Lecture 4

Stationary processes and Linear Processes

Readings: CC08 Chapter 4.1 - 4.3; BD16 Chapter 1.4, 1.6, 2.2; SS17 Ch 1.5-1.6

MATH 8090 Time Series Analysis
Week 4

Whitney Huang
Clemson University

Agenda

1. **Estimation of Mean and Autocovariance Function**
2. **Testing Temporal Dependence**
3. **Linear Processes**
4. **MA(q) and AR(p) Processes**

The Sampling Distribution of $\bar{\eta}$

Let

$$v_T = \sum_{h=-(T-1)}^{(T-1)} \left(1 - \frac{|h|}{T}\right) \gamma(h)$$

- If $\{\eta_t\}$ is Gaussian we have

$$\sqrt{T}(\bar{\eta} - \mu) \sim \mathrm{N}(0, v_T)$$

- The result above is approximate for many non-Gaussian time series
- In practice we also need to estimate $\gamma(h)$ from the data

Confidence Intervals for μ

- If $\gamma(h) \to 0$ as $h \to \infty$ then

$$v = \lim_{T\to\infty} v_T = \sum_{h=-\infty}^{\infty} \gamma(h) \text{ exists.}$$

- Further, if $\{\eta_t\}$ is Gaussian and

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty,$$

then an approximate large-sample 95% CI for μ is given by

$$\left[\bar{\eta} - 1.96\sqrt{\frac{v}{T}}, \bar{\eta} + 1.96\sqrt{\frac{v}{T}}\right]$$

Strategies for Estimating v

- Parametric:
  - Assume a parametric model $\gamma_{\boldsymbol{\theta}}(\cdot)$, and calculate

$$\hat{v} = \sum_{h=-\infty}^{\infty} \gamma_{\hat{\boldsymbol{\theta}}}(h)$$

    based on the ACVF for that model
  - The standard error, v, will depend on the parameters $\boldsymbol{\theta}$ of the parametric model
- Nonparametric:
  - Estimate v by

$$\hat{v} = \sum_{h=-\infty}^{\infty} \hat{\gamma}(h),$$

    where $\hat{\gamma}(\cdot)$ is an nonparametric estimate of ACVF

Examples of Parametric Forms for v

- i.i.d. Gaussian Noise: $v = \gamma(0) = \sigma^2 \Rightarrow$ CI reduces to the classical case:

$$\left[\bar{\eta} - 1.96\sqrt{\frac{\sigma^2}{T}}, \bar{\eta} + 1.96\sqrt{\frac{\sigma^2}{T}}\right]$$

- MA(1) process: We have

$$\begin{aligned} v = \sum_{h=-\infty}^{\infty} \gamma(h) &= \gamma(-1) + \gamma(0) + \gamma(1) \\ &= \gamma(0) + 2\gamma(1) \\ &= \sigma^2(1 + \theta^2 + 2\theta) = \sigma^2(1+\theta)^2 \end{aligned}$$

- **Exercise**: Show for an AR(1) process we have

$$v = \frac{\sigma^2}{(1-\phi)^2}$$

An Estimator of $\gamma(\cdot)$

Goal: Want to estimate

$$\gamma(h) = \mathbb{C}\mathrm{ov}(\eta_t, \eta_{t+h}) = \mathbb{E}\left[(\eta_t - \mu)(\eta_{t+h} - \mu)\right]$$

using data $\{\eta_t\}_{t=1}^{T}$

- For $|h| < T$, consider $\hat{\gamma}(h) = \frac{1}{T}\sum_{t=1}^{T-|h|}(\eta_t - \bar{\eta})(\eta_{t+|h|} - \bar{\eta})$. We call $\hat{\gamma}(h)$ the sample ACVF
- The sample ACVF is a biased estimator of $\gamma(h)$, but, it is used as the **standard** estimate of $\gamma(h)$
- $\hat{\gamma}(h)$ are even and non-negative definite

