INDIAN INSTITUTE OF SCIENCE

STOCHASTIC HYDROLOGY

Lecture -17

Course Instructor : Prof. P. P. MUJUMDAR

Department of Civil Engg., IISc.

Summary of the previous lecture

- Behavior of AR and MA process
- Parameter estimation
  - Matlab function “armax”
- Model selection
  - Maximum likelihood rule

ARIMA Models

Mean square error, MSE (Prediction approach):

- Using a portion of available data (N/2) estimate the parameters of different models
- Forecast the series one step ahead by using the candidate models
- Estimate the MSE corresponding to each model
- The model with least value of MSE is selected for prediction

ARIMA Models

The one step ahead forecast for ARMA(p, q) is

$$\hat{X}_{t+1} = \sum_{j=1}^{p} \phi_j X_{t-j} + \sum_{j=1}^{q} \phi_j e_{t-j}$$

The error for one step ahead forecast is

$$e_{t+1} = X_{t+1} - \hat{X}_{t+1}$$

If the series consists on N observations, the first N/2 observations are used for parameter estimation and N/2+1 to N are used for error series calculation.

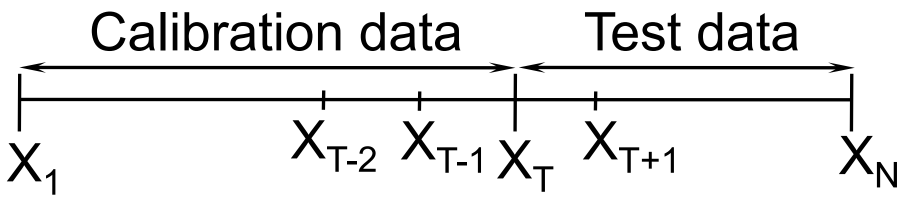
ARIMA Models

The MSE for model is

$$MSE = \frac{\sum_{i=\frac{N}{2}+1}^{N} e_i^2}{N/2}$$

ARIMA Models

3. Model testing / Validation:



First 'T' values are used to build the model (say 50% of the available data) and the rest of data is used to validate the model.

All the tests are carried out on the residual series only.

ARIMA Models

The tests are performed to examine whether the

valid for the model selection

- The residual series has zero mean
- No significant periodicities are present in the residual series
- The residual series is uncorrelated

$$\underbrace{e_t}_{\text{Residual}} = \underbrace{X_t}_{\text{Data}} - \underbrace{\left(\sum_{j=1}^{m_1} \phi_j X_{t-j} + \sum_{j=1}^{m_2} \theta_j e_{t-j} \right)}_{\text{Simulated from the model}}$$

ARIMA Models

Validation tests are listed here

- Significance of residual mean
- Significance of periodicities
- Cumulative periodogram test or Bartlett's test
- White noise test
  - Whittle's test
  - Portmanteau test

Significance of residual mean

Significance of residual mean:

- This test examines the validity of the assumption that the error series e(t) has zero mean
- A statistic η(e) is defined as

$\{e_t\}$

$$\eta(e) = \frac{N^{1/2}\overline{e}}{\hat{\rho}^{1/2}}$$

Variance

N: no. of values

Where

$\overline{e}$ is the estimate of the residual mean

$\hat{\rho}$ is the estimate of the residual variance

Ref: Kashyap R.L. and Ramachandra Rao.A, "Dynamic stochastic models from empirical data", Academic press, New York , 1976

Significance of residual mean

- The statistic $\eta(e)$ is approximately distributed as t (α, N–1), where α is the significance level at which the test is being carried out.
- If the value of $\eta(e) \leq t(\alpha, N–1)$, then the mean of the residual series is not significantly different from zero – series passes the test.

