Time Series Analysis - Forecasting and control

Presentation

What is a time series..?

- A ***time series*** is a set of observations generated sequentially in time.
    - Ex: weather data over days/months/years, daily closing stock prices,etc.
- $Z_1, Z_2, Z_3, \ldots Z_N$ are the observations made at t_0+h, t_0+2h, $t_0+3h, \ldots, t_0+Nh$, times being equispaced by an interval h, taking t_0 as the origin time.

Continuous and discrete time series

- If the set is continuously observed or recorded at every instant of time, the time series is said to be ***continuous***;
  - Ex: sine wave, continuous i/p of gas
- ***Discrete*** time series arise in two ways:
  - *Sampling*: By sampling a continuous time series
    - Ex: continuous i/p of gas sampled at every 10 seconds
  - *Accumulating*: By accumulating a variable over a period of time
    - Ex: rainfall measure accumulated over a day/month/year

Time Series Analysis

- The procedure of fitting a time series to a proper model in called ***Time Series Analysis***.
- The model may be ***deterministic*** or ***stochastic***.
- A time series analysis seeks to draw inferences from the data.
- It comprises of methods that attempt to understand the nature of the series and is often useful for future forecasting and simulation.

3 important areas of application for time series:

- Forecasting future values of a series from current and past values (our subject of interest).
- Determination of the transfer function of system (input-output)
- Design of simple feed-forward and feedback control schemes by means of which potential deviations of the output from a desired target may be compensated.

1. Forecasting Time Series

- The use of available observations from a time series to forecast its value at some future time $t+l$ can provide basis for
  - Economic and business planning
  - Production planning
  - Inventory and production control, etc.
- Forecasts are usually needed over a period known as lead time l which varies with each problem.
  - Lead time: period between start and end of process
- Ex: compute how much inventory is needed for future demand using previous data.

Forecasting Time Series (contd...)

- Consider sales forecasting problem, the current month t & the sales z_{t-1} ,z_{t-2} , ... in previous months might be used to forecast sales for lead times $l=1,2,...,12$ months ahead.
- Forecasts made at origin 't' is denoted by $z'_t(l)$ of the sales z_{t+l} at some future time $t+l$.
- The function $z'_t(l)$ that provides the forecasts at origin t for all future lead times will be called the *forecast function at origin t.*
- Objective: obtain a forecast function which is such that the mean square of the deviations z_{t+l} - $z'_t(l)$ between the actual and forecast value.

Deterministic and stochastic models

- If exact calculation were possible, such a model would be entirely ***deterministic***.
    - Ex: trajectory of a missile launched in a known direction with known velocity
- If we derive a model that can be used to calculate the probability of a future value lying between specified limits, (or) a model that includes the probabilistic structure and gives probabilistic outputs, it is said to be a ***probability model*** or a ***stochastic model***.
    - Ex: sales forecasting, weather forecasting, etc.

- A time series $Z_1, Z_2, Z_3, \ldots Z_N$ of N successive observations is regarded as a sample realization from an infinite population of such time series that could have been generated by the *stochastic process;* also called as a *process.*

Stationary and Non-Stationary stochastic models

- ***Stationary models*** is an important class of stochastic models, which assume that the process remains in equilibrium about a constant mean level.
- Nevertheless, many practical use-cases such as those of industry, business, economics, etc are represented as ***non-stationary***, having no natural mean.

Some Simple Operators

- Backward Shift Operator (B) :
    - $Bz_t = z_{t-1}$

      => B^k means "backshift k times".

      Examples

      $$B^k z_t = z_{t-k}$$

- Forward Shift Operator (F) :
    - $Fz_t = z_{t+1}$ => $F^k z_t = z_{t+k}$
    - $F = B^{-1}$

Some Simple Operators (contd...)

- Backward difference operator:

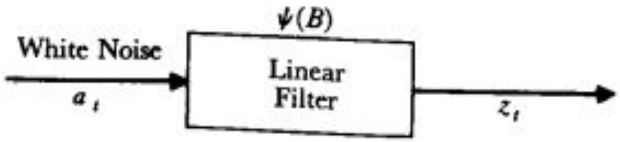
$$\nabla z_t = z_t - z_{t-1} = (1 - B)z_t$$

- Backward difference operator for its inverse the summation operation S:

$$\begin{aligned}\nabla^{-1} z_t = S z_t &= \sum_{j=0}^{\infty} z_{t-j} \\ &= z_t + z_{t-1} + z_{t-2} + \cdots \\ &= (1 + B + B^2 + \cdots)z_t \\ &= (1 - B)^{-1} z_t\end{aligned}$$

Linear Filter Model

- The stochastic model we employ are based on the idea that, a time series in which successive values are highly dependent and as generated from a series of “independent shocks”, a_t.
- These shocks are random drawaings from a fixed distribution, usually assumed Normal.
- Such a sequence of random variables a_t, a_{t-1}, a_{t-2},... is called a white noise process.
- The white noise process a_t, is transformed into the process z_t , by what is called a linear filter.



