Financial Time Series

Topic 3: Non-stationary Processes and ARIMA models

Hung Chen
Department of Mathematics
National Taiwan University
3/20/2000

OUTLINE

1. Non-stationary Processes
   - Non-stationarity in variance
   - Box-Cox transformation
   - Non-stationarity in mean
   - Eliminate the trend term by differencing
2. ARIMA Models
   - Description
   - Properties
   - Data Modelling
   - Forecasting

Nonstationary Time Series Models

- Weakly stationary implies that the mean, variance and autocovariances of the process are invariant under time translation
- Figure 2.14 plots monthly observations from January 1965 December 1990 of the FTA (Financial times-Actuaries) All Share index.
  - It shows that the series to exhibit a prominent upward, but not linear, trend, with pronounced and persistent fluctuations about it, which increase in variability as the level of the series increases.
  - non-stationarity in variance:
    Write a time series as the sum of a **non-stochastic** mean level and a random error component:

$$X_t = \mu_t + \epsilon_t, \qquad (1)$$

    and we suppose that the variance of the errors is functionally related to the mean level μ_t by

$$V(X_t) = V(\epsilon_t) = h^2(\mu_t)\sigma^2,$$

where h is some known function.

- Box and Cox (1964) class of power transformations

$$g(x_t) = \frac{x_t^\lambda - 1}{\lambda}$$

and

$$\begin{aligned} &\lim_{\lambda \to 0} \frac{x_t^\lambda - 1}{\lambda} \\ &= \lim_{\lambda \to 0} \frac{\exp(\lambda \ln x_t) - 1}{\lambda} = \ln x_t. \end{aligned}$$

Idea: Consider

$$g(x_t) \approx g(\mu_t) + (x_t - \mu_t)g'(\mu_t)$$

and

$$V(g(x_t)) \approx [g'(\mu_t)]^2 h^2(\mu_t)\sigma^2.$$

Choose $g'(\mu_t) = 1/h(\mu_t)$ to stabilize the variance.

- Apply a logarithmic transformation to FTA all Share index.

– Figure 2.14 indicates that the transformation linearize the trend and stabilize the variance.

– When $h(\mu_t) = \mu_t$, $g(\mu_t) = \ln \mu_t$.
The natural logarithms of x_t can be used to stabilize the variance.

– The use of logarithms is a popular (why?) transformation for financial time series, a constant variance is rarely completely induced by this transformation alone. More to be seen in Chapters 4 and 7.

- non-stationarity in mean:
How do we model the non-constant mean level in (1)?

  – Figure 2.14 indicates that the transformation linearize the trend and stabilize the variance.

  – Assume that the mean evolves as a polynomial of order d in time.

  – $\{x_t\}$ is decomposed into a trend component, given by the polynomial, and a stochastic, stationary, but possibly autocorrelated, zero mean error component.

$$x_t = \sum_{j=0}^{d} \beta_j t^j + \phi(B) a_t. \qquad (2)$$

Note that $E(\epsilon_t) = \phi(B)E(a_t) = 0$ and hence

$$E(X_t) = E(\mu_t) = \sum_{j=0}^{d} \beta_j t^j.$$

– Consider the linear trend $(d = 1)$

$$x_t = \beta_0 + \beta_1 t + a_t. \qquad (3)$$

Lagging (3) one period and subtracting this from (3) yields

$$x_t - x_{t-1} = \beta_1 + (a_t - a_{t-1}). \qquad (4)$$

Let $w_t = x_t - x_{t-1} = (1-B)x_t = \triangle x_t$. Then

$$w_t = \triangle x_t = \beta_1 + \triangle a_t,$$

which is stationary $(E(W_t) = \beta_1)$ but not invertible $MA(1)$ process.

– Differencing:

$$\begin{aligned} \triangle &= 1 - B && \text{the first difference operator} \\ \triangle^d &= (1-B)^d && \text{the } d\text{th difference operator} \end{aligned}$$

- $\triangle^d \Sigma_{j=0}^d \beta_j t^j = d!\beta_d$.
  Any polynomial trend of degree d can be reduced to a constant by application of the operator.
- Suggestion: Given any sequence $\{x_t\}$ of data, apply the operator $\triangle$ repeatedly until we find a sequence $\{\triangle^d x_t\}$ which can plausibly be modelled as a realization of a stationary process.
- A series $\{x_t\}$ is nonstationary but its dth differenced series $\{(1-B)^d x_t\}$ for some integer $d \geq 1$, is stationary.
- Typically, $d = 1$ or 2.
  Note that $\triangle^2 x_t = x_t - 2x_{t-1} + x_{t-2}$.
- Example

$$\begin{aligned} x_t &: \{26.8, 34.7, 25.4, \ldots, 38.1, 39.5\} \\ (1-B)x_t &: \{7.9, -9.3, \ldots, 1.4\} \end{aligned}$$

ARIMA Models

- A series may need first differencing d times to attain stationarity and the obtained series may itself be autocorrelated.
- Suppose this autocorrelation can be modeled by an $ARMA(p, q)$ process.
- The model for the original series is of the form

$$\phi(B) \bigtriangleup^d x_t = \theta_0 + \theta(B) a_t, \qquad (5)$$

  where $\theta_0 = d! \beta_d$. It is said to be an autoregressive-integrated-moving average process of orders p, d and q, or $ARIMA(p, d, q)$.
- X_t is said to be integrated of order d, denoted $I(d)$.