Significance of periodicities

Significance of periodicities:

- This test ensures that no significant periodicities are present in the residual series
- The test is conducted for different periodicities and the significance of each of the periodicities is tested.
- A statistic η(e) is defined as

$$\eta(e) = \frac{\gamma_k^2 (N-2)}{4\hat{\rho}_1}$$

γ_k corresponds to the periodicity being tested

Significance of periodicities

Where $\gamma_k^2 = \alpha_k^2 + \beta_k^2$

$$\hat{\rho}_1 = \frac{1}{N}\left[\sum_{t=1}^{N}\left\{e_t - \hat{\alpha}_k \cos\left(\omega_k t\right) - \hat{\beta}_k \sin\left(\omega_k t\right)\right\}^2\right]$$

$$\alpha_k = \frac{2}{N}\sum_{t=1}^{n} e_t \cos\left(\omega_k t\right)$$

$$\beta_k = \frac{2}{N}\sum_{t=1}^{n} e_t \sin\left(\omega_k t\right)$$

$2\pi/\omega_k$ is the periodicity for which test is being carried out.

Significance of periodicities

- The statistic $\eta(e)$ is approximately distributed as $F_{\alpha}(2, N–2)$, where α is the significance level at which the test is being carried out.
- If the value of $\eta(e) \leq F_{\alpha}(2, N–2)$, then the periodicity is not significant.

Bartlett's test

Cumulative periodogram test or Bartlett's test :

- This test is also carried out to examine significant periodicities in the residual series
- This test is more convenient computationally and is preferred because of its ability to test all the periodicities at a time.

Bartlett's test

$$\gamma_k^2 = \left\{\frac{2}{N}\sum_{t=1}^{N} e_t \cos\left(\omega_k t\right)\right\}^2 + \left\{\frac{2}{N}\sum_{t=1}^{N} e_t \sin\left(\omega_k t\right)\right\}^2$$

k = 1,2,……N/2

$$g_k = \frac{\sum_{j=1}^{k} \gamma_j^2}{\sum_{k=1}^{N/2} \gamma_k^2} \qquad 0 \leq g_k \leq 1$$

