Time Series

Sudipta Das

Assistant Professor,
Department of Computer Science,
Ramakrishna Mission Vivekananda Educational & Research Institute

Outline I

1. Forecasting Stationary Time Series
   - Recursive Forecasting

Forecasting Stationary Time Series I

- We consider the problem of predicting the values $X_{n+h}, h > 0$, of a stationary time series with known mean μ and known autocovariance function $\gamma(\cdot)$ in terms of the values $\{X_n, \ldots, X_1\}$, up to time n.
    - Forecasting as **AR** model
- Our goal is to find the linear combination of $1, X_n, X_{n-1}, \ldots, X_1$, ($\hat{X}_{n+h} = a_0 + a_1 X_n + \cdots + a_n X_1 = X_{n+h}^n$) that forecasts X_{n+h} with minimum mean squared error, i.e.

$$E\left(X_{n+h} - a_0 - a_1 X_n - \cdots - a_n X_1\right)^2$$

is minimized.

- Minimization yields
  - Normal Equations

$$E\left[X_{n+h} - a_0 - \sum_{i=1}^{n} a_i X_{n+1-i}\right] = 0$$

and

$$E\left[\left(X_{n+h} - a_0 - \sum_{i=1}^{n} a_i X_{n+1-i}\right) X_{n+1-j}\right] = 0, \text{ for } j = 1, \ldots, n$$

- Solutions

$$a_0 = \mu(1 - \sum_{i=1}^{n} a_i)$$

and

$$\mathbf{a_n} = [a_1, \ldots, a_n]'$$

as the solution of the equation

$$\Gamma_n \mathbf{a_n} = \gamma_{\mathbf{n}}(h),$$

where

  - $\gamma_{\mathbf{n}}(h) = [\gamma(h), \gamma(h+1), \ldots, \gamma(h+n-1)]'$ and
  - $\Gamma_n = [\gamma(i-j)]_{i,j=1}^{n}$

Forecasting Stationary Time Series IV

- Best Linear Unbiased Estimator

$$X_{n+h}^{n} = \mu + \mathbf{a_n}' \left(\mathbf{X_n} - \mu \mathbf{1_n} \right),$$

where $\mathbf{X_n} = [X_n, X_{n-1}, \ldots, X_1]'$ and $\mathbf{1_n} = [\underbrace{1, \ldots, 1}_{n-times}]'$

- Expected value of the prediction error (i.e., first normal equation)

$$E[X_{n+h} - X_{n+h}^{n}] = 0$$

- Mean square prediction error

$$\begin{aligned}
E\left(X_{n+h} - X_{n+h}^{n}\right)^2 &= E\left[(X_{n+h} - \mu) - \mathbf{a_n}'(\mathbf{X_n} - \mu\mathbf{1_n})\right]^2 \\
&= \gamma(0) - 2\mathbf{a_n}'\gamma_{\mathbf{n}}(h) + \mathbf{a_n}'\Gamma_n(h)\mathbf{a_n} \\
&= \gamma(0) - \mathbf{a_n}'\gamma_{\mathbf{n}}(h) \\
&= \gamma(0) - \gamma_{\mathbf{n}}'(h)\Gamma_n^{-1}\gamma_{\mathbf{n}}(h)
\end{aligned}$$

Forecasting Stationary Time Series V

- Example: One-step prediction of an AR(1) series

$$X_t = \phi X_{t-1} + Z_t, t = 0, \pm 1, \dots$$

where $|\phi| < 1$ and $Z_t \sim WN(0, \sigma^2)$.

- Solution:

$$a_0 = 0$$

and

$$X_{n+1}^n = \mathbf{a_n}'\mathbf{X_n},$$

where $\mathbf{X_n} = [X_n, X_{n-1}, \dots, X_1]'$ and $\mathbf{a_n} = [\phi, 0, \dots, 0]'$ is the solution of

$$\begin{aligned} \Gamma_n \mathbf{a_n} &= \gamma_\mathbf{n}(1) \\ \begin{bmatrix} 1 & \phi & \phi^2 & \dots & \phi^{n-1} \\ \phi & 1 & \phi & \dots & \phi^{n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \phi^{n-1} & \phi^{n-2} & \phi^{n-3} & \dots & 1 \end{bmatrix} \mathbf{a_n} &= \begin{bmatrix} \phi \\ \phi^2 \\ \vdots \\ \phi^n \end{bmatrix} \end{aligned}$$

