Time Series

Sudipta Das

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

Outline I

Stationary Process I

- Going beyond *i.i.d* stochastic process (time series)
- Stationary Process
  - Strict Stationary Process
  - Weak Stationary Process

Strict Stationary Process I

- $\{X_t\}$ is a strictly stationary process if

$$(X_1, X_2, \ldots, X_n)' \stackrel{D}{=} (X_{1+h}, X_{2+h} \ldots, X_{n+h})'$$

  for all all integers h and $n \geq 1$.

  - i.e.

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_{X_{1+h},X_{2+h}\ldots,X_{n+h}}(x_1, x_2 \ldots, x_n)$$

    for all all integers h and $n \geq 1$.

Strict Stationary Process II

- Properties of a Strictly Stationary Process $\{X_t\}$:
  - The random variables X_t are identically distributed
    - Not necessarily independent
  - An *i.i.d* sequence is also strictly stationary
  - $(X_t, X_{t+h})' \overset{D}{=} (X_1, X_{1+h})'$ for all all integers t and h.

(Weak) Stationary Process I

- $\{X_t\}$ is a (weakly) stationary process if all of the following three conditions hold
  1. $E[X_t^2] < \infty$
     - Finite second order moment
  2. $\mu_X(t)$ is independent of t,
     - where $\mu_X(t)$ is the **mean function** of $\{X_t\}$ and is defined as

$$\mu_X(t) = E(X_t)$$

  3. $\gamma(t+h, t)$ is independent of t for each h,
     - where $\gamma(t+h, t)$ is the **covariance function** of $\{X_t\}$ and is defined as

$$\begin{aligned}\gamma(t+h, t) &= \text{Cov}(X_{t+h}, X_t) \\ &= E\left[(X_{t+h} - \mu_X(t+h))\,(X_t - \mu_X(t))\right]\end{aligned}$$

(Weak) Stationary Process II

- Remarks:
  - If $\{X_t\}$ is strictly stationary and $E[X_t^2] < \infty$ for all t, then $\{X_t\}$ is also weakly stationary
  - In this course, whenever we use the term stationary we shall mean weakly stationary, unless we specifically indicate otherwise.
  - We use the term covariance function with reference to a stationary time series $\{X_t\}$ we shall mean the function γ_X of one variable defined by

$$\gamma_X(h) := \gamma_X(h, 0) = \gamma_X(t + h, t)$$

  - The function $\gamma_X(\cdot)$ will be referred to as the **autocovariance function** (ACVF) of X_t and $\gamma_X(h)$ as its value at *lag* h.

(Weak) Stationary Process III

- Formally, the **autocovariance function** (ACVF) of $\{X_t\}$ at lag h is defined as

$$\gamma_X(h) = E\left[(X_{t+h} - \mu_X)(X_t - \mu_X)\right]$$

  - Note that

$$\gamma_X(0) \geq 0$$

    and

$$\gamma_X(h) = \gamma_X(-h)$$

- The **autocorrelation function** (ACF) of $\{X_t\}$ at lag h is defined as

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}$$

- Basic Properties of $\rho(\cdot)$:
  - $|\rho(h)| \leq 1$
  - $\rho(\cdot)$ is even, i.e., $\rho(h) = \rho(-h)$ for all h.

(Weak) Stationary Process: Examples I

- Some Elementary Stationary processes
  1. iid noise: $\{X_t\} \sim IID(0, \sigma^2)$, with $E(X_t^2) = \sigma^2 < \infty$
     - ACVF:
       $$\gamma_X(t+h, t) = \gamma_X(h) = \begin{cases} \sigma^2, & \text{if } h = 0, \\ 0 & \text{if } h \neq 0. \end{cases}$$
     - ACF
       $$\rho_X(t+h, t) = \rho_X(h) = \begin{cases} 1, & \text{if } h = 0, \\ 0, & \text{if } h \neq 0. \end{cases}$$

(Weak) Stationary Process: Examples II

2. White Noise: $\{X_t\} \sim WN(0, \sigma^2)$
   - It's a sequence of uncorrelated random variables, each with zero mean and variance σ^2,
   - ACVF:

$$\gamma_X(h) = \begin{cases} \sigma^2, & \text{if } h = 0, \\ 0 & \text{if } h \neq 0. \end{cases}$$