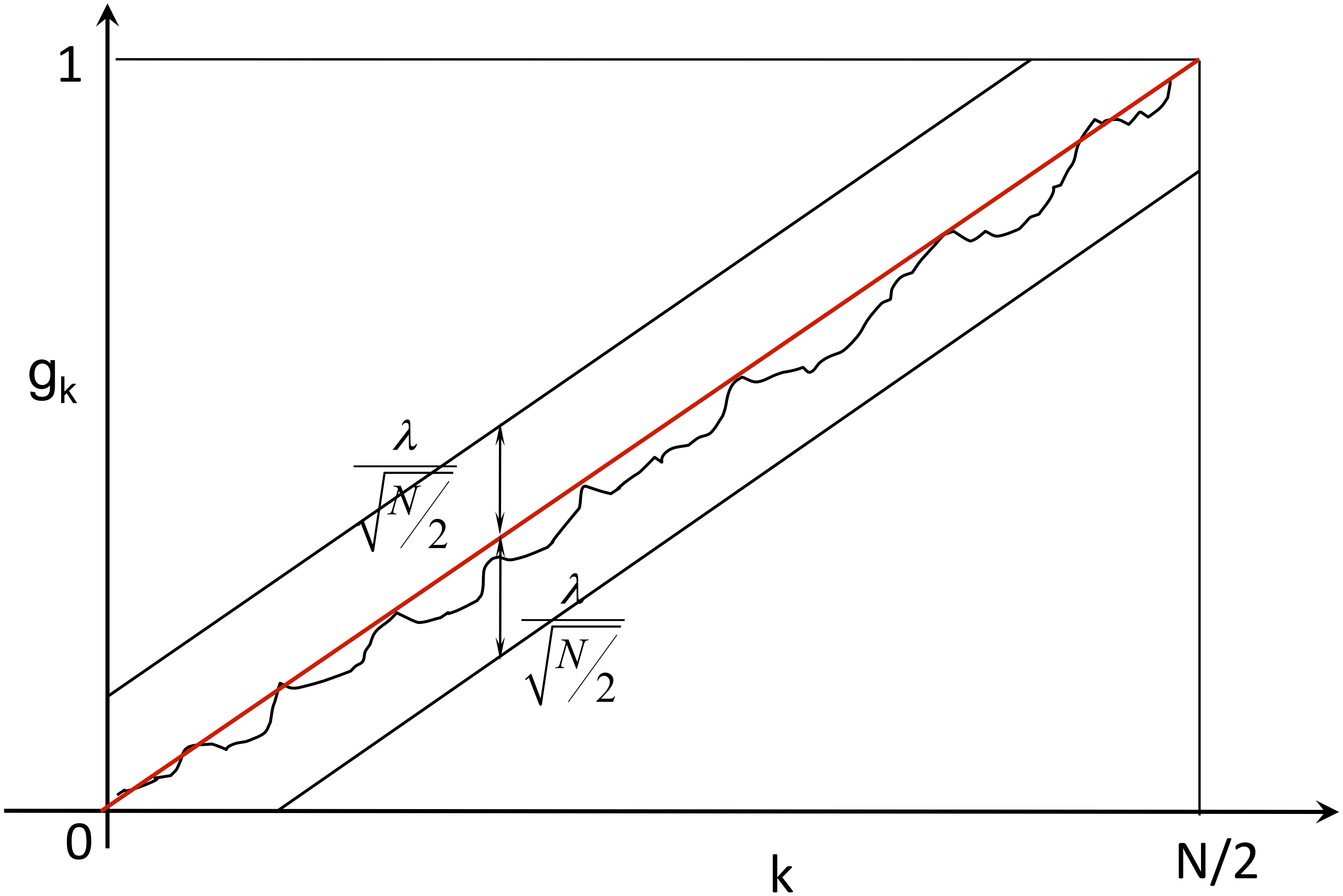
The plot of g_k vs k is called as cumulative periodogram

Ref: Kashyap R.L. and Ramachandra Rao.A, "Dynamic stochastic models from empirical data", Academic press, New York , 1976

Bartlett's test

- On the cumulative periodogram two confidence limits ($\lambda/(N/2)^{1/2}$) are drawn on either side of line joining (0, 0) and (N/2, 1)
- The value of λ prescribed for 95% confidence limits is 1.35 and for 99% confidence limits is 1.65
- If all the values of g_k lie within the significance band, there is no significant periodicities in the series.
- If a value of g_k lies outside the significance band, the periodicity corresponding to that value of g_k is significant.

Bartlett’s test



Whittle's test for white noise

White noise test (Whittle's test):

- This test is carried out to test the absence of correlation in the series.
- The covariance r_k at lag k of the error series e(t)

$$r_k = \frac{1}{N-k} \sum_{j=k+1}^{N} e_j e_{j-k} \qquad k = 0, 1, 2, \ldots\ldots k_{max}$$

- The value of k_{max} is normally chosen as 0.15N

Ref: Kashyap R.L. and Ramachandra Rao.A, "Dynamic stochastic models from empirical data", Academic press, New York

Whittle's test for white noise

- The covariance matrix is

$$\Gamma_{n1} = \begin{bmatrix} r_0 & r_1 & r_2 & . & . & r_{k_{\max}} \\ r_1 & r_0 & r_1 & . & . & r_{k_{\max}-1} \\ r_2 & & & & & \\ . & & & & & \\ . & & & & & \\ r_{k_{\max}} & r_{k_{\max}-1} & & & & r_0 \end{bmatrix}_{k_{max} \times k_{max}}$$

Whittle's test for white noise

- A statistic η(e) is defined as

$$\eta(e) = \frac{N}{n1-1}\left(\frac{\hat{\rho}_0}{\hat{\rho}_1} - 1\right) \qquad n1 = k_{max}$$

Where $\hat{\rho}_0$ is the lag zero correlation =1, and

$$\hat{\rho}_1 = \frac{\det \Gamma_{n1}}{\det \Gamma_{n1-1}}$$

The matrix Γ_{n1-1} is constructed by eliminating the last row and the last column from the Γ_{n1} matrix.