Forecasting Stationary Time Series VI

- Therefore the best linear predictor of X_{n+1} in terms of $\{X_1, \ldots, X_n\}$ is

$$X_{n+1}^n = \mathbf{a_n}'\mathbf{X_n} = \phi X_n$$

- The mean square error is

$$\begin{aligned} E\left(X_{n+1} - X_{n+1}^n\right)^2 &= \gamma(0) - \mathbf{a_n}'\gamma_\mathbf{n}(1) \\ &= \gamma(0)\left[1 - \phi\rho(1)\right] \\ &= \sigma^2 \end{aligned}$$

Forecasting Stationary Time Series VII

- Remark: For stationary time series $\{Y_t\}$ with non-zero mean μ, the best linear predictor of Y_{n+h} can be determined by the following steps
  - Subtract μ from the series Y_t to get the zero-mean series X_t $[X_t = Y_t - \mu,]$
  - Finding the best linear predictor of X_{n+h} in terms of $X_n, \ldots, X_1$ and
  - Then adding μ to it.
- We, therefore, restrict attention to zero-mean stationary time series.

Recursive Forecasting I

- $h-$step forecasting

$$X_{n+h}^{n} = \mathbf{a_n}'\mathbf{X_n}$$

- Potential problem: Determination of $\mathbf{a_n}$ from the set of linear equation $\Gamma_n \mathbf{a_n} = \gamma_\mathbf{n}(h)$, may be difficult and time-consuming.
- Remedy: Go for recursive algorithm
  - We start with finding one-step predictor X_{n+1}^{n} based on n observations
  - then find the two-step predictor X_{n+2}^{n+1}, based on $n+1$ previous observations (n observed and 1 predicted observation among them)
  - and continue till the $h-$step predictor X_{n+h}^{n+h-1},

Recursive Forecasting II

- One step Predicting equation

$$X_{n+1}^{n} = \boldsymbol{\phi_n}'\mathbf{X_n} = \phi_{n1}X_n + \cdots + \phi_{nn}X_1,$$

where $\boldsymbol{\phi_n} = [\phi_{n1}, \ldots, \phi_{nn}]' = \Gamma_n^{-1}\boldsymbol{\gamma_n}$ and $\boldsymbol{\gamma_n} = [\gamma(1), \gamma(2), \ldots, \gamma(n)]'$ with the corresponding MSE

$$v_n := E(X_{n+1} - X_{n+1}^{n})^2 = \gamma(0) - \boldsymbol{\phi_n}'\boldsymbol{\gamma_n}$$

  - Again Determination of $\boldsymbol{\phi_n}$ involves matrix inversion.
- Therefore, we go for recursive solution for one step prediction

Durbin-Levinson algorithm I

- One step Recursive Forecast (Durbin-Levinson algorithm)
  - Set a one step predicting equation based on single (current) observation

$$X_{n+1}^{n,n} = \phi_{11} X_n$$

  - Compute ϕ_{11} and v_0 as follows

$$\phi_{11} = \frac{\gamma(1)}{\gamma(0)}$$

and

$$v_0 = \gamma(0).$$

Durbin-Levinson algorithm II

- Recursively, set one step predicting equations based on (current) n observation

$$X_{n+1}^{n} = X_{n+1}^{1,n} = \phi_{n1}X_n + \cdots + \phi_{nn}X_1,$$

and

- Compute the coefficients $\phi_{n1}, \ldots, \phi_{nn}$ recursively from the following equations

$$\phi_{nn} = \left[\gamma(n) - \sum_{j=1}^{n-1} \phi_{n-1,j}\gamma(n-j)\right] v_{n-1}^{-1},$$

$$\begin{bmatrix} \phi_{n1} \\ \vdots \\ \phi_{n,n-1} \end{bmatrix} = \begin{bmatrix} \phi_{n-1,1} \\ \vdots \\ \phi_{n-1,n-1} \end{bmatrix} - \phi_{nn} \begin{bmatrix} \phi_{n-1,n-1} \\ \vdots \\ \phi_{n-1,1} \end{bmatrix}$$

and

$$v_n = v_{n-1}[1 - \phi_{nn}^2]$$

Durbin-Levinson algorithm III

- Alternative compact form

$$\phi_{nn} = \left[\gamma(n) - \phi_{\mathbf{n-1}}^{(\mathbf{r})\,\prime}\boldsymbol{\gamma}_{\mathbf{n-1}}\right] v_{n-1}^{-1}, \tag{1}$$

$$\begin{bmatrix} \phi_{n1} \\ \vdots \\ \phi_{n,n-1} \end{bmatrix} = \boldsymbol{\phi}_{\mathbf{n-1}} - \phi_{nn}\boldsymbol{\phi}_{\mathbf{n-1}}^{(\mathbf{r})}, \tag{2}$$

$$v_n = v_{n-1}[1 - \phi_{nn}^2] \tag{3}$$

where $\boldsymbol{\phi}_{\mathbf{k}}^{(\mathbf{r})} = [\phi_{k,k}, \phi_{k,k-1}, \ldots, \phi_{k1}]'$