The Sample Autocorrelation Function

- The sample autocorrelation function (ACF) is defined for $|h| < T$ by

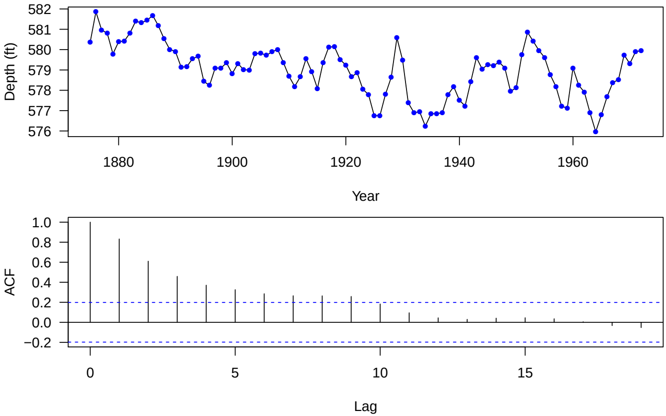
$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}.$$

- **Rule of thumb**: Box and Jenkins (1976) recommend using $\hat{\rho}(h)$ and $\hat{\gamma}(h)$ only for $\frac{|h|}{T} \leq \frac{1}{4}$ and $T \geq 50$

- This is because estimates $\hat{\rho}(h)$ and $\hat{\gamma}(h)$ are unstable for large $|h|$ as there will be no enough data points going into the estimator

Calculating the Sample ACF in R

- We use `acf` function to calculate the sample ACF
- Lake Huron Example



Asymptotic Distribution of the Sample ACF [Bartlett, 1946]

Let $\{\eta_t\}$ be a stationary process we suppose that the ACF

$$\boldsymbol{\rho} = (\rho(1), \rho(2), \cdots, \rho(k))^T$$

is estimated by

$$\hat{\boldsymbol{\rho}} = (\hat{\rho}(1), \hat{\rho}(2), \cdots, \hat{\rho}(k))^T$$

- For large T

$$\hat{\boldsymbol{\rho}} \dot{\sim} \mathrm{N}_k(\boldsymbol{\rho}, \frac{1}{T} W),$$

where N_k is the K-variate normal distribution and W is an $k \times k$ covariance matrix with (i, j) element defined by

$$w_{ij} = \sum_{k=1}^{\infty} a_{ik} a_{jk},$$

where $a_{ik} = \rho(k+i) + \rho(k-i) - 2\rho(k)\rho(i)$

Using the ACF as a Test for i.i.d. Noise

When $\{\eta_t\}$ is an i.i.d. process with finite variance, Bartlett's result simplifies for each $h \neq 0$

$$\hat{\rho}(h) \dot{\sim} \mathrm{N}(0, \frac{1}{T}).$$

This suggests a diagnostic for i.i.d. noise:

1. Plot the lag h versus the sample ACF $\hat{\rho}(h)$
2. Draw two horizontal lines at $\pm\frac{1.96}{\sqrt{T}}$ (blue dashed lines in `R`)
3. About 95% of the $\{\hat{\rho}(h) : h = 1, 2, 3, \cdots\}$ should be within the lines if we have i.i.d. noise

The Portmanteau Test [Box and Pierce, 1970] for i.i.d. Noise

Suppose we wish to test:

$$H_0 : \{\eta_1, \eta_2, \cdots, \eta_T\} \text{ is an i.i.d. noise sequence}$$
$$H_1 : H_0 \text{ is false}$$

- Under H_0,

$$\hat{\rho}(h) \dot{\sim} \mathrm{N}(0, \frac{1}{T}) \stackrel{d}{=} \frac{1}{\sqrt{T}}\mathrm{N}(0,1)$$

- Hence

$$Q = T\sum_{i=1}^{k} \hat{\rho}^2(h) \dot{\sim} \chi^2_{df=k}$$

- We reject H_0 if $Q > \chi^2_k(1-\alpha)$, the $1-\alpha$ quatile of the chi-squared distribution with k degrees of freedom

Ljung-Box Test [Ljung and Box, 1978]

Ljung and Box [1978] showed that

$$Q_{LB} = T(T-2)\sum_{h=1}^{k}\frac{\hat{\rho}^2(h)}{T-h} \mathrel{\dot{\sim}} \chi^2_k.$$

The Ljung-Box test can be more powerful than the Portmanteau test

Both the Portmanteau Test (aka Box-Pierce test) and Ljung-Box test can be carried out in `R` using the function `Box.test`