Whittle's test for white noise

- The statistic $\eta(e)$ is approximately distributed as $F_{\alpha}(n1, N–n1)$, where α is the significance level at which the test is being carried out.
- If the value of $\eta(e) \leq F_{\alpha}(n1, N–n1)$, then the residual series is uncorrelated.

Portmanteau test for white noise

White noise test (Portmanteau test):

- This test is also carried out to test the absence of correlation in the series.
- This test also uses the covariance r_k defined earlier.
- A statistic η(e) is defined as

$$\eta(e) = (N - n1) \sum_{k=1}^{n1} \left(\frac{r_k}{r_0} \right)^2$$

Ref: Kashyap R.L. and Ramachandra Rao.A, “Dynamic stochastic models from empirical data”, Academic press, New York

Portmanteau test for white noise

- The statistic $\eta(e)$ is approximately distributed as $\chi^2_\alpha(n1)$, where α is the significance level at which the test is being carried out.
- The value of n1 is normally chosen as 0.15N
- If the value of $\eta(e) \leq \chi^2_\alpha(n1)$, then the residual series is uncorrelated.
- Kashyap & Rao(1976) have proved that the Portmanteau test is uniformly inferior to Whittle's test and recommended the latter for applications.

Ref: Kashyap R.L. and Ramachandra Rao.A, "Dynamic stochastic models from empirical data", Academic press, New York

ARIMA Models

Data Generation:

Consider AR(1) model,

$$X_t = \phi_1 X_{t-1} + e_t$$

e.g., $\phi_1 = 0.5$: AR(1) model is

$$X_t = 0.5X_{t-1} + e_t$$

Choose e_t terms with zero mean and uncorrelated

ARIMA Models

Say $X_1 = 3.0$

$X_2 = 0.5*3.0 + 0.335$
$= 1.835$

$X_3 = 0.5*1.835 + 1.226$
$= 2.14$

And so on…

ARIMA Models

Consider ARMA(1, 1) model,

$$X_t = \phi_1 X_{t-1} + \theta_1 e_{t-1} + e_t$$

e.g., $\phi_1 = 0.5$, $\theta_1 = 0.4$: ARMA(1, 1) model is written as

with zero mean and uncorrelated

$$X_t = 0.5X_{t-1} + 0.4e_{t-1} + e_t$$

Choose e_{t-1} terms as previous e_t and set initial value as zero

ARIMA Models

Say $X_1 = 3.0$

$X_2 = 0.5*3.0 + 0.4*0 + 0.667$
$= 2.167$

$X_3 = 0.5*2.167 + 0.4*0.667 + 1.04$
$= 2.39$

$X_4 = 0.5*2.39 + 0.4*1.04 + 2.156$
$= 3.767$ and so on...

ARIMA Models

Data Forecasting:

Consider AR(1) model,

$X_t = \phi_1 X_{t-1} + e_t$

Expected value is considered.