Durbin-Levinson algorithm IV

- Proof
  - $\Gamma_1\phi_1 = \gamma_1$ follows from $\gamma(0)\phi_1 = \gamma(1)$
  - Let $\Gamma_n\boldsymbol{\phi}_\mathbf{n} = \boldsymbol{\gamma}_\mathbf{n}$ be true for $n = k$, then

$$
\begin{aligned}
\Gamma_{k+1}\boldsymbol{\phi}_{\mathbf{k+1}} &= \begin{bmatrix} \Gamma_k & \boldsymbol{\gamma}_\mathbf{k}^{(\mathbf{r})} \\ \boldsymbol{\gamma}_\mathbf{k}^{(\mathbf{r})\prime} & \gamma(0) \end{bmatrix} \begin{bmatrix} \boldsymbol{\phi}_\mathbf{k} - \phi_{k+1,k+1}\boldsymbol{\phi}_\mathbf{k}^{(\mathbf{r})} \\ \phi_{k+1,k+1} \end{bmatrix} \\
&= \begin{bmatrix} \Gamma_k\boldsymbol{\phi}_\mathbf{k} - \phi_{k+1,k+1}\Gamma_k\boldsymbol{\phi}_\mathbf{k}^{(\mathbf{r})} + \phi_{k+1,k+1}\boldsymbol{\gamma}_\mathbf{k}^{(\mathbf{r})} \\ \boldsymbol{\gamma}_\mathbf{k}^{(\mathbf{r})\prime}\boldsymbol{\phi}_\mathbf{k} - \phi_{k+1,k+1}\boldsymbol{\gamma}_\mathbf{k}^{(\mathbf{r})\prime}\boldsymbol{\phi}_\mathbf{k}^{(\mathbf{r})} + \gamma(0)\phi_{k+1,k+1} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{\gamma}_\mathbf{k} - \phi_{k+1,k+1}\boldsymbol{\gamma}_\mathbf{k}^{(\mathbf{r})} + \phi_{k+1,k+1}\boldsymbol{\gamma}_\mathbf{k}^{(\mathbf{r})} \\ \boldsymbol{\gamma}_\mathbf{k}^{(\mathbf{r})\prime}\boldsymbol{\phi}_\mathbf{k} + \phi_{k+1,k+1}\left(\gamma(0) - \boldsymbol{\gamma}_\mathbf{k}^{(\mathbf{r})\prime}\boldsymbol{\phi}_\mathbf{k}^{(\mathbf{r})}\right) \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{\gamma}_\mathbf{k} \\ \boldsymbol{\gamma}_\mathbf{k}^{(\mathbf{r})\prime}\boldsymbol{\phi}_\mathbf{k} + \phi_{k+1,k+1}v_k \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{\gamma}_\mathbf{k} \\ \gamma(k+1) \end{bmatrix}, \text{[by (1)]} \\
&= \boldsymbol{\gamma}_{\mathbf{k+1}}
\end{aligned}
$$

Therefore, true for all n.

Durbin-Levinson algorithm V

- The mean squared errors:
  Let $v_n = v_{n-1}[1 - \phi_{nn}^2]$ be true for $n = k$, then

$$
\begin{aligned}
v_{k+1} : \quad &= \gamma(0) - \boldsymbol{\phi}_{\mathbf{k+1}}{}' \boldsymbol{\gamma}_{\mathbf{k+1}} \\
&= \gamma(0) - [\phi_{k+1,1}, \ldots, \phi_{k+1,k}]\boldsymbol{\gamma}_{\mathbf{k}} - \phi_{k+1,k+1}\gamma(k+1) \\
&= \gamma(0) - \left(\boldsymbol{\phi}_{\mathbf{k}}{}' - \phi_{k+1,k+1} \boldsymbol{\phi}_{\mathbf{k}}^{(\mathbf{r})\prime} \right) \boldsymbol{\gamma}_{\mathbf{k}} - \phi_{k+1,k+1}\gamma(k+1), \text{[by (2)]} \\
&= v_k - \phi_{k+1,k+1} \left(\gamma(k+1) - \boldsymbol{\phi}_{\mathbf{k}}^{(\mathbf{r})\prime} \boldsymbol{\gamma}_{\mathbf{k}} \right), \text{[by assumption]} \\
&= v_k - \phi_{k+1,k+1} \left(\phi_{k+1,k+1} v_k \right), \text{[by (1)]} \\
&= v_k \left(1 - \phi_{k+1,k+1}^2 \right)
\end{aligned}
$$

  Therefore, true for all n.