   - ACF

$$\rho_X(h) = \begin{cases} 1, & \text{if } h = 0, \\ 0, & \text{if } h \neq 0. \end{cases}$$

(Weak) Stationary Process: Examples III

Note that every $IID(0, \sigma^2)$, sequence is $WN(0, \sigma^2)$ but not conversely, like the following
Let $\{Z_t\}$ be iid$\sim N(0, 1)$ noise and define

$$X_t = \begin{cases} Z_t, & \text{if } t \text{ is even,} \\ (Z_{t-1}^2 - 1)/\sqrt{2} & \text{if } t \text{ is odd.} \end{cases}$$

Here, that $\{X_t\}$ is WN$(0, 1)$ but not iid$(0, 1)$ noise.

(Weak) Stationary Process: Examples IV

3. First-order moving average or MA(1) process:

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

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

- $EX_t = 0$
- ACVF

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

- ACF

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

(Weak) Stationary Process: Examples V

4. First-order autoregressive or AR(1) process:

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

where $\{Z_t\} \sim WN(0, \sigma^2), |\phi| < 1$ and Z_t is uncorrelated with X_s for each $s < t$

- $EX_t = 0$
- ACVF

$$\gamma(h) = \phi^{|h|} \frac{\sigma^2}{1 - \phi^2}$$

- ACF

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

(Weak) Stationary Process: Examples VI

5. First-order ARMA or ARMA(1,1) process:

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

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

- $EX_t = 0$
- ACVF

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

(Weak) Stationary Process: Examples VII

- Is Random Walk a stationary process?

$$X_t = X_{t-1} + Z_t,$$

where $Z_t \sim IID(\mu, \sigma^2)$

(Weak) Stationary Process: Examples VIII

- NO
  - $\mu_X(t) = \mu t$
  - $Cov(X_m, X_n) = \min(m, n) \times \sigma^2$

(Weak) Stationary Process: Examples IX

- Three higher order stationary processes
  1. 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
  2. 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.

3. ARMA(p,q) process:

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}, t = 0, \pm 1, \ldots,$$

where $\{Z_t\} \sim WN(0, \sigma^2)$, with all roots of the polynomial $(1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p)$ lie outside the unit circle and the polynomials $(1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p)$ and $(1 + \theta_1 z + \ldots + \theta_q z^q)$ have no common factors.

Linear Process I

- Linear processes: It includes the class of autoregressive moving-average (ARMA) process,
- Definition: The process $\{X_t\}$ is a linear process if it has the representation

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j},$$

for all t, where $\{Z_t\} \sim WN(0, \sigma^2)$ and $\{\psi_j\}$ is a sequence of constants with $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$

Linear Process II

- Alternate representation by backward shift operator:

$$X_t = \Psi(B) Z_t,$$

  where $\Psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$

  - The operator $\Psi(B)$ can be thought of as a linear filter, which when applied to the white noise "input" series $\{Z_t\}$ produces the "output" $\{X_t\}$
- Note: every second-order stationary process is either a linear process or can be transformed to a linear process by subtracting a deterministic component.

Linear Process III

Remarks

1. The condition $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ (absolute summability) ensures that the infinite sum

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j},$$

converges (with probability one)

Linear Process IV

- Sketch of proof:-

  Let $X_t^n = \sum_{j=-n}^{n} \psi_j Z_{t-j}$, and for small $\epsilon > 0$, define

$$A_n(\epsilon) = \left\{ |X_t^n - X_t| > \epsilon \right\} = \left\{ \left| \sum_{|j|>n} \psi_j Z_{t-j} \right| > \epsilon \right\}.$$

  By Chebyshev's inequality,

$$P(A_n) \leq E\left[\left| \sum_{|j|>n} \psi_j Z_{t-j} \right|^2 \right] / \epsilon^2.$$