$$E\left[X_t\right] = \phi_1 E\left[X_{t-1}\right] + E\left[e_t\right]$$

$$\hat{X}_t = \phi_1 X_{t-1}$$

Expected value of e_t is zero

ARIMA Models

Consider ARMA(1, 1) model,

$$X_t = \phi_1 X_{t-1} + \theta_1 e_{t-1} + e_t$$

$$E[X_t] = \phi_1 X_{t-1} + \theta_1 e_{t-1} + 0$$

Error in forecast in the previous period

e.g., $\phi_1 = 0.5$, $\theta_1 = 0.4$: Forecast model is written as

$$X_t = 0.5X_{t-1} + 0.4e_{t-1}$$

ARIMA Models

Say $X_1 = 3.0$

Initial error assumed to be zero

$$\hat{X}_2 = 0.5 \times 3.0 + 0.4 \times 0$$
$$= 1.5$$

$X_2 = 2.8$

Error $e_2 = 2.8 - 1.5 = 1.3$

$$\hat{X}_3 = 0.5 \times 2.8 + 0.4 \times 1.3$$
$$= 1.92$$

Actual value to be used

ARIMA Models

$X_3 = 1.8$

Error $e_3 = 1.8 – 1.92 = -0.12$

$$\hat{X}_4 = 0.5 \times 1.8 + 0.4 \times \left(-0.12\right)$$
$$= 0.852$$

and so on.

Markov Chains

Markov Chains:

- Markov chain is a stochastic process with the property that value of process X_t at time t depends on its value at time t-1 and not on the sequence of other values (X_{t-2} , X_{t-3},……. X_0) that the process passed through in arriving at X_{t-1}.

$$P\left[X_t / X_{t-1}, X_{t-2}, \ldots.. X_0\right] = P\left[X_t / X_{t-1}\right]$$

Single step Markov chain

Markov Chains

$$P\left[X_t = a_j / X_{t-1} = a_i\right]$$

- The conditional probability gives the probability at time t will be in state 'j', given that the process was in state 'i' at time t-1.
- The conditional probability is independent of the states occupied prior to t-1.
- For example, if X_{t-1} is a dry day, what is the probability that X_t is a dry day or a wet day.
- This probability is commonly called as transitional probability

Markov Chains

$$P\left[X_t = a_j / X_{t-1} = a_i\right] = P_{ij}^t$$

- Usually written as P_{ij}^t indicating the probability of a step from a_i to a_j at time 't'.
- If P_{ij} is independent of time, then the Markov chain is said to be homogeneous.

i.e., $P_{ij}^t = P_{ij}^{t+\tau} \quad \forall \quad t$ and τ

the transitional probabilities remain same across time

Markov Chains

Transition Probability Matrix(TPM):

$$
P = \begin{array}{c} t \\ \downarrow \\ 1 \\ 2 \\ 3 \\ . \\ . \\ m \end{array}
\begin{array}{c} t+1 \rightarrow \;\; 1 \quad\quad 2 \quad\quad 3 \quad\quad . \quad\quad . \quad\quad m \\ \\ \left[\begin{array}{cccccc} P_{11} & P_{12} & P_{13} & . & . & P_{1m} \\ P_{21} & P_{22} & P_{23} & . & . & P_{2m} \\ P_{31} & & & & & \\ . & & & & & \\ . & & & & & \\ P_{m1} & P_{m2} & & & & P_{mm} \end{array} \right]_{m \times m} \end{array}
$$

Markov Chains

$$\sum_{j=1}^{m} P_{ij} = 1 \quad \forall \text{ j}$$

- Elements in any row of TPM sum to unity (stochastic matrix)
- TPM can be estimated from observed data by tabulating the number of times the observed data went from state 'i' to 'j'
- $P_j^{(n)}$ is the probability of being in state 'j' in the time step 'n'.

Markov Chains

- $p_j^{(0)}$ is the probability of being in state 'j' in period t = 0.

$$p^{(0)} = \begin{bmatrix} p_1^{(0)} & p_2^{(0)} & . & . & p_m^{(0)} \end{bmatrix}_{1\times m}$$

…. Probability vector at time 0

$$p^{(n)} = \begin{bmatrix} p_1^{(n)} & p_2^{(n)} & . & . & p_m^{(n)} \end{bmatrix}_{1\times m}$$

…. Probability vector at time 'n'