Representation of a time series as the output from a linear filter

Linear Filter Model (contd...)

- The linear filter model simply takes the weighted sum of previous observations, so that,

$$\begin{aligned} z_t &= \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots \\ &= \mu + \psi(B) a_t \end{aligned}$$

Where Mu is the level of process and,

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \cdots$$

AutoRegressive Model

- The current value of the process is expressed as a finite, linear aggregate of previous values of the process and a shock a_t.

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \cdots + \phi_p \tilde{z}_{t-p} + a_t$$

- represents the autoregressive process of order p.

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$$

- Is the autoregressive operator.

- Autoregressive model then can be written as $\phi(B)\tilde{z}_t = a_t$

Moving Average Models

- Moving average model of order *q*

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}$$

- Moving average operator:

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$$

- Moving average model can be represented as:

$$\tilde{z}_t = \theta(B) a_t$$

Autoregressive Moving Average (ARMA)

- Obtained by mixing Autoregressive and Moving Average models, for greater flexibility in fitting of actual time series.

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \cdots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$

$$\phi(B)\tilde{z}_t = \theta(B)a_t$$

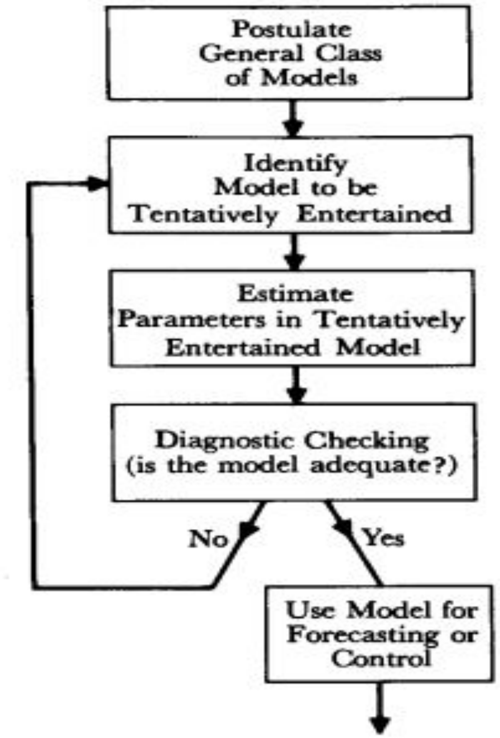
Which has $p+q+2$ parameters.

Basic Ideas in Model Building

- **Parsimony:**
  - Means extreme unwillingness to spend money or use resources(Wiki)
  - The principle of parsimony *stresses to use the smallest possible number of parameters for adequate representation*.
  - There is generally a tradeoff between goodness of fit and parsimony; low parsimony (ie, models with many parameters) tend to have a better fit than high parsimony.
  - This is not usually a good thing; adding more parameters usually results in a good fit for data at hand but may perform poor for predicting other data sets.

Basic Ideas in Model Building (contd...)

- **Iterative stages in the selection of a model :**



Stages in the iterative approach to model building

Stochastic Models & Forecasting

- A model that describes the probability structure of a sequence of observations is called a stochastic process.
- Stationary process : a class of stochastic process. They are assumed to be in a specific form of statistical equilibrium, and in particular, vary about a mean.
    - Useful devices to describe behaviour of stationary process are the autocorrelation function and spectrum.
- A major objective of statistical inference is to infer properties of population from the sample.

Recap…

- A central feature in development of time series models is an assumption of some form of statistical equilibrium. A particular assumption of this kind is that of stationarity.
- A stationary time series is described by its mean, variance and autocorrelation function.

Stochastic Process

- If a series can’t be forecasted deterministically and makes use of probabilistic structure, then it is called as a *statistical time series*.
- A statistical phenomenon that evolves in time according to probabilistic laws is called a *stochastic process*, often as *process.*
- The time series to be analysed may be thought of as a particular realization of the underlying probabilistic mechanism.

Stochastic Process (contd...)

- We can regard the observation z_t at a given time *t*, say *t=25*, as a realization of random variable z_t with probability density function $p(z_t)$.
- Similarly, observations at any two times, say $t_1=25$, $t_2=27$ may be regarded as realization of two random variables z_{t1} ,z_{t2} with the joint probability function $p(z_{t1} ,z_{t2})$.

