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
2. ACF and PACF of Stationary Time Series
   - ACF of Stationary Time Series
   - PACF of Stationary Time Series

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.

Forecasting Stationary Time Series II

- Minimization yields
  - Normal Equations

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

and

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

Forecasting Stationary Time Series III

- 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

- 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

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

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

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

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

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

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

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

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

  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 $\theta_{n1}, \dots \theta_{nn}$ recursively from the following equations

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

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

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 $U_1, U_2, \dots, U_n$ 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

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

Innovations algorithm V

- Hence,

$$
\begin{aligned}
\theta_{n,n-k} &= \nu_k^{-1} < X_{n+1}, X_{k+1} - X_{k+1}^{k} > \\
&= \nu_k^{-1} \left(\gamma(n-k) - \sum_{j=1}^{k} \theta_{k,j} < X_{n+1}, X_{k+1-j} - X_{k+1-j}^{k-j} > \right), \\
&\quad \text{by replacing } n \text{ by } k \text{ in (4)} \\
&= \nu_k^{-1} \left(\gamma(n-k) - \sum_{j=0}^{k-1} \theta_{k,j+1} < X_{n+1}, X_{k-j} - X_{k-j}^{k-j-1} > \right), \\
&= \nu_k^{-1} \left(\gamma(n-k) - \sum_{j=0}^{k-1} \theta_{k,k-j} < X_{n+1}, X_{j+1} - X_{j+1}^{j} > \right), \\
&\quad \text{by replacing } (k-j) \text{ by } (j+1) \\
&= \nu_k^{-1} \left(\gamma(n-k) - \sum_{j=0}^{k-1} \theta_{k,k-j}\theta_{n,n-j}\nu_j \right), \text{ by (5).}
\end{aligned}
$$

- 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 for Forecasting ARMA I

- 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

       $$W_t = \begin{cases} \sigma^{-1}X_t, & t = 1, \ldots, m \\ \sigma^{-1}\phi(B)X_t = \sigma^{-1}[X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p}] & \\ \quad = \sigma^{-1}[Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}] & t > m \end{cases} \quad (6)$$

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

Innovations algorithm for Forecasting ARMA II

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

$$\hat{W}_{n+1}^{n} = \begin{cases} \sum_{j=1}^{n} \theta_{nj}(W_{n+1-j} - W_{n+1-j}^{n-j}), & n = 1, \dots, m-1 \\ \sum_{j=1}^{q} \theta_{nj}(W_{n+1-j} - W_{n+1-j}^{n-j}), & n \geq m \end{cases}, \quad (7)$$

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 (Use 6)

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

Innovations algorithm for Forecasting ARMA III

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 as same as the best linear predictor of Y in terms of $\{1, W_1, \ldots, W_n\}$.
   - Thus, (by 6)

$$W_t^{t-1} = \begin{cases} \sigma^{-1} X_t^{t-1}, & t = 1, \ldots, m \\ \sigma^{-1}\left[X_t^{t-1} - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p}\right],, & t > m \end{cases} \tag{8}$$

   - Also,

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

- Therefore, X_{n+1} can be predicted by (Using 7 in 8)

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

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.

ACF of Stationary Time Series I

- The autocorrelation function (ACF) of a stationary process, X_n, denoted as $\rho(h)$, for $h = 0, 1, 2, \ldots$, is defined as follows

$$\begin{aligned} \rho(h) &= cor(X_{n+h}, X_n) \\ &= \frac{E(X_{n+h}X_n)}{\sqrt{E[X_{n+h}^2]E[X_n^2]}} \end{aligned}$$

- Remarks
  - The autocorrelation matrix R_n is positive definite for all n, where

$$R_n = \begin{bmatrix} 1 & \rho(1) & \cdots & \rho(n-1) \\ \rho(1) & 1 & \cdots & \rho(n-2) \\ \vdots & \vdots & \vdots & \vdots \\ \rho(n-1) & \rho(n-2) & \cdots & 1 \end{bmatrix}$$

ACF of MA(q) process I

- q-order moving average or MA(q) process:

$$X_t = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}, t = 0, \pm 1, \ldots,$$

  where $\{Z_t\} \sim WN(0, \sigma^2)$ and $\theta_1, \ldots, \theta_q$ are real valued constants