Examples in R

```
> Box.test(rnorm(100), 20)

        Box-Pierce test

data:  rnorm(100)
X-squared = 12.197, df = 20, p-value = 0.9091

> Box.test(LakeHuron, 20)

        Box-Pierce test

data:  LakeHuron
X-squared = 182.43, df = 20, p-value < 2.2e-16

> Box.test(LakeHuron, 20, type = "Ljung")

        Box-Ljung test

data:  LakeHuron
X-squared = 192.6, df = 20, p-value < 2.2e-16
```

Linear Processes

- A time series $\{\eta_t\}$ is a linear process with mean μ if we can write it as

$$\eta_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j Z_j, \quad \forall t,$$

  where μ is a real-valued constant, $\{Z_t\}$ is a $\mathrm{WN}(0, \sigma^2)$ process and $\{\psi_j\}$ is a set of absolutely summable constants[1]

- Absolute summability of the constants guarantees that the infinite sum converges

[1] A set of real-valued constants $\{\psi_j : j \in \mathbb{Z}\}$ is absolutely summable if $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$

Example: Moving Average Process of Order q, MA(q)

Let $\{Z_t\}$ be a $\mathrm{WN}(0, \sigma^2)$ process. For an integer $q > 0$ and constants $\theta_1, \cdots, \theta_q$ with $\theta_q \neq 0$, define

$$\begin{aligned} \eta_t &= Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \\ &= \theta_0 Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \\ &= \sum_{j=0}^{q} \theta_j Z_{t-j}, \end{aligned}$$

where we let $\theta_0 = 1$

$\{\eta_t\}$ is known as the moving average process of order q, or the MA(q) process, and, by definition, is a linear process

Defining Linear Processes with Backward Shifts

- Recall the backward shift operator, B, is defined by $B\eta_t = \eta_{t-1}$
- We can represent a linear process using the backward shift operator as $\eta_t = \mu + \psi(B)Z_t$, where we let $\psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$
- Example: we can write a mean zero MA(1) process as

$$\eta_t = \mu + \psi(B)Z_t,$$

where $\mu = 0$ and $\psi(B) = 1 + \theta B$

Linear Filtering Preserves Stationarity

- Let $\{Y_t\}$ be a time series and $\{\psi_j\}$ be a set of absolutely summable constants that does not depend on time
- **Definition**: A linear time invariant filtering of $\{Y_t\}$ with coefficients $\{\psi_j\}$ that do not depend on time is defined by

$$X_t = \psi(B)Y_t$$

- **Theorem**: Suppose $\{Y_t\}$ is a zero mean stationary series with ACVF $\gamma_Y(\cdot)$. Then $\{X_t\}$ is a zero mean stationary process with ACVF

$$\gamma_X(h) = \sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\psi_j\psi_k\gamma_Y(j-k+h)$$

Example: The MA(q) Process is Stationary

By the filtering preserves stationarity result, the MA(q) process is a stationary process with mean zero and ACVF

$$\gamma(h) = \sigma^2 \sum_{j=0}^{q} \theta_j \theta_{j+h}$$

Example: The MA(q) Process is Stationary

By the filtering preserves stationarity result, the MA(q) process is a stationary process with mean zero and ACVF

$$\gamma(h) = \sigma^2 \sum_{j=0}^{q} \theta_j \theta_{j+h}$$

$$\begin{aligned}
\gamma(h) &= \sum_{j=0}^{q} \sum_{k=0}^{q} \theta_j \theta_k \gamma_Z(j-k+h) \\
&= \sigma^2 \sum_{j=0}^{q} \sum_{k=0}^{q} \theta_j \theta_k \mathbb{1}(k = j+h) \\
&= \sigma^2 \sum_{j=0}^{q} \theta_j \theta_{j+h}
\end{aligned}$$

Processes with a Correlation that Cuts Off

- A time series η_t is q-correlated if

$$\eta_t \text{ and } \eta_s \text{ are uncorrelated } \forall |t-s| > q,$$

i.e., $\mathbb{C}\mathrm{ov}(\eta_t, \eta_s) = 0, \forall |t-s| > q$

- A time series $\{\eta_t\}$ is q-dependent if

$$\eta_t \text{ and } \eta_s \text{ are independent } \forall |t-s| > q.$$

- **Theorem**: if $\{\eta_t\}$ is a stationary q-correlated time series with zero mean, then it can be always be represented as an MA(q) process