Innovations algorithm I

- Here, we consider the problem of predicting the values X_{n+h}, $h > 0$, of a stationary time series with known mean and autocovariance function in terms of the values of successive differences in prediction $\{X_n - X_n^{n-1}\}$, up to time n.
    - Forecasting as **MA** model

Innovations algorithm II

- One step Recursive Forecast (The Innovations Algorithm)
  - Set the predicting equation of time series at time $n+1$ depending on the previous n observations as follows

$$X_{n+1}^{n} = \sum_{j=1}^{n} \theta_{nj} \left(X_{n+1-j} - X_{n+1-j}^{n-j} \right) \text{ for } n = 1, 2, \ldots, \tag{4}$$

with $X_1^0 = 0$

Innovations algorithm III

- Compute the coefficients $\theta_{n1}, \ldots \theta_{nn}$ recursively from the following equations

$$v_0 = \gamma(0),$$

$$\theta_{n,n-k} = v_k^{-1}\left(\gamma(n-k) - \sum_{j=0}^{k-1} \theta_{k,k-j}\theta_{n,n-j}v_j\right), \text{ for } 0 \leq k < n$$

and

$$v_n = \gamma(0) - \sum_{j=0}^{n-1} \theta_{n,n-j}^2 v_j$$

- Remarks:
  - The one step prediction error, $U_n = X_n - X_n^{n-1}$ is named as innovation at time n
  - Innovations $U_1, U_2, \ldots, U_n$ are uncorrelated.

Innovations algorithm IV

- Proof
  - Innovations $X_1 - X_1^0, X_2 - X_2^1, \ldots, X_n - X_n^{n-1}$ are orthogonal by definition.
  - Taking the inner product on both sides of (4) with $X_{k+1} - X_{k+1}^k,\ 0 \leq k < n$ we have

$$< X_{n+1}^n, X_{k+1} - X_{k+1}^k > \quad = \quad \theta_{n,n-k}\nu_k$$

  - Since $\left(X_{n+1} - X_{n+1}^n\right) \perp \left(X_{k+1} - X_{k+1}^k\right)$, for $k = 0, \ldots, n-1$, thus

$$\begin{aligned} < X_{n+1}, X_{k+1} - X_{k+1}^k > &= < X_{n+1}^n, X_{k+1} - X_{k+1}^k > \\ &= \theta_{n,n-k}\nu_k. \end{aligned} \tag{5}$$

Innovations algorithm V

- Hence,

$$
\begin{aligned}
\theta_{n,n-k} &= \nu_k^{-1} < X_{n+1}, X_{k+1} - X_{k+1}^k > \\
&= \nu_k^{-1}\left(\gamma(n-k) - \sum_{j=1}^{k} \theta_{k,j} < X_{n+1}, X_{k+1-j} - X_{k+1-j}^{k-j} >\right),
\end{aligned}
$$

by replacing n by k in (4)

$$
\begin{aligned}
&= \nu_k^{-1}\left(\gamma(n-k) - \sum_{j=0}^{k-1} \theta_{k,j+1} < X_{n+1}, X_{k-j} - X_{k-j}^{k-j-1} >\right), \\
&= \nu_k^{-1}\left(\gamma(n-k) - \sum_{j=0}^{k-1} \theta_{k,k-j} < X_{n+1}, X_{j+1} - X_{j+1}^{j} >\right),
\end{aligned}
$$

by replacing $(k-j)$ by $(j+1)$

$$
= \nu_k^{-1}\left(\gamma(n-k) - \sum_{j=0}^{k-1} \theta_{k,k-j}\theta_{n,n-j}\nu_j\right), \text{ by (5).}
$$

Innovations algorithm VI

- The mean squared errors:

$$
\begin{aligned}
v_n : &= ||X_{n+1} - X_{n+1}^n||^2 \\
&= ||X_{n+1}||^2 - ||X_{n+1}^n||^2 \\
&= \gamma(0) - \sum_{j=1}^{n} \theta_{nj}^2 \nu_{n-j}, \text{ [by (4)]} \\
&= \gamma(0) - \sum_{j=0}^{n-1} \theta_{n,j+1}^2 \nu_{n-j-1} \\
&= \gamma(0) - \sum_{j=0}^{n-1} \theta_{n,n-j}^2 \nu_j, \text{ by replacing } (n-j-1) \text{ by } j.
\end{aligned}
$$

