Time Series

Author: Wenxiao Yang

Institute: Haas School of Business, University of California Berkeley

Date: 2024

All models are wrong, but some are useful.

Contents

Chapter 1 Time Series Analysis

1.1 Goals and Terminology

Goals and Challenge

Data in time series is denoted by

$$\{\underbrace{y_t}_{n\times 1} : 1 \le t \le T\}$$

Assumption *Each y_t is the realization of some random vector Y_t.*

The **objective** is to provide data-based answers to questions about the distribution of $\{Y_t : 1 \le t \le T\}$.

The **challenge** we face is $Y_1, Y_2, \ldots, Y_T$ are *not necessarily independent*. Time series analysis gives the models and methods that can accommodate dependence.

Terminology

Some terminologies we need to know:

Definition 1.1 (Stochastic Process)

A **stochastic process** is a collection $\{Y_t : t \in \mathcal{T}\}$ of random variables/vectors (defined on the same probability space).

1. $\{Y_t : t \in \mathcal{T}\}$ is **discrete time process** if $\mathcal{T} = \{1, \ldots, T\}$ or $\mathcal{T} = \mathbb{N} = \{1, 2, \ldots\}$ or $\mathcal{T} = \mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$.
2. $\{Y_t : t \in \mathcal{T}\}$ is **continuous time process** if $\mathcal{T} = [0, 1]$ or $\mathcal{T} = \mathbb{R}_+$ or $\mathcal{T} = \mathbb{R}$.

Observed data Y_t is a realization of a discrete time process with $\mathcal{T} = \{1, \ldots, T\}$.

Definition 1.2 (Strictly Stationary (Discrete and Scalar Process))

A scalar[a] process $\{Y_t : t \in \mathbb{Z}\}$ is **strictly stationary** *if and only if*

$$(Y_t, \ldots, Y_{t+k}) \underbrace{\sim}_{\text{"is distributed as"}} (Y_0, \ldots, Y_k)\,, \ \forall t \in \mathbb{Z}, k \ge 0$$

[a] i.e., Y_t is 1×1

Note

1. *If $Y_t \sim i.i.d.$, then $\{Y_t : t \in \mathbb{Z}\}$ is strictly stationary.*
2. *If $\{Y_t : t \in \mathbb{Z}\}$ is strictly stationary, then Y_t are identically distributed (i.e., "marginal stationary").*

Example 1.1 Strictly Stationary and Dependent

A constant process that $\ldots = Y_{-1} = Y_0 = Y_1 = \ldots$ *is strictly stationary.*

All these above hold for strictly stationary vector process.

Lemma 1.1 (Property of Strictly Stationary)

If $\{Y_t : t \in \mathbb{Z}\}$ is strictly stationary with $\mathbb{E}[Y_t^2] < \infty$ $(\forall t)$, then

1. Same Expectation:

$$\mathbb{E}[Y_t] = \mu, \ \forall t \text{ (for some constant } \mu) \tag{*}$$

2. Covariance only depends on time length:

$$\text{Cov}(Y_t, Y_{t-j}) = \gamma(j), \ \forall t, j \text{ (for some function } \gamma(\cdot)) \tag{**}$$

   Note $\gamma(0) = \text{Var}(Y_t), \forall t$.

A subset of strictly stationary processes that has second moment (i.e., $\mathbb{E}[Y_t^2] < \infty$) can be defined as **covariance stationary**.

Definition 1.3 (Covariance Stationary)

A process $\{Y_t : t \in \mathbb{Z}\}$ is **covariance stationary** *iff* $\mathbb{E}[Y_t^2] < \infty$ $(\forall t)$ and it satisfies (*) and (**).

Note *Not every strictly stationary process is covariance stationary. (e.g., if it does not have second moment).*

Definition 1.4 (Autocovariance and Autocorrelation Functions)

$\gamma(\cdot)$ in (**) is called **autocovariance function** of $\{Y_t : t \in \mathbb{Z}\}$.