Stationary Stochastic Process

- Based on the assumption that the process in a particular state of statistical equilibrium.
- A Stochastic process is said to be *strictly stationary* if its *properties are unaffected by a change of time origin,* ie, *if the joint probability distribution of m observations* z_{t1}, z_{t2},... z_{tm} *is same as that of the* z_{t1+l}, z_{t2+k},... z_{tm+k}.

Mean and Variance of a stationary process

- When m=1, the stationarity assumption implies that the probability distribution $p(z_t)$ is the same for all times t and may be written as $p(z)$.
- Hence the stochastic process has a constant mean

$$\mu = E[z_t] = \int_{-\infty}^{\infty} z p(z)\, dz$$

And a constant variance

$$\sigma_z^2 = E[(z_t - \mu)^2] = \int_{-\infty}^{\infty} (z - \mu)^2 p(z)\, dz$$

Mean and Variance of a stationary process(contd...)

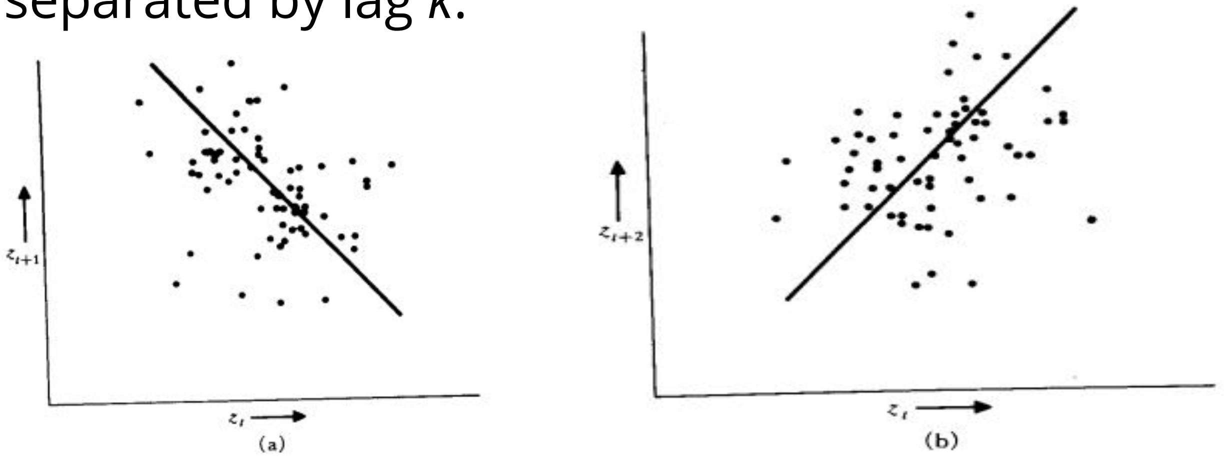
- In addition, the mean and variance of a stationary process may be represented as

Mean: $$\bar{z} = \frac{1}{N}\sum_{t=1}^{N} z_t$$

Variance: $$\hat{\sigma}_z^2 = \frac{1}{N}\sum_{t=1}^{N} (z_t - \bar{z})^2$$

Autocovariance and Autocorrelation coefficients

- The stationarity assumption also implies that the joint probability distribution *$p(z_{t1},z_{t2})$* is the same for all times z_{t1},z_{t2} which are constant interval apart.
- It follows that the nature of this joint distribution can be inferred by plotting a scatter diagram using pairs of values *$p(z_{t1},z_{t2})$* of time series, separated by lag *k*.



- The plots represent negatively and positively correlated data points respectively.

Autocovariance and Autocorrelation coefficients …

- ***Covariance*** :
    - *Measure of how variables change together*.
    - *Expected value of variations of two random variables from their expected values*.
    - Can be +ve(X,Y vary together) or -ve(X,Y vary in opposite directions)
- Let E(X) = m_x and E(Y) = m_y ;
    - *Cov(X,Y) = E(XY) -* $m_x m_y$
- The covariance between z_t and z_{t+k} is called the autocovariance at lag *k* and is defi

$$\gamma_k = \text{cov}\,[z_t, z_{t+k}] = E[(z_t - \mu)(z_{t+k} - \mu)]$$

    -

Autocovariance and Autocorrelation coefficients ...

- ***Correlation***:
    - *Measure of telling how related two variables are*
    - *Simply, it is the scaled version of covariance*.
- Let E(X) = m_x and E(Y) = m_y ;
    - *Cov(X,Y) = E(XY) - $m_x m_y$*
- The correlation between z_t and z_{t+k} is called the autocorrelation at lag *k* and is defined by

$$\rho_k = \frac{E[(z_t - \mu)(z_{t+k} - \mu)]}{\sqrt{E[(z_t - \mu)^2]E[(z_{t+k} - \mu)^2]}} = \frac{E[(z_t - \mu)(z_{t+k} - \mu)]}{\sigma_z^2}$$

=>

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

(for a stationary process)

Positive Definiteness & Autocovariance Matrix

- The covariance matrix associated with a stationary process for observations $(z_{t1}, z_{t2}, \ldots z_{tm})$ is
  -

$$\Gamma_n = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{n-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_{n-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \gamma_{n-3} & \cdots & \gamma_0 \end{bmatrix}$$

$$= \sigma_z^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{n-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{n-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \cdots & 1 \end{bmatrix} = \sigma_z^2 \mathbf{P}_n$$

$\mathbf{P}_n$ is the autocorrelation matrix.

