Lecture 4

Model-Free Estimation of Stationary Means and Covariances

Readings: CC08 Chapter 4.3, 3.2, 3.6; BD16 Chapter 2.4; SS17 Chapter 1.5

MATH 8090 Time Series Analysis
Week 4

Whitney Huang
Clemson University

Agenda

1. The Autoregressive Process

2. Model-Free Estimation of Stationary Means and Covariances

3. Testing Temporal Dependence

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

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

$\Rightarrow$ **This is another way to show that AR(1) is a linear process**

- Here $1 - \phi_1 B$ is the AR characteristic polynomial, which can be used to check whether the process is stationary and causal

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:

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

$h = 1, 2, \cdots$

The Autocorrelation Function for the AR(2) Process

- Yule-Walker equations:

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

$h = 1, 2, \cdots$

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

The Autocorrelation Function for the AR(2) Process

- Yule-Walker equations:

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

$h = 1, 2, \cdots$

- 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

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

The General Autoregressive Processes

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:

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

Estimating the Mean of Stationary Processes

Let $\{\eta_t\}$ be stationary with mean μ and ACVF $\gamma(s,t) = \gamma(s-t)$

- A natural estimator of μ is the sample mean

$$\bar{\eta} = \frac{1}{T}\sum_{t=1}^{T}\eta_t.$$

$\bar{\eta}$ is an unbiased estimator of μ, i.e.

- To conduct inference for μ, we need the standard error:

$$\begin{aligned}\mathrm{Var}(\bar{\eta}) &= \frac{1}{T^2}\mathrm{Var}\left(\sum_{i=1}^{T}\eta_t\right)\\ &= \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\mathrm{Cov}(\eta_s,\eta_t)\\ &= \frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\gamma(s-t)\end{aligned}$$

- **Exercise**: Show

$$\mathrm{Var}(\bar{\eta}) = \frac{1}{T}\sum_{h=-(T-1)}^{T-1}\left(1-\frac{|h|}{T}\right)\gamma(h)$$

AR(1) Example

Suppose $\{\eta_1, \eta_2, \eta_3\}$ is an AR(1) process with $|\phi| < 1$ and innovation variance σ^2. Show that the variance of $\bar{\eta}$ is $\frac{\sigma^2}{9(1-\phi^2)}(3 + 4\phi + 2\phi^2)$

Solution:

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 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

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. We will return to this later

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

- 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 depends 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: Estimate the autocovariance function (ACVF)

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

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

- For $|h| < T$, define the sample ACVF:

$$\hat{\gamma}(h) = \frac{1}{T} \sum_{t=1}^{T-|h|} (\eta_t - \bar{\eta})(\eta_{t+|h|} - \bar{\eta}).$$

An Estimator of $\gamma(\cdot)$

Goal: Estimate the autocovariance function (ACVF)

$$\gamma(h) = \mathrm{Cov}(\eta_t, \eta_{t+h}) = \mathbb{E}\big[(\eta_t - \mu)(\eta_{t+h} - \mu)\big],$$

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

- For $|h| < T$, define the sample ACVF:

$$\hat{\gamma}(h) = \frac{1}{T}\sum_{t=1}^{T-|h|}(\eta_t - \bar{\eta})(\eta_{t+|h|} - \bar{\eta}).$$

- **Properties:**

An Estimator of $\gamma(\cdot)$

Goal: Estimate the autocovariance function (ACVF)

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

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

- For $|h| < T$, define the sample ACVF:

$$\hat{\gamma}(h) = \frac{1}{T} \sum_{t=1}^{T-|h|} (\eta_t - \bar{\eta})(\eta_{t+|h|} - \bar{\eta}).$$

- **Properties:**
  - $\hat{\gamma}(h)$ is a biased estimator of $\gamma(h)$, but it is the **standard** estimator used in practice

An Estimator of $\gamma(\cdot)$

Goal: Estimate the autocovariance function (ACVF)

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

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

- For $|h| < T$, define the sample ACVF:

$$\hat{\gamma}(h) = \frac{1}{T} \sum_{t=1}^{T-|h|} (\eta_t - \bar{\eta})(\eta_{t+|h|} - \bar{\eta}).$$

- **Properties:**
  - $\hat{\gamma}(h)$ is a biased estimator of $\gamma(h)$, but it is the **standard** estimator used in practice
  - The collection $\{\hat{\gamma}(h)\}$ is **symmetric** (i.e., $\hat{\gamma}(h) = \hat{\gamma}(-h)$, $\forall h$)

An Estimator of $\gamma(\cdot)$

Goal: Estimate the autocovariance function (ACVF)

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

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

- For $|h| < T$, define the sample ACVF:

$$\hat{\gamma}(h) = \frac{1}{T}\sum_{t=1}^{T-|h|}(\eta_t - \bar{\eta})(\eta_{t+|h|} - \bar{\eta}).$$

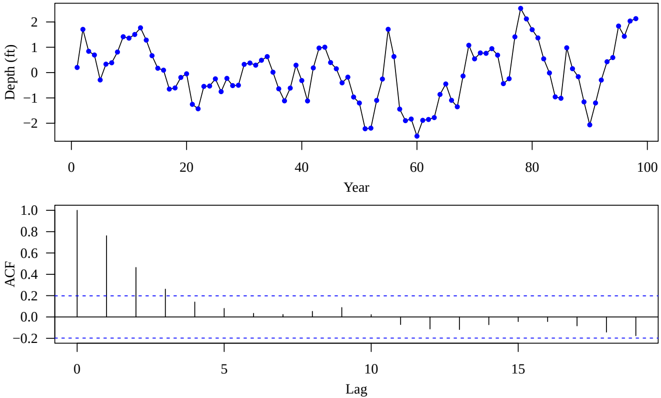
- **Properties:**
  - $\hat{\gamma}(h)$ is a biased estimator of $\gamma(h)$, but it is the **standard** estimator used in practice
  - The collection $\{\hat{\gamma}(h)\}$ is **symmetric** (i.e., $\hat{\gamma}(h) = \hat{\gamma}(-h)$, $\forall h$)
  - It is non-negative definite; that is, for all vectors $\mathbf{a}$, $\mathbf{a}^T\Sigma, \mathbf{a} \geq 0$, where Σ is the Toeplitz covariance matrix whose (i,j)-th element is $\gamma(i-j)$

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

$$w_{ij} = \sum_{h=1}^{\infty} a_{ih} a_{jh}, \quad 1 \le i \le k, \quad 1 \le j \le k$$

where $a_{ih} = \rho(h+i) + \rho(h-i) - 2\rho(h)\rho(i)$

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:

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

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

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`

New Weighted Portmanteau Tests [Fisher & Gallagher, 2012]

$$Q_W = n \sum_{h=1}^{k} w_h \, \hat{\rho}(h)^2, \qquad w_h = \frac{k+1-h}{k} \quad \text{(linearly decreasing)}.$$

- **Why:** Better finite-sample behavior and power than Box-Pierce/Ljung-Box, especially for larger k or underfit ARMA
- **Null:** $Q_W \rightsquigarrow$ weighted sum of χ_1^2; use a Gamma approximation for p-values
- **Nonlinearity:** Apply to squared residuals to detect ARCH/GARCH
- **R:** `WeightedPortTest::Weighted.Box.test(resid, m=20[, sqrd = TRUE])`