- Let $p^{(0)}$ is given and TPM is given

$$p^{(1)} = p^{(0)} \times P$$

Markov Chains

$$p^{(1)} = \begin{bmatrix} p_1^{(0)} & p_2^{(0)} & . & . & p_m^{(0)} \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} & P_{13} & . & . & P_{1m} \\ P_{21} & P_{22} & P_{23} & . & . & P_{2m} \\ P_{31} & & & & & \\ . & & & & & \\ P_{m1} & P_{m2} & & & & P_{mm} \end{bmatrix}$$

$$= p_1^{(0)} P_{11} + p_2^{(0)} P_{21} + ... + p_m^{(0)} P_{m1}$$ …. Probability of going to state 1

$$= p_1^{(0)} P_{12} + p_2^{(0)} P_{21} + ... + p_m^{(0)} P_{m2}$$ …. Probability of going to state 2

And so on…

Markov Chains

Therefore

$$p^{(1)} = \begin{bmatrix} p_1^{(1)} & p_2^{(1)} & . & . & p_m^{(1)} \end{bmatrix}_{1\times m}$$

$$\begin{aligned} p^{(2)} &= p^{(1)} \times P \\ &= p^{(0)} \times P \times P \\ &= p^{(0)} \times P^2 \end{aligned}$$

In general,

$$p^{(n)} = p^{(0)} \times P^n$$

Markov Chains

- As the process advances in time, $p_j^{(n)}$ becomes less dependent on $p^{(0)}$

  The probability of being in state 'j' after a large number of time steps becomes independent of the initial state of the process.
- The process reaches a steady state ay very large n

$$p = p \times P^n$$

- As the process reach steady state, TPM remains constant

Example – 1

Consider the TPM for a 2-state (state 1 is non-rainfall day and state 2 is rainfall day) first order homogeneous Markov chain as

$$TPM = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}$$

Obtain the

1. probability of day 1 is non-rainfall day / day 0 is rainfall day
2. probability of day 2 is rainfall day / day 0 is non-rainfall day
3. probability of day 100 is rainfall day / day 0 is non-rainfall day

Example – 1 (contd.)

1. probability of day 1 is non-rainfall day / day 0 is rainfall day

$$TPM = \begin{array}{c c} & \begin{array}{cc} \text{No rain} & \text{rain} \end{array} \\ \begin{array}{c} \text{No rain} \\ \text{rain} \end{array} & \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \end{array}$$

The probability is 0.4

2. probability of day 2 is rainfall day / day 0 is non-rainfall day

$$p^{(2)} = p^{(0)} \times P^2$$

Example – 1 (contd.)

$$p^{(2)} = \begin{bmatrix} 0.7 & 0.3 \end{bmatrix} \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}$$

$$= \begin{bmatrix} 0.61 & 0.39 \end{bmatrix}$$

The probability is 0.39

3. probability of day 100 is rainfall day / day 0 is non-rainfall day

$$p^{(n)} = p^{(0)} \times P^n$$

Example – 1 (contd.)

$$P^2 = P \times P$$

$$= \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} = \begin{bmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{bmatrix}$$

$$P^4 = P^2 \times P^2 = \begin{bmatrix} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{bmatrix}$$

$$P^8 = P^4 \times P^4 = \begin{bmatrix} 0.5715 & 0.4285 \\ 0.5714 & 0.4286 \end{bmatrix}$$

$$P^{16} = P^8 \times P^8 = \begin{bmatrix} 0.5714 & 0.4286 \\ 0.5714 & 0.4286 \end{bmatrix}$$

Example – 1 (contd.)

Steady state probability

$$p = \begin{bmatrix} 0.5714 & 0.4286 \end{bmatrix}$$

For steady state,

$$p = p \times P^n$$

$$= \begin{bmatrix} 0.5714 & 0.4286 \end{bmatrix} \begin{bmatrix} 0.5714 & 0.4286 \\ 0.5714 & 0.4286 \end{bmatrix}$$

$$= \begin{bmatrix} 0.5714 & 0.4286 \end{bmatrix}$$