  - ACF

$$\rho(h) = \begin{cases} \frac{1}{(1+\theta_1^2+\cdots+\theta_q^2)} \sum_{j=0}^{q-|h|} \theta_j \theta_{j+|h|}, & \text{if } |h| \leq q, \\ 0, & \text{if } |h| > q. \end{cases}$$

    - where θ_0 is defined to be 1
- ACF of $MA(q)$ process is **ZERO** for lags greater than q.
  - Cut-off to zero after lag q

ACF of AR(1) process I

- 1st order autoregressive or AR(1) process:

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

  where $\{Z_t\} \sim WN(0, \sigma^2)$ and $|\phi| < 1$
- The ACF of an $AR(1)$ process

$$\rho(h) = \phi^{|h|}$$

  - Tails off to zero

ACF of ARMA(1) process I

- 1st order ARMA or ARMA(1,1) process:

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

  where $\{Z_t\} \sim WN(0, \sigma^2), |\phi| < 1, Z_t$ is uncorrelated with X_s for each $s < t$ and $\phi + \theta \neq 0$
- The ACF of an $ARMA(1,1)$ process

$$\rho(h) = \begin{cases} 1, & \text{if } h = 0, \\ \frac{(\theta+\phi)(1-\phi^2)+(\theta+\phi)^2\phi}{(1-\phi^2)+(\theta+\phi)^2}, & \text{if } h = \pm 1 \\ \phi^{|h|-1}\rho(1), & \text{if } |h| \geq 2. \end{cases}$$

  - Tails off to zero

PACF of Stationary Time Series I

- The partial autocorrelation function (PACF) of a stationary process, X_n, denoted as $\alpha(h)$, for $h = 0, 1, 2, \ldots$ is defined as follows

$$\alpha(0) = 1, \alpha(1) = \rho(1)$$

and

$$\alpha(h) = cor(X_{n+h} - X_{n+h}^{n+1,n+h-1}, X_n - X_n^{n+1,n+h-1}), h \geq 2$$

- Remarks
  - The PACF, $\alpha(h)$, is the correlation between X_{n+h} and X_n with the linear dependence of $\{X_{n+1}, \dots, X_{n+h-1}\}$ on each, removed.
  - Both $(X_{n+h} - X_{n+h}^{n+1,n+h-1})$ and $(X_n - X_n^{n+1,n+h-1})$ are uncorrelated with $\{X_{n+1}, \dots, X_{n+h-1}\}$.
  - If the process X_n is Gaussian, then

$$\alpha(h) = cor(X_{n+h}, X_n | X_{n+1}, \dots, X_{n+h-1}).$$

    - That is, $\alpha(h)$ is the correlation coefficient between X_{n+h} and X_n in the bivariate distribution of (X_{n+h}, X_n) conditional on $\{X_{n+1}, \dots, X_{n+h-1}\}$.

PACF of Stationary Time Series III

- Theorem: $\alpha(h) = \phi_{hh}$
  - Recall, ϕ_{hh} is the last element of the vector $\phi_{\mathbf{h}}$ and $\Gamma_h \phi_{\mathbf{h}} = \gamma_{\mathbf{h}}$
- Proof:-
  - Forward MSE:

$$E\left[\left(X_{n+h} - \sum_{i=1}^{h-1} a_i X_{n+h-i}\right)^2\right]$$

  - Normal Equations

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

  - Solution:

$$\gamma_{\mathbf{h-1}} = \Gamma_{h-1} \mathbf{a}_{\mathbf{h-1}}$$

PACF of Stationary Time Series IV

- Backward MSE:

$$E\left[\left(X_n - \sum_{i=1}^{h-1} b_i X_{n+i}\right)^2\right]$$

- Normal Equations

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

- Solution

$$\gamma_{\mathbf{h-1}} = \Gamma_{h-1}\mathbf{b_{h-1}}$$

- Therefore,

$$\mathbf{a_{h-1}} = \mathbf{b_{h-1}} = \phi_{\mathbf{h-1}}$$

PACF of Stationary Time Series V

- As a result,

$$
\begin{aligned}
\alpha(h) &= cor(X_{n+h} - X_{n+h}^{n+h-1,n+1}, X_n - X_n^{n+1,n+h-1}) \\
&= \frac{E\left[\left(X_{n+h} - \phi'_{\mathbf{h-1}}\mathbf{X_{n+h-1,n+1}}\right)\left(X_n - \phi'_{\mathbf{h-1}}\mathbf{X_{n+1,n+h-1}}\right)\right]}{\sqrt{E\left[\left(X_{n+h} - \phi'_{\mathbf{h-1}}\mathbf{X_{n+h-1,n+1}}\right)^2\right] E\left[\left(X_n - \phi'_{\mathbf{h-1}}\mathbf{X_{n+1,n+h-1}}\right)^2\right]}} \\
&= \frac{E\left[\left(X_{n+h} - \phi'_{\mathbf{h-1}}\mathbf{X_{n+h-1,n+1}}\right)\left(X_n - \phi'^{(\mathbf{r})}_{\mathbf{h-1}}\mathbf{X_{n+h-1,n+1}}\right)\right]}{\sqrt{E\left[\left(X_{n+h} - \phi'_{\mathbf{h-1}}\mathbf{X_{n+h-1,n+1}}\right)^2\right] E\left[\left(X_n - \phi'^{(\mathbf{r})}_{\mathbf{h-1}}\mathbf{X_{n+h-1,n+1}}\right)^2\right]}} \\
&= \frac{\gamma(h) - \phi'_{\mathbf{h-1}}\gamma^{(\mathbf{r})}_{\mathbf{h-1}} - \phi'^{(\mathbf{r})}_{\mathbf{h-1}}\gamma_{\mathbf{h-1}} + \phi'_{\mathbf{h-1}}\Gamma_{h-1}\phi^{(\mathbf{r})}_{\mathbf{h-1}}}{\sqrt{\left[\gamma(0) - \phi'_{\mathbf{h-1}}\Gamma_{h-1}\phi_{\mathbf{h-1}}\right]\left[\gamma(0) - \phi'^{(\mathbf{r})}_{\mathbf{h-1}}\Gamma_{h-1}\phi^{(\mathbf{r})}_{\mathbf{h-1}}\right]}} \\
&= \frac{\gamma(h) - \phi'_{\mathbf{h-1}}\gamma^{(\mathbf{r})}_{\mathbf{h-1}} - \phi'_{\mathbf{h-1}}\gamma^{(\mathbf{r})}_{\mathbf{h-1}} + \phi'_{\mathbf{h-1}}\gamma^{(\mathbf{r})}_{\mathbf{h-1}}}{\gamma(0) - \phi'_{\mathbf{h-1}}\Gamma_{h-1}\phi_{\mathbf{h-1}}} \\
&= \frac{\gamma(h) - \phi'_{\mathbf{h-1}}\gamma^{(\mathbf{r})}_{\mathbf{h-1}}}{v_{h-1}} = \phi_{hh}
\end{aligned}
$$

PACF of AR(p) process I

- p-order autoregressive or AR(p) process:

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + Z_t, t = 0, \pm 1, \ldots,$$

where $\{Z_t\} \sim WN(0, \sigma^2),\ Z_t$ is uncorrelated with X_s for each $s < t$ and all roots of the polynomial $(1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p)$ lie outside the unit circle.

PACF of AR(p) process II

- PACF of causal $AR(p)$
  - For $h \geq p$ the best linear predictor of X_{h+1} in terms of $1, X_1, \ldots, X_h$ is
$$X_{h+1} = \phi_1 X_h + \phi_2 X_{h-1} + \cdots + \phi_p X_{h+1-p}.$$
Since the coefficient ϕ_{hh} of X_1 is ϕ_p if $h = p$ and 0 if $h > p$, we conclude that the
$$\alpha(h) = \phi_p \text{ for } h = p$$
and
$$\alpha(h) = 0 \text{ for } h > p$$
- PACF of a causal $AR(p)$ process is **ZERO** for lags greater than p.
  - Cut-off to zero after lag p

PACF of MA(1) process I

- 1st order moving average or MA(1) process:

$$X_t = Z_t + \theta Z_{t-1}, \; t = 0, \pm 1, \ldots,$$

  where $\{Z_t\} \sim WN(0, \sigma^2)$ and θ is a real constant.

- The PACF of an $MA(1)$ process

$$\alpha(h) = \phi_{hh} = -(-\theta)^h / (1 + \theta^2 + \cdots + \theta^{2h}).$$

  - Tails off to zero

ACF & PACF of Stationary Time Series I

- Behavior of the ACF and PACF for Causal and Invertible ARMA Models

	AR(p)	MA(q)	ARMA(p, q)
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off