The autoregressive process of order p, AR(p)

- This process is attributed to George Udny Yule. The AR(1) process has also been called the Markov process
- Let $\{Z_t\}$ be a $\mathrm{WN}(0, \sigma^2)$ process and let $\{\phi_1, \cdots, \phi_p\}$ be a set of constants for some integer $p > 0$ with $\phi_p \neq 0$
- The AR(p) process is defined to be the solution to the equation

$$\eta_t = \sum_{j=1}^{p} \phi_j \eta_{t-j} + Z_t \Rightarrow \underbrace{\eta_t - \sum_{j=1}^{p} \eta_{t-j}}_{\phi(B)\eta_t} = Z_t,$$

where we let $\phi(B) = 1 - \sum_{j=1}^{p} \phi_j B^j$

A Stationary Solution for AR(1)

- We want the solution to the AR equation to yield a stationary process. Let's first consider AR(1). We will demonstrate that **a stationary solution exists for** $|\phi_1| < 1$.

- We first write

$$
\begin{aligned}
\eta_t = \phi_1 \eta_{t-1} + Z_t &= \phi_1(\phi_1 \eta_{t-2} + Z_{t-1}) + Z_t \\
&= \phi_1^2 \eta_{t-2} + \phi_1 Z_{t-1} + Z_t \\
&\vdots \\
&= \phi_1^k \eta_{t-k} + \sum_{j=0}^{k-1} \phi_1^j Z_{t-j} \\
&\vdots \\
&= \sum_{j=0}^{\infty} \phi_1^j Z_{t-j}
\end{aligned}
$$

AR(1) Example Cont'd

- Now let $\psi_j = \phi_1^j$. We then have

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

Using the fact that, for $|a| < 0$, $\sum_{j=0}^{\infty} a^j = \frac{1}{1-a}$, the sequence $\{\psi_j\}$ is absolutely summable

- Thus, since $\{\eta_t\}$ is a linear process, it follows by the filtering preserves stationarity result that $\{\eta_t\}$ is a zero mean stationary process with ACVF

$$\begin{aligned}\gamma(h) &= \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h} \\ &= \sigma^2 \sum_{j=0}^{\infty} \phi_1^j \phi_1^{j+h} \\ &= \sigma^2 \phi^h \sum_{j=0}^{\infty} (\phi_1^2)^j\end{aligned}$$

AR(1) Example Cont'd

Now $|\phi_1| < 1$ implies that $|\phi_1^2| < 1$ and therefore we have

$$\gamma(h) = \frac{\sigma^2 \phi_1^h}{1 - \phi_1^2}$$

When $|\phi_1| \geq 1$

- No stationary solutions exist for $|\phi_1| = 1$
- When $|\phi_1| > 1$, dividing by ϕ_1 for both sides we get

$$\phi_1^{-1} \eta_t = \eta_{t-1} + \phi_1^{-1} Z_t$$
$$\Rightarrow \eta_{t-1} = \phi_1^{-1} \eta_t - \phi_1^{-1} Z_t$$

  A linear combination of **future** Z_t's $\Rightarrow$ we have a stationary solution, but, η_t depends on future $\{Z_t\}$'s–This process is said to be not causal

- If we assume that η_s and Z_t are uncorrelated for each $t > s$, $|\phi_1| < 1$ is the only stationary solution to the AR equation

The Autoregressive Operator

- AR(1) process

$$\eta_t = \phi_1 \eta_{t-1} + Z_t \Rightarrow (1 - \phi_1 B)\eta_t = Z_t \Rightarrow \eta_t = (1 - \phi_1 B)^{-1} Z_t$$

- Recall $\sum_{j=0}^{\infty} a^j = \frac{1}{1-a} = (1-a)^{-1}$. We have

$$\eta_t = \sum_{j=0}^{\infty} (\phi_1 B)^j Z_t = \sum_{j=0}^{\infty} \phi_1^j B^j Z_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j}$$

