Chapter 10 Deterministic trends and seasonality: Nonstationary series

1. Wide sense stationarity — The first two moments of data, namely the mean and the covariance, are independent of the time origin.

   Nonstationary trends and seasonal patterns can be modeled by decomposing the data into: deterministic (to representing the mean of the series)
   and
   stochastic parts (w/ zero mean)

   i.e., removing the deterministic part causing the nonstationarity.

   In other words, the nonstationary series can be modeled by relaxing the first assumption of zero or fixed mean.

   Many of the nonstationary data can be modeled by explicitly including polynomial, exponential, or sinusoidal functions, dependent on the time origin, to represent the mean of the series.

2. Linear trends

   (a) In a preliminary fit, linear trend will be obtained by the least squares method assuming uncorrelated errors.

   (b) The residuals will be examined together with their sample autocorrelations to check that no other nonstationary trends are present.

   (c) The residuals will be then modeled by ARMA.

   (d) The parameters of the final deterministic plus stochastic model can be estimated together by starting with the estimates from the separate models as initial values. This is especially necessary when large proportion of stochastic parts exist in the data.

Example: Crack propagation data

page. 357, Fig. 10.1,

(1) <u>deterministic trend</u>

$y_t=\beta_0+\beta_1 t+\varepsilon_t$ $\hat{\beta}_0=11.03\pm0.05$ $\hat{\beta}_1=0.107\pm0.0099$ RSS=1.326

After removing the deterministic trend, the residuals are plotted in Fig. 10.2.

(2) <u>stochastic part:</u> magnitude is relatively small, but has a strong dependence

• autocorrelation of residuals from deterministic model is plotted in Fig. 10.3

AR(1) model seems to be adequate: $\phi_1=0.9299\pm0.079$ RSS = 0.18986

• autocorrelation of residuals from AR(1) model is shown in Fig. 10.4. (much reduced)

(3) <u>complete model:</u>

$$Y_t = \beta_0 + \beta_1 t + X_t \quad \text{and } X_t = \phi_1 X_{t-1} + a_t$$

If the data are uniformly sampled from a continuous system, the equivalent continuous model will be:

$$Y(t) = \beta_0 + \beta_1 t + X(t) \quad \text{and } \frac{dX(t)}{dt} + \alpha_0 X(t) = Z(t)$$

$$\alpha_0 = -\ln(0.9299) = 0.07268$$

For the purpose of estimating the final model parameters, we re-write:

$$Y_t = \beta_0 + \beta_1 t + \phi_1 X_{t-1} + a_t \qquad X_t = Y_t - \beta_0 - \beta_1 t$$

or

$$Y_t = \beta_0 + \beta_1 t + \phi_1 [Y_{t-1} - \beta_0 - \beta_1 (t-1)] + a_t$$

This involves the nonlinear least squares routine to minimize the sum of squares of a's.

$\hat{\beta}_0 = 10.8896 \pm 0.085$ $\hat{\beta}_1 = 0.1079 \pm 0.0024$ $\hat{\phi}_1 = 0.9592 \pm 0.0679$ RSS=0.1694

For this particular case, the final estimations are not much different from the parameters of the separate models due to the relatively small portion of noise.

(4) physical interpretation:

A conjecture about the stochastic part of the model based on the fact that the stochastic part represents a very small percentage of the actual observations.

Assume: the stochastic part is contributed by the measurement instrument.

The instrument cannot respond immediately and will involve some time lag or inertia giving rise to instrument error or noise.

$$\tau = \frac{1}{\alpha_0} = 24.01$$

Observations: $Y(t) = C(t) + X(t)$

Crack length: $\frac{dC(t)}{dt} = \beta_1 = 0.108$

$C(0) = \beta_0 = 10.89$

Instrument: $\frac{dX(t)}{dt} + 0.04166 X(t) = Z(t)$ $\sigma_z^2 = 0.00195$

(5) dynamic calibration of measuring instruments

3. Exponential trends

Ex trends generally arise from physical/engineering systems governed by differential equations with constant coefficients. (unknown order, coefficients, noise)

For real roots, the system response to a step or impulse input will be a sum of exponentials with added noise.

procedures:

(a) basic strategy: fitting models of successively higher order until no further improvement can be obtained.

(b) apply the strategy to both deterministic and stochastic models, i.e., assuming n exponentials for the deterministic part, plus ARMA(n,n-1) for the stochastic part.

Example: Papermaking process — basis weight response to step input

page 414, Fig. 11.1 diagram of papermaking process

basis weight : the weight of dry paper per unit area.

Page 365, Fig. 10.6, basis weight response to a step input

(1) first order dynamics

$$y_t = A_0 + g(1 - e^{-\frac{t}{\tau}}) + \varepsilon_t$$

All parameters can be initially estimated from the raw data.

initial value --> A_0

after 30 seconds, data fluctuate around 40.6 --> $A_0 + g$

at $t=\tau$, $A_0+(1-.37)g = 40$ $\tau=27$ sec.

(2) The parameters A_0, g, τ are estimated by nonlinear least square method.

$\hat{A}_0=38.76\pm0.184$ $\hat{g}=1.9321\pm0.1661$ $\hat{\tau}=27.17\pm6.259$ RSS=4.029

(3) Residuals and autocorrelations are shown in Figs. 10.9 /10.10

AR(1) model is fitted.