Positive Definiteness & Autocovariance Matrix...

- Now consider any linear function of the random variables *(z_t, z_{t-1},... z_{t-n+1})*

$$L_t = l_1 z_t + l_2 z_{t-1} + \cdots + l_n z_{t-n+1}$$

Since $\mathrm{cov}[z_i, z_j] = \gamma_{|j-i|}$ for a stationary process, the variance of L_t is

$$\mathrm{var}[L_t] = \sum_{i=1}^{n} \sum_{j=1}^{n} l_i l_j \gamma_{|j-i|}$$

Which is necessarily greater than zero if *l's* aren't all zeros.

It follows that autocovariance and autocorrelation matrices are positive definite for stationary process.

Conditions satisfied by autocorrelations of a stationary process

- The positive definiteness of the autocorrelation matrix implies that its determinant and all principle minors are greater than zero.

for $n = 2$

$$\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix} > 0$$

so that

$$1 - \rho_1^2 > 0$$

and hence

$$-1 < \rho_1 < 1$$

Similarly, for $n = 3$, we must have

$$\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix} > 0, \qquad \begin{vmatrix} 1 & \rho_2 \\ \rho_2 & 1 \end{vmatrix} > 0$$

$$\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix} > 0$$

which implies

$$-1 < \rho_1 < 1$$

$$-1 < \rho_2 < 1$$

$$-1 < \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} < 1$$

Autocovariance and autocorrelation functions:

Plot of *gamma$_k$ versus lag k* is called ***the autocovariance function***; and that of *rho$_k$ versus lag k* is called ***the autocorrelation function*** of the stochastic process.

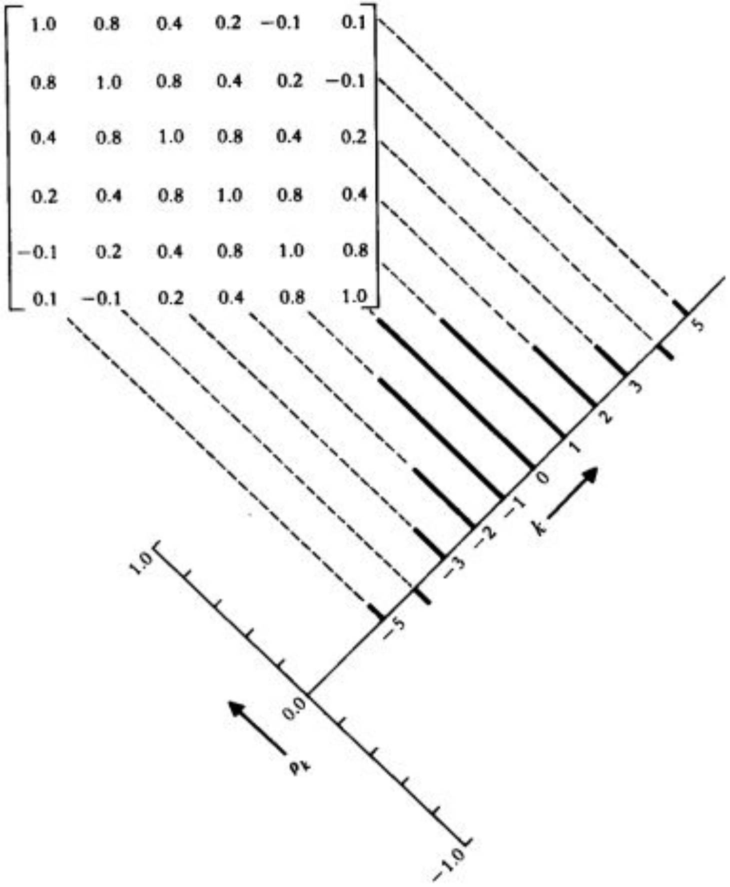


FIG. 2.5 An autocorrelation matrix and the resulting autocorrelation function

Estimation of Autocovariance and Autocorrelation Function

- The most satisfactory estimate of k^{th} lag autocorrelation is
  - $r_k = c_k / c_0$
  - Where

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (z_t - \bar{z})(z_{t+k} - \bar{z}), \qquad k = 0, 1, 2, \ldots, K$$

Thank You