Innovations algorithm VII

- $h-$ step Recursive Forecast
  - The predicting equation of time series at time $n+h$ depending on the n observations is as follows

$$X_{n+h}^{n} = \sum_{j=h}^{n+h-1} \theta_{n+h-1,j} \left(X_{n+h-j} - X_{n+h-j}^{n+h-1-j} \right)$$

  for $n = 1, 2, \ldots,$ with $X_1^0 = 0$

  - Corresponding mean squared error

$$E(X_{n+h} - X_{n+h}^{n})^2 = \gamma(0) - \sum_{j=h}^{n+h-1} \theta_{n+h-1,j}^2 v_{n+h-1-j}$$

- Innovations algorithm can also help to forecast an ARMA(p, q) process.
- It works in two phases, here
  1. First:
     - Transform an causal ARMA process X_t, where

$$\phi(B)X_t = \theta(B)Z_t,\ Z_t \sim WN(0, \sigma^2)$$

to a MA process as

$$W_t = \begin{cases} \sigma^{-1}X_t, & t = 1, \dots, m \\ \sigma^{-1}\phi(B)X_t = \sigma^{-1}[X_t - \phi_{t-1}X_{t-1} - \cdots - \phi_{t-p}X_{t-p}], & t > m \end{cases}$$

where $m = \max(p, q)$

Innovations algorithm for Forecasting ARMA II

- Then apply the innovations algorithm to the process $\{W_t\}$ to obtain

$$W_{n+1}^{n} = \begin{cases} \sum_{j=1}^{n} \theta_{nj}(W_{n+1-j} - W_{n+1-j}^{n-j}), & n = 1, \ldots, m-1 \\ \sum_{j=1}^{q} \theta_{nj}(W_{n+1-j} - W_{n+1-j}^{n-j}), & n \geq m \end{cases},$$

where the coefficients θ_{nj} and the mean squared errors $r_n = E(W_{n+1} - W_{n+1}^{n})^2$ are found recursively from the innovations algorithm with γ_W defined as follows

$$\gamma_W(i-j) = \begin{cases} \sigma^{-1}\gamma_X(i-j), & 1 \leq i,j \leq m \\ \sum_{r=0}^{q} \theta_r \theta_{r+|i-j|}, & m < \min(i,j) \\ \sigma^{-1}\left[\gamma_X(i-j) - \sum_{r=1}^{p} \phi_r \gamma_X(r-|i-j|)\right], & \min(i,j) \leq m < \max(i,j) \leq 2m \\ 0, & \text{otherwise.} \end{cases}$$

2. Second:
   - Note that, by definition each $X_n, n \geq 1$, to be written as a linear combination of $W_j, 1 \leq j \leq n$, and vice-versa.
   - Therefore, the best linear predictor of any random variable Y in terms of $\{1, X_1, \ldots, X_n\}$ is the same as the best linear predictor of Y in terms of $\{1, W_1, \ldots, W_n\}$.
   - Thus,

$$W_t^{t-1} = \begin{cases} \sigma^{-1} X_t^{t-1}, & t = 1, \ldots, m \\ \sigma^{-1} \left[X_t^{t-1} - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p}\right], , & t > m \end{cases},$$

   - Also,

$$\sigma^{-1}\left(X_t - X_t^{t-1}\right) = \left(W_t - W_t^{t-1}\right), \text{ for } t \geq 1$$

Innovations algorithm for Forecasting ARMA IV

- Therefore, X_{n+1} can be predicted by

$$X_{n+1}^{n} = \begin{cases} \sum_{j=1}^{n} \theta_{nj} \left(X_{n+1-j} - X_{n+1-j}^{n-j} \right), & 1 \leq n < m \\ \phi_1 X_n + \cdots + \phi_p X_{n+1-p} + \sum_{j=1}^{q} \theta_{nj} \left(X_{n+1-j} - X_{n+1-j}^{n-j} \right), & n \geq m \end{cases}$$

  with MSE:

$$E\left(X_t - X_t^{t-1}\right)^2 = \sigma^2 E\left(W_t - W_t^{t-1}\right)^2 = \sigma^2 r_n$$

- Note that:
  - ϕ_is are known
  - while θ_{nj}s and r_ns are calculated by Innovation algorithm!
  - The one-step predictors $X_2^1, X_3^2, \ldots, X_{n+1}^n$ are calculated recursively.