- Here $1 - \phi_1 B$ is the AR characteristic polynomial

The Second-Order Autoregressive Process

Now consider the series satisfying

$$\eta_t = \phi_1 \eta_{t-1} + \phi_2 \eta_{t-2} + Z_t,$$

where, again, we assume that Z_t is independent of $\eta_{t-1}, \eta_{t-2}, \cdots$

- The AR characteristic polynomial is

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2$$

- The corresponding AR characteristic equation is

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 = 0$$

Stationarity of the AR(2) Process

- A stationary solution exists if and only if the roots of the AR characteristic equation exceed 1 in absolute value
- For the AR(2) the roots of the quadratic characteristic equation are

$$\frac{\phi_1 \pm \sqrt{\phi_1^2 - 4\phi_2}}{-2\phi_2}$$

These roots exceed 1 in absolute value if

$$\phi_1 + \phi_2 < 1, \quad \phi_2 - \phi_1 < 1, \quad \text{and } |\phi_2| < 1$$

- We say that the roots should lie outside the unit circle in the complex plane. This statement will generalize to the AR(p) case

The Autocorrelation Function for the AR(2) Process

- Yule-Walker equations:

$$
\begin{aligned}
\eta_t &= \phi_1 \eta_{t-1} + \phi_2 \eta_2 + Z_t \\
\Rightarrow \eta_t \eta_{t-h} &= \phi_1 \eta_{t-1}\eta_{t-h} + \phi_2 \eta_{t-2}\eta_{t-h} + Z_t \eta_{t-h} \\
\Rightarrow \gamma(h) &= \phi_1 \gamma(h-1) + \phi_2 \gamma(h-2) \\
\Rightarrow \rho(h) &= \phi_1 \rho(h-1) + \phi_2 \rho(h-2),
\end{aligned}
$$

$h = 1, 2, \cdots$

- Setting $h = 1$, we have $\rho(1) = \phi_1 \underbrace{\rho(0)}_{=1} + \phi_2 \underbrace{\rho(-1)}_{=\rho(1)} \Rightarrow \rho(1) = \frac{\phi_1}{1-\phi_2}$
- $\rho(2) = \phi_1 \rho(1) + \phi_2 \rho(0) = \frac{\phi_2(1-\phi_2)+\phi_1^2}{1-\phi_2}$

The Variance for the AR(2) Model

Taking the variance of both sides of AR(2) equations:

$$\eta_t = \phi_1 \eta_{t-1} + \phi_2 \eta_{t-2} + Z_t,$$

yields

$$\begin{aligned}
\gamma(0) &= (\phi_1^2 + \phi_2^2)\gamma(0) + 2\phi_1\phi_2\gamma(1) + \sigma^2 \\
&= \frac{(1-\phi_2)\sigma^2}{(1-\phi_2)(1-\phi_1^2-\phi_2^2) - 2\phi_2\phi^2} \\
&= \left(\frac{1-\phi_2}{1+\phi_2}\right)\frac{\sigma^2}{(1-\phi_2)^2-\phi_1^2}
\end{aligned}$$

The General Autoregressive Process

Consider now the pth-order autoregressive model:

$$\eta_t = \phi_1\eta_{t-1} + \phi_2\eta_{t-2} + \cdots + \phi_p\eta_{t-p} + Z_t$$

- AR characteristic polynomial:
  $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$.
  AR characteristic equation: $1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p = 0$

- Yule-Walker equations:

$$\begin{aligned}
\rho(1) &= \phi_1 + \phi_2\rho(1) + \cdots + \phi_p\rho(p-1) \\
\rho(2) &= \phi_1\rho(1) + \phi_2 + \cdots + \phi_p\rho(p-2) \\
&\vdots \\
\rho(p) &= \phi_1\rho(p-1) + \phi_2\rho(p-2) + \cdots + \phi_p
\end{aligned}$$

- Variance:

$$\begin{aligned}
\gamma(0) &= \phi_1\gamma(1) + \phi_2\gamma(2) + \cdots + \phi_p\gamma(p) + \sigma^2 \\
&= \frac{\sigma^2}{1 - \phi_1\rho(1) - \cdots - \phi_p\rho(p)}
\end{aligned}$$