Linear Process V

Thus,

$$
\begin{aligned}
\sum_{n=1}^{\infty} P(A_n) &\leq \frac{1}{\epsilon^2}\sum_{n=1}^{\infty} E\left|\sum_{|j|>n}\psi_j Z_{t-j}\right|^2 \\
&< \sum_{n=1}^{\infty} E\left|\sum_{|j|>n}\psi_j Z_{t-j}\right|^2 \\
&= \sum_{n=1}^{\infty}\left[\sum_{|i|>n}\sum_{|k|>n}\psi_i\psi_k E\left(Z_{t-i}Z_{t-k}\right)\right] \\
&\leq \sum_{n=1}^{\infty}\left[\sum_{|i|>n}\sum_{|k|>n}|\psi_i||\psi_k|\left|E\left(Z_{t-i}Z_{t-k}\right)\right|\right] \text{ ; triangular inequality} \\
&\leq \sum_{n=1}^{\infty}\left[\sum_{|i|>n}\sum_{|k|>n}|\psi_i||\psi_k|\left|E\left(Z_{t-i}^2\right)^{1/2}\right|E\left|\left(Z_{t-k}^2\right)^{1/2}\right|\right] \text{ ; Cauchy-Schwarz inequality} \\
&\leq \sigma^2\sum_{n=1}^{\infty}\left[\sum_{|i|>n}\sum_{|k|>n}|\psi_i||\psi_k|\right] \text{ ; Stationarity} \\
&= \sigma^2\sum_{n=1}^{\infty}\left[\sum_{|j|>n}|\psi_j|\right]^2 < \infty; \text{ absolute summability}
\end{aligned}
$$

Linear Process VI

Therefore, by Borel Cantelli Lemma

$$\sum_{n=1}^{\infty} P(A_n) < \infty \Rightarrow P\left(A^{(S)}\right) = 0, \text{ where } A^{(S)} = \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} A_m = \limsup_n A_n$$

- Event $A^{(S)}$ is called the lim sup event of the infinite sequence $\{A_n\}$.
- Event $A^{(S)}$ occurs if and only if for all $n \geq 1$, there exists an $m \geq n$ such that A_m occurs,
- equivalently, Event $A^{(S)}$ occurs if and only if infinitely many of the A_n occur.

By definition of limit, $\omega \in \left\{\lim_n X_n = X\right\}$ if and only if for all $u \geq 1$ there exists $n \geq 1$ such that for every $m \geq n$, $|X_m(\omega) - X(\omega)| \leq \frac{1}{u}$. Equivalently, it holds if and only if

$$\omega \in \cap_{u=1}^{\infty} \cup_{n=1}^{\infty} \cap_{m=n}^{\infty} \left[A_m\left(\frac{1}{u}\right)\right]^c = \left(\cup_{u=1}^{\infty} \limsup_n A_n\left(\frac{1}{u}\right)\right)^c.$$

Thus,

$$\begin{aligned} P\left(\omega : \lim_n X_t^n(\omega) = X_t(\omega)\right) = P\left(\left(\cup_{u=1}^{\infty} \limsup_n A_n(1/u)\right)^c\right) &= 1 - P\left(\cup_{u=1}^{\infty} \limsup_n A_n(1/u)\right) \\ &\geq 1 - \sum_{u=1}^{\infty} P\left(\limsup_n A_n(1/u)\right) = 1 \end{aligned}$$

Hence, $X_t^n \overset{a.s.}{\to} X_t$.

Linear Process VII

2. The condition $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ ensures that $\sum_{j=-\infty}^{\infty} \psi_j^2 < \infty$ therefore, the series $\sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$ converges in mean square to X_t, i.e.,

$$X_t^n = \sum_{j=-n}^{n} \psi_j Z_{t-j} \overset{m.s.}{\to} X_t$$

3. In generally, let $\{Y_t\}$ be a stationary process with mean 0 and covariance function γ_Y. If $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$, then the process

$$X_t = \Psi(B) Y_t = \sum_{j=-\infty}^{\infty} \psi_j Y_{t-j},$$

Linear Process VIII

is also stationary with mean 0 and autocovariance function as

$$\begin{aligned}
\gamma_X(h) &= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k E[Y_{t-j} Y_{t+h-k}] \\
&= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_Y(h-k+j) \\
&= \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+h} \sigma^2, \text{ (if } X_t \text{ is linear)}
\end{aligned}$$

Linear Process IX

4. The filters of the form $\alpha(B) = \sum_{j=-\infty}^{\infty} \alpha_j B^j$ and $\beta(B) = \sum_{j=-\infty}^{\infty} \beta_j B^j$ with absolutely summable coefficients can be applied successively to a stationary series $\{Y_t\}$ to generate a new stationary series

$$\begin{aligned} W_t &= \sum_{j=-\infty}^{\infty} \alpha_j \left(\sum_{k=-\infty}^{\infty} \beta_k Y_{(t-j)-k} \right) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \alpha_j \beta_k Y_{(t-j)-k} \\ &= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \alpha_{j-k} \beta_k Y_{t-j}, \text{replacing } j \text{ by } j-k \\ &= \sum_{j=-\infty}^{\infty} \psi_j Y_{t-j} \end{aligned}$$