The **autocorrelation function** is

$$\rho(j) = \text{Corr}(Y_t, Y_{t-j}) = \frac{Cov(Y_t, Y_{t-j})}{\sqrt{\text{Var}(Y_t)\text{Var}(Y_{t-j})}} = \frac{\gamma(j)}{\gamma(0)}.$$

Lemma 1.2

The autocovariance function satisfies the following properties:

1. $\gamma(\cdot)$ is **even** i.e., $\gamma(j) = \gamma(-j)$.
2. $\gamma(\cdot)$ is **positive semi-definite** (psd) i.e., for any $n \in \mathbb{N}$ and any $a_1, \ldots, a_n$,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \gamma(i-j) = \text{Var}(\sum_{i=1}^{n} a_i Y_i) \geq 0$$

1.2 Moving-Average Process

Definition 1.5 (White Noise)

A process $\{\epsilon_t : t \in \mathbb{Z}\}$ is a **white noise** process iff it is covariance stationary with $\mathbb{E}[\epsilon_t] = 0$ and

$$\text{Cov}(\epsilon_t, \epsilon_{t-j}) = \begin{cases} \sigma^2, & \text{if } j = 0 \\ 0, & \text{otherwise} \end{cases}$$

We use *notation* $\epsilon_t \sim \text{WN}(0, \sigma^2)$.

Note

1. *If* $\epsilon_t \sim \text{i.i.d.}(0, \sigma^2)$*, then* $\{\epsilon_t : t \in \mathbb{Z}\}$ *is white noise, i.e.,* $\epsilon_t \sim \text{WN}(0, \sigma^2)$.
2. *Gauss-Markov theorem assumes WN errors.*
3. *WN terms are used as "building blocks": often a variable can be generated as*

$$Y_t = h(\epsilon_t, \epsilon_{t-1}, \ldots) \text{ for some function } h(\cdot) \text{ and some } \epsilon_t \sim \text{WN}(0, \sigma^2).$$

1.2.1 Moving-Average Process

Definition 1.6 (Finite Moving-Average Process)

1. First-order moving average process: $Y_t \sim \text{MA}(1)$ iff

$$Y_t = \mu + \epsilon_t + \theta\epsilon_{t-1}$$

where $\epsilon_t \sim \text{WN}(0, \sigma^2)$.

Claim 1.1

$\{Y_t\}$ is covariance stationary: $\mathbb{E}[Y_t] = \mu$ and its autocovariance function is

$$\gamma(j) := \text{Cov}(Y_t, Y_{t-j}) = \begin{cases} (1+\theta^2)\sigma^2, & j = 0 \\ \theta\sigma^2, & j = 1 \\ 0, & j \geq 2 \end{cases}$$

2. $Y_t \sim \text{MA}(q)$ (for some $q \in \mathbb{N}$) iff

$$Y_t = \mu + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$$

where $\epsilon_t \sim \text{WN}(0, \sigma^2)$.

Claim 1.2

$\{Y_t\}$ is covariance stationary: $\mathbb{E}[Y_t] = \mu$ and its autocovariance function is

$$\gamma(j) := \text{Cov}(Y_t, Y_{t-j}) = \begin{cases} \left(\sum_{i=0}^{q-j} \theta_i \theta_{i+j}\right) \sigma^2, & j \leq q \\ 0, & j \geq q+1 \end{cases}$$

where $\theta_0 = 1$.

Definition 1.7 (Infinite Moving-Average Process)

$Y_t \sim \text{MA}(\infty)$ iff

$$Y_t = \mu + \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}$$

where $\epsilon_t \sim \text{WN}(0, \sigma^2)$ and $\underline{\sum_{i=0}^{\infty} \psi_i^2 < \infty}$

1.2.2 Conditions for Infinite Moving-Average Process

Note *Conjecture:*

1. $\{Y_t\}$ *is covariance stationary;*
2. $\mathbb{E}[Y_t] = \mu$ *and*
3. *its autocovariance function is*

$$\gamma(j) := \text{Cov}(Y_t, Y_{t-j}) = (\sum_{i=0}^{\infty} \psi_i \psi_{i+j}) \sigma^2, \forall j \geq 0.$$

The necessary condition to make these conjectures correct is

$$\begin{aligned} \mathbb{E}[Y_t^2] &= (\mathbb{E}[Y_t])^2 + \Gamma(0) \\ &= \mu^2 + (\sum_{i=0}^{\infty} \psi_i^2) \sigma^2 < \infty \\ \Leftrightarrow \sum_{i=0}^{\infty} \psi_i^2 &< \infty \end{aligned}$$

Claim 1.3

With the `right' definition of ``$\sum_{i=0}^{\infty}$'', the conjecture is correct.

Remark

1. If $X_0, X_1, \ldots$ are i.i.d. with $X_0 = 0$, then $\sum_{i=0}^{\infty} X_i$ denote $\lim_{n\to\infty} \sum_{i=0}^{n} X_i$ (assuming the limit exists).
2. $\exists$ various models of stochastic convergence.
3. There: convergence in mean square.

