a_{t-j}$$

$$G_j = [1+j(1-\frac{\theta_1}{\lambda})]\,\lambda^j$$

For a continuous 2nd order system with repeated roots:

$G(t) = (C_1 + C_2 t)\, e^{\mu t}$

$G_j \to 1 + (1-\theta_1)\, j \quad \text{as } \lambda_1, \lambda_2 \to 1$

$G(t) \to C_1 + C_2 t \quad \text{as } \mu_1, \mu_2 \to 0$

If a parsimonious model is desired,

$(1-B)^2 X_t = (1-\theta_1 B) a_t \quad \text{or} \quad \nabla^2 X_t = (1-\theta_1 B) a_t$

From Fig. 9.2, although the series tends to behave linearly in straight line, the slope is slowly changing due to the stochastic nature of the trend.

$(1-B) = \nabla$ not asymptotically stable, but stable $G_j = g(1)^j$

$(1-B)^2 = \nabla^2$ not stable, $G_j = 1 + (1-\theta)j \to \infty$ as $j \to \infty$

The difference in stability can be seen in Fig. 9.1 and 9.2 by the ranges of the two plots.

Advantage of using a continuous model:

If a continuous model is used, it can provide valuable information whether the trend arises as a result of genuine limiting system parameters, such as α, or simply because the sampling interval is tool small.

e.g.,

$$G_j = g_1 e^{-(\zeta + \sqrt{\zeta^2 - 1})\, \omega_n \Delta j} + g_2 e^{-(\zeta - \sqrt{\zeta^2 - 1})\, \omega_n \Delta j}$$

If ω_n is very small, both roots tends to zero --> indicate a linear trend

i) when roots are complex, sine wave with very small frequency

ii) when roots are real, exponential with a very slow decaying factors

(3) Polynomial trend:

nth order polynomial trend can be represented by a model with n+1 discrete real roots close to one.

$G_j = 1 + g_1 j + g_2 j^2 + \ldots + g_n j^n \quad \text{or} \quad G(t) = C_1 + C_2 t + \ldots + C_{n+1} t^n$

e.g.,

$(1-B)^3 X_t = (1-\theta_1 B - \theta_2 B^2) a_t$

shown in Fig. 9.3

Remarks: The range is a clear indication of increased instability of the system.

2. Stochastic seasonality (or periodic trends)

   Seasonal or periodic trends may arise when the roots are complex and of absolute value one in the discrete case or the real parts tend to zero in the continuous case.

When $\phi_2 = -\lambda_1\lambda_2 = -1$ and $\phi_1^2 + 4\phi_2 < 0$

$$\lambda_1, \lambda_2 = \frac{1}{2}\left(\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}\right) = r\,e^{i\omega}$$

$$r = \sqrt{\left[\frac{\phi_1}{2}\right]^2 + \left[\frac{\sqrt{-(\phi_1^2 + 4\phi_2)}}{2}\right]^2} = \sqrt{-\phi_2}$$

$$\cos\omega = \frac{\phi_1}{2\sqrt{-\phi_2}}$$

(1) Periodicity of 12 (yearly)

since

$$\cos\left(\frac{2\pi}{12}\right) = \frac{\sqrt{3}}{2}$$

then

$$\phi_1 = \frac{\sqrt{3}}{2} \times 2\sqrt{-\phi_2} = \sqrt{3}$$

thus

$$\lambda_1, \lambda_2 = \frac{1}{2}\left(\sqrt{3} \pm i\right)$$

ARMA(2,1) models with

$$(1 - \sqrt{3}\,B + B^2)X_t = (1 - \theta_1 B)a_t$$

e.g., Sunspot data

(2) Periodicity of 3 (quarterly)

$$\phi_2 = -1 \qquad \phi_1 = 2\cos\left(\frac{2\pi}{3}\right) = -1 \qquad \lambda_1, \lambda_2 = \frac{1}{2}\left(-1 \pm i\sqrt{3}\right)$$

$$(1 + B + B^2)X_t = (1 - \theta_1 B)a_t$$

(3) Arbitrary period

$$\phi_1 = 2\cos\left(\frac{2\pi}{p}\right)$$

Differences between ω_n and seasonality (deterministic and stochastic):

ω_n is a system's natural frequency regardless whether there is a visible oscillation in response or not. For instance, for an overdamped system, the system response to a step input will not oscillate but it still has a natural frequency.

That is, for any given pair of system roots, one can always calculate a corresponding natural frequency for the continuous system.

Deterministic seasonality is due to the superposition of strict sine cosine waves to correlated signals. There is a predictable cyclic change in the signal.

Stochastic seasonality refers to the phenomena that sometimes the data fluctuate around a fixed mean and the periodic tendency is present but not explicitly clear. An overall period or periods are discernible (for example, by counting the average number of peaks), but the periodicity appears to be more or less changing from one part of the data to another.

The lecture ended with a presentation of the term project on:

- Birth rates in Singapore (interesting topic, but quite a few errors in the project)