- Therefore, $\psi_j = \sum_{k=-\infty}^{\infty} \alpha_{j-k}\beta_k = \sum_{k=-\infty}^{\infty} \alpha_k \beta_{j-k}$
- Alternate form $W_t = \alpha(B)\beta(B)Y_t = \beta(B)\alpha(B)Y_t = \psi(B)Y_t$

Linear Process X

- Forms of (stable) linear process:
  - Causal: A linear process $\{X_t\}$ is causal if X_t can be expressed in terms of the current and past values $Z_s, s \leq t$,

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}.$$

  - Invertible: A linear process $\{X_t\}$ is invertible if Z_t can be expressed in terms of the current and past values $X_s, s \leq t$,

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}.$$

Revisiting ARMA Proces I

- Let X_t be an ARMA(p,q) process

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

  where $\{Z_t\} \sim WN(0, \sigma^2)$, and the polynomials $(1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p)$ and $(1 + \theta_1 z + \ldots + \theta_q z^q)$ have no common factors.

- Representing X_t as linear process

$$X_t = (1 - \phi_1 B - \cdots - \phi_p B^p)^{-1}(1 + \theta_1 B + \cdots + \theta_q B^q) Z_t$$

Revisiting ARMA Proces II

- Condition for stability of X_t:
  The coefficients of linear process expression of X_t (i.e. $X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$) are absolutely summable.
- Equivalent Condition:

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0 \text{ for all } |z| = 1$$

  - No roots of $\phi(z)$ on the unit circle

Revisiting ARMA Proces III

- Example: AR(1)

$$X_t - \phi X_{t-1} = Z_t$$

  - if $|\phi| < 1$, $X_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j}$ is stable
  - if $|\phi| > 1$, $X_t = \sum_{j=1}^{\infty} \phi^{-j} Z_{t+j}$ is stable

Revisiting ARMA Proces IV

- Condition for causality of X_t:
  Process X_t can be expressed in terms of the current and past values $Z_s, s \leq t$, (i.e., $X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$)
- Equivalent Condition

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0 \text{ for all } |z| < 1$$

  - No roots of $\phi(z)$ inside the unit circle

Revisiting ARMA Proces V

- Example: AR(1)

$$X_t - \phi X_{t-1} = Z_t$$

  - if $|\phi| < 1$, $X_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j}$ is stable and causal
  - if $|\phi| > 1$, $X_t = \sum_{j=1}^{\infty} \phi^{-j} Z_{t+j}$ is stable but non-causal

Revisiting ARMA Proces VI

- Condition for invertibility of X_t:
  Process Z_t can be expressed in terms of the current and past values $X_s, s \leq t$, (i.e., $Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}$)
- Equivalent Condition

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q \neq 0 \text{ forall } |z| < 1$$

  - No roots $\theta(z)$ inside the unit circle

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

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

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

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

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

$$\Gamma_n \mathbf{a_n} = \gamma_\mathbf{n}(1)$$

$$\begin{bmatrix} 1 & \phi & \phi^2 & \ldots & \phi^{n-1} \\ \phi & 1 & \phi & \ldots & \phi^{n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \phi^{n-1} & \phi^{n-2} & \phi^{n-3} & \ldots & 1 \end{bmatrix} \mathbf{a_n} = \begin{bmatrix} \phi \\ \phi^2 \\ \vdots \\ \phi^n \end{bmatrix}$$

- 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

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

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

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

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

$$v_n := E(X_{n+1} - X_{n+1}^{n})^2 = \gamma(0) - \phi_{\mathbf{n}}{}'\gamma_{\mathbf{n}}$$

  - Again Determination of $\phi_{\mathbf{n}}$ 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

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

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

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

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 $\Gamma_n \boldsymbol{\phi_n} = \boldsymbol{\gamma_n}$ be true for $n = k$, then

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

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

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

  Therefore, true for all n.

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:

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

  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:

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

  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

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

PACF of Stationary Time Series II

- 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

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

- Therefore,

$$\mathbf{a}_{\mathbf{h-1}} = \mathbf{b}_{\mathbf{h-1}} = \phi_{\mathbf{h-1}}$$

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:

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

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