Forecasting using ARIMA models:

Given a realization $\{x_t\}_{1-d}^{T}$ from a general $ARIMA(p, d, q)$ process

$$\phi(B) \bigtriangleup^d x_t = \theta_0 + \theta(B) a_t.$$

How do we forecast a future value X_{T+h}?

- Let

$$\begin{aligned} \alpha(B) &= \phi(B) \bigtriangleup^d \\ & \quad \left(1 - \alpha_1 B - \alpha_2 B^2 - \cdots - \alpha_{p+d} B^{p+d}\right). \end{aligned}$$

- Denote a minimum mean square error (MMSE) forecast based on the data up to time T by $f_{T,h}$. Then

$$\begin{aligned} f_{T,h} &= E(\alpha_1 X_{T+h-1} + \alpha_2 X_{T+h-2} + \cdots \\ & \quad + \alpha_{p+d} X_{T+h-p-d} + \theta_0 + a_{T+h} - \theta_1 a_{T+h-1} - \cdots \\ & \quad - \theta_q a_{T+h-q} | x_T, x_{T-1}, \cdots). \end{aligned}$$

- Note that

$$E(X_{T+j} | x_T, x_{T-1}, \cdots) = \begin{cases} x_{T+j}, & j \le 0 \\ f_{T,j}, & j > 0 \end{cases}$$

and

$$E(a_{T+j} | x_T, x_{T-1}, \cdots) = \begin{cases} a_{T+j}, & j \le 0 \\ 0, & j > 0 \end{cases}$$

- Algorithm:
  - Replace past expectations ($j \leq 0$) by known values, x_{T+j} and a_{T+j}.
  - Replace future expectations ($J > 0$) by forecast values, $f_{T,j}$ and 0.

Examples:

1st Example:

- $AR(2)$ model

  $$(1-\phi_1 B-\phi_2 B^2)x_t=\theta_0+a_t$$

  Hence, $\alpha(B)=1-\phi_1 B-\phi_2 B^2$.

- Note that

  $$x_{T+h}=\phi_1 x_{T+h-1}+\phi_2 x_{T+h-2}+\theta_0+a_{T+h}$$

  and

  $$f_{T,h}=(\phi_1+\phi_2)f_{T,h-1}-\phi_2(f_{T,h-1}-f_{T,h-2})+\theta_0.$$

- By repeated substitution, we have

  $$\begin{aligned} f_{T,h} &= \theta_0\sum_{j=0}^{h-1}(\phi_1+\phi_2)^j+(\phi_1+\phi_2)^h x_T \\ &\quad -\phi_2\sum_{j=0}^{h-1}(\phi_1+\phi_2)^j(f_{T,h-1-j}-f_{T,h-2-j}) \end{aligned}$$

  where $f_{T,0}=x_T$ and $f_{T-1}=x_{T-1}$.

- As $h\to\infty$,

  $$f_{T,h}=\frac{\theta_0}{1-\phi_1-\phi_2}=E(X_t)=\mu$$

  since $\phi_1+\phi_2<1$ and $|\phi_2|<1$.

- The best forecast of a future observation with large lead time is eventually the mean of the process.

2nd Example:

- $ARIMA(0,1,1)$ model
$$\nabla x_t = (1 - \theta B)a_t.$$
Hence, $\alpha(B) = 1 - B$.

- Note that
$$\begin{aligned} x_{T+h} &= x_{T+h-1} + a_{T+h} - \theta a_{T+h-1}, \\ f_{T,1} &= x_T - \theta a_T \end{aligned}$$
and, for $h > 1$,
$$f_{T,h} = f_{T,h-1}.$$

- Note that
$$a_T = (1 - B)(1 - \theta B)^{-1} x_T$$
and
$$\begin{aligned} f_{T,h} &= (1-\theta)(1-\theta B)^{-1} x_T \\ &= (1-\theta)(x_T + \theta x_{T-1} + \theta^2 x_{T-2} + \cdots). \end{aligned}$$

- The forecast for all future values of x is an exponentially weighted moving average of current and past values.