(4) final model:

$$y_t = A_0 + g(1 - e^{-\frac{t}{\tau}}) + X_t \qquad X_t = \phi_1 X_{t-1} + a_t$$

(5) To estimate the parameters of the final model, get initial values

$\hat{\phi}_1 = \hat{\rho}_1 = 0.755$

then use nonlinear least square methods

$\hat{A}_0= 38.9974\pm0.2682$ $\hat{g}=1.7992\pm0.4228$ $\hat{\tau}= 35.3912\pm23.56$ $\hat{\phi}_1=0.7887\pm0.127$

(6) Large difference is due to the large proportion of noise in the data leading to significant errors in estimates when the residuals are incorrectly assumed to be uncorrelated in the deterministic modeling.

(7) Residuals from AR(1) model in Fig. 10.11.

$$\tau_x = \frac{1}{\alpha_0} = \frac{1}{-\ln(0.7887)} = 4.2128$$

4. Periodic trends: seasonality

For some or all complex conjugate roots, the system response may be a sum of exponentials, periodic waves, with added noise.

case a) Strict sine cosine waves or their combinations superimposed with correlated noise form a special case of these models.
case b) The data fluctuate around a fixed mean but the periodicity appears to be more or less changing from one part of the data to another. These stochastic tendency should be modeled by the methods in Ch.9.
case c) The periodic tendency repeats regularly plus a deterministic trend.

Procedures:

(a) The growth (or decay) trends can be removed in the preliminary analysis by the models containing exponentials with real roots.

(b) The periodic trends can then be added to the above model by pairs of complex conjugate exponentials, with the imaginary parts of the roots containing the dominant frequency and its multiples.

(c) After the addition of further periods and trends fail to result in a significant improvement in the RSS, the residuals can be checked for stationarity and ARMA model fitted to the residuals.

(d) The parameters of the combined model will be re-estimated

Example: International airline passenger data

Page 377, Fig. 10.15

The data contains: an overall growth tendency, growing periodic trends.
peaks in the month of March, July, and August.

$$y_t = \sum_{j=1}^{l} R_j e^{r_j t} + \sum_{j=1}^{k} B_j e^{b_j t} \sin(j\omega t + \psi_j) + X_t$$

$$s = l + 2k$$

$$y_t = \sum_{j=1}^{l} R_j e^{r_j t} + \sum_{j=1}^{k} B_j e^{b_j t} [C_j \sin(j\omega t) + \sqrt{1 - C_j^2} \cos(j\omega t)] + X_t$$

$$C_j = \cos\psi_j$$

(a) exponential growth trend

$$y_t = R_1 e^{r_1 t} + \varepsilon_t$$

$$t=0,\ y_0 \approx R_1 = 130$$

$$t = \frac{1}{r_1},\ \ y_t = R_1 e^1 = 130(2.7) = 351 \rightarrow t \approx 100 \text{ months}$$

$r_1=0.01$

Use nonlinear least square methods with above R and r as initial values.

Table 10.2.

Residuals from a single exponential model is shown in Fig. 10.16, which indicates that no more exponentials are necessary.

(b) addition of periodic trends

$$y_t= R_1e^{r_1 t} + B_1e^{b_1 t}[C_1\sin(\omega t)+\sqrt{1-C_1^2}\cos(\omega t)] +\varepsilon_t \qquad \omega=\frac{2\pi}{12}$$

From Fig. 10.16, we can roughly estimate the initial values for:

$\hat{B}_1=-20$, $\hat{b}_1=0.01$, $\hat{C}_1=0.7$

After nonlinear least squares estimates, the results are shown in Table 10.2.

RSS = 296,250 to 95,783

Then, successively try k=2,3,..

$$y_t= R_1e^{r_1 t} + \sum_{j=1}^{k} B_je^{b_j t}[C_j\sin(j\omega t) + \sqrt{1-C_j^2}\cos(j\omega t)] + \varepsilon_t$$

until the reduction in RSS is insignificant. k=5

These five periods account for yearly, 4 monthly, quarterly, and 2-2/5 monthly periods.

Residuals are shown in Fig. 10.17.

(c) ARMA(3,3) model.

5. General nonstationary models

$$y_t=\sum_{j=0}^{s} A_je^{K_j t} +X_t \qquad (1)$$

where X is the response of an ARMA(n,n-1) system for discrete time

$$(1-\phi_1B -\phi_2B^2-\ldots-\phi_nB^n)X_t=(1-\theta_1B-\theta_2B^2-\ldots-\theta_{n-1}B^{n-1})a_t$$

or of an AM(n,n-1) system for continuous time.

$$(D^n+\alpha_{n-1}D^{n-1}+\ldots+\alpha_0)X(t) = (1+b_1D+\ldots+b_{n-1}D^{n-1})Z(t)$$

In model (1),

- when all A_j's are zero, --> it becomes a zero mean stationary series.
- When only one A_j is nonzero and the corresponding k_j is zero, --> constant mean stationary series.
- when the exponents k_j are real, very small but nonzero, --> polynomial trends

- with real, large and positive or negative exponents, --> increasing or decreasing exponential trends
- with complex conjugate exponents k_j

    negative real parts, --> damped sine cosine trends.

    zero real parts, --> exact sine cosines with noise.

    positive real parts --> sine cosine with growing amplitude.

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 <u>stochastic</u> 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.

(3) Polynomial trend:

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

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

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

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

then

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

thus

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

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)

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

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

(3) Arbitrary period

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

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.