Module 4.1
Standard Models - ARMA(p,q) Family

S. Lakshmivarahan

School of Computer Science
University of Oklahoma
Norman, OK, 73071
USA

Model Building - Goal of DM

- The ultimate goal of DM is to build models that can mimic the data generation process
- Recall: time series represent discrete time observation of a (dynamic) stochastic process
- Much like differential equation provide a natural framework for the analysis of continuous time process, the theory of difference equations provide a mathematical framework for the analysis of discrete time phenomena

Difference Equations - A Classification

- Deterministic vs. stochastic
- Linear vs. nonlinear
- Homogeneous vs. non-homogeneous
- Time varying vs. time invariant
- Based on order

Examples

- $x_{k+1} = x_k + x_{k-1}$, with $x_0 = 0, x_1 = 1$ as initial conditions - linear, second-order, homogeneous (no external forcing), time invariant deterministic model - Fibonacci sequence generator
- $x_k = \phi_1 x_{k-1} + \phi_2 x_{k-2} + \epsilon_k$ with (x_0, x_1) specified is a second order, non-homogeneous time invariant, stochastic, linear difference equations, when ϵ_k is the white noise sequence

A Class of Linear Stochastic Models - ARMA(p,q)

- Develop and analyze a class of linear, time invariant, non-homogeneous stochastic models of various orders
- These are called an autoregressive (AR) and moving average (MA) models

AR(p) Models - (p - Order of the Model)

- AR(1): $x_k = c + \phi_1 x_{k-1} + \epsilon_k$
- AR(2): $x_k = c + \phi_1 x_{k-1} + \phi_2 x_{k-2} + \epsilon_k$
- AR(p): $x_k = c + \phi_1 x_{k-1} + \phi_2 x_{k-2} + \cdots + \phi_p x_{k-p} + \epsilon_k$
- The constant c is related to mean $\mu = \mathbf{E}(x_k)$
- AR(p) has p (constant) coefficients $\phi_1, \phi_2, \cdots, \phi_p$
- Need p initial conditions and $\{\epsilon_k\}$ to generate x_k
- $\{\epsilon_k\}$ is white noise: $\mathbf{E}(\epsilon_k) = 0$, $\mathbf{Var}\{\epsilon_k\} = \sigma^2$, $\mathbf{Cov}(x_i, x_k) = 0$, for $i \neq k$
- Since x_k is expressed as a linear combination of its past p values, this is called autoregressive (AR) models

MA(q) Models - (q - Order of Model)

- In MA(q) model, x_k is expressed as a linear combination of $(q+1)$ values $\epsilon_k, \epsilon_{k-1}, \cdots, \epsilon_{k-q}$ of white noise sequence and hence the name MA
- MA(1): $x_k = \epsilon_k + \theta_1 \epsilon_{k-1}$
- MA(2): $x_k = \epsilon_k + \theta_1 \epsilon_{k-1} + \theta_2 \epsilon_{k-2}$
- MA(q): $x_k = \epsilon_k + \theta_1 \epsilon_{k-1} + \theta_2 \epsilon_{k-2} + \cdots \theta_q \epsilon_{k-q}$
- MA(q) has q constant coefficient $\theta_1, \theta_2, \cdots, \theta_q$
- Again need the specification of $(q+1)$ values of $\epsilon_j, \ \ (k-q \leq j \leq k)$ to compute x_k

ARMA(p,q) Models - (p,q) Order of the Model

- This model is specified by $(p+q)$ constant coefficients
- ARMA(2,2):
  $x_k = c + \phi_1 x_{k-1} + \phi_2 x_{k-2} + \epsilon_k + \theta_1 \epsilon_{k-1} + \theta_2 \epsilon_{k-2}$
- Needs the specification of p initial values of x_j and $(q+1)$ values of ϵ_j

Lag Operator

- Recall $Lx_k = x_{k-1}$
- Accordingly, AR(2) is

$$\phi(L)\, x_k = (1 - \phi_1 L - \phi_2 L^2)\, x_k = \epsilon_k$$
$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 \text{ - AR polynomial of degree 2}$$

- MA(2):

$$x_k = \psi(L)\epsilon_k = (1 + \theta_1 L + \theta_2 L^2)\, \epsilon_k$$
$$\psi(L) = 1 + \theta_1 L + \theta_2 L^2 \text{ - MA polynomial of degree 2}$$

- It turns out that the roots of the AR polynomial $\phi(x) = 0$ and MA polynomial $\psi(x) = 0$ play a critical role in the analysis of the models

ARMA(p,q) Model

- Let

$$\phi(L) : 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$$

$$\psi(L) : 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$$

- $\phi(L)x_k = \psi(L)\epsilon_k$ then represents the ARMA(p,q) model in terms of AR and MA polynomials

Examples of Non-Linear Models

- GARCH(1,1) model: $x_k = \sqrt{h_k}\, \epsilon_k, \quad \epsilon_k \sim \textbf{IID}\, N(0, 1)$
  $h_k = a_0 + a_1 h_{k-1} + a_2 \epsilon_{k-1}^2$
- Bilinear model: $x_k = \theta_1 x_{k-1} \epsilon_{k-1} + \epsilon_k$
- Nonlinear MA model: $x_k = \epsilon_k + \theta_1 \epsilon_{k-1} \epsilon_{k-2}$
- Nonlinear AR model: $x_k = (1 + c\epsilon_{k-1}) x_{k-1} + \epsilon_k$
- Threshold model:

$$x_k = \begin{cases} - \quad \phi_1 x_{k-1} + \epsilon_k & \text{if } \ x_{k-1} < 1 \\ + \quad \phi_2 x_{k-1} + \epsilon_k & \text{otherwise} \end{cases}$$

- Exponential nonlinearity
  $x_k = \left(a_1 + a_2 \exp\left[\frac{-x_{k-1}^2}{2} \right] \right) x_{k-1} + \epsilon_k$
- Chaotic time series analysis

References - Nonlinear Time Series

- Patterson, Douglas M and Ashley, Richard A, "A nonlinear time series workshop: A toolkit for detecting and identifying nonlinear serial dependence" *Kluwer Academic Publishers, Boston, MA, USA* (2000)
- Holger and Schreiber, Thomas, "Nonlinear time series analysis", *Cambridge University Press* New York (2004)