Definition 1.8 (Stochastic Convergence in Mean Square)

If $X_0, X_1, \dots$ are random (with $\mathbb{E}[X_i^2] < \infty, \forall i$), then $\sum_{i=0}^{\infty} X_i$ denotes any S such that $\lim_{n\to\infty} \mathbb{E}[(S - \sum_{i=0}^{n} X_i)^2] = 0$.

Lemma 1.3

The properties of the S are

1. S is ``essentially unique.''
2. $\mathbb{E}[S] = \sum_{i=0}^{\infty} \mathbb{E}[X_i] = \lim_{n\to\infty} \sum_{i=0}^{n} \mathbb{E}[X_i]$
3. $\mathrm{Var}[S] = \dots = \lim_{n\to\infty} \mathrm{Var}[\sum_{i=0}^{n} X_i]$
4. (Higher order moments of S are similar) $\cdots$

Theorem 1.1 (Cauchy Criterion)

$\sum_{i=0}^{\infty} X_i$ exists iff

$$\lim_{n\to\infty} \sup_{m>n} \mathbb{E}[(S_m - S_n)^2] = 0,$$

where $S_n = \sum_{i=0}^{n} X_i$.

In the $MA(\infty)$ context: The condition that can make

$$\lim_{n\to\infty} \sup_{m>n} \mathbb{E}[(Y_{t,m} - Y_{t,n})^2] = 0$$

where $Y_{t,n} = \mu + \sum_{i=0}^{n} \psi_i \epsilon_{t-i}$.

This condition is given as: If $m > n$,

$$\begin{aligned}
Y_{t,m} - Y_{t,n} &= \sum_{i=n+1}^{m} \psi_i \epsilon_{t-i} \\
\Rightarrow \mathbb{E}[(Y_{t,m} - Y_{t,n})^2] &= \mathbb{E}\left[\left(\sum_{i=n+1}^{m} \psi_i \epsilon_{t-i}\right)^2\right] = \left(\sum_{i=n+1}^{m} \psi_i^2\right)\sigma^2 \\
\Rightarrow \sup_{m>n} \mathbb{E}[(Y_{t,m} - Y_{t,n})^2] &= \left(\sum_{i=n+1}^{\infty} \psi_i^2\right)\sigma^2 \\
\Rightarrow \lim_{n\to\infty} \sup_{m>n} \mathbb{E}[(Y_{t,m} - Y_{t,n})^2] &= \left(\lim_{n\to\infty} \sum_{i=n+1}^{\infty} \psi_i^2\right)\sigma^2
\end{aligned}$$

Thus,

$$\begin{aligned}
\lim_{n\to\infty} \sup_{m>n} \mathbb{E}[(Y_{t,m} - Y_{t,n})^2] = 0 \text{ iff } &\lim_{n\to\infty} \sum_{i=n+1}^{\infty} \psi_i^2 = 0 \\
\text{iff } &\sum_{i=0}^{\infty} \psi_i^2 < \infty
\end{aligned}$$

1.2.3 Remarks about $MA(\infty)$ models

1. $MA(\infty)$ models are useful in theoretical work.
2. The $MA(\infty)$ class is "large": Wold decomposition (theorem).
3. Parametric $MA(\infty)$ models are useful in inference.

1.3 Autoregressive Model (Special Case of $MA(\infty)$)

Autoregressive model is an example of well-defined $MA(\infty)$ model.

Example 1.2 (Autoregressive model)

Suppose

$$Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}, \ \forall t$$

where

- $\epsilon_t \sim \mathrm{WN}(0, \sigma^2)$;
- $\psi_i = \phi^i$ ($\forall i \geq 0$) for some $|\phi| < 1$.

Checking the condition: $\lim_{n\to\infty} \sum_{i=0}^{n} \psi_i^2 = \lim_{n\to\infty} \sum_{i=0}^{n} \phi^{2i} = \lim_{n\to\infty} \frac{1-\phi^{2(n+1)}}{1-\phi^2} = \frac{1}{1-\phi^2} < \infty$.

Lemma 1.4 (Property of $MA(\infty)$)

For $j \geq 0$, the autocovariance function is

$$\begin{aligned}\gamma(j) := \mathrm{Cov}(Y_t, Y_{t-j}) &= (\sum_{i=0}^{\infty} \psi_i \psi_{i+j})\sigma^2 = \phi^j (\sum_{i=0}^{\infty} \phi^{2i})\sigma^2 = \phi^j \frac{\sigma^2}{1-\phi^2} \\ &= \phi^j \gamma(0)\end{aligned}$$

Note

1. *$\gamma(j) \neq 0, \forall j$ if $\phi \neq 0$.*
2. *$\gamma(j) \propto \phi^j$ decays exponentially.*

Definition 1.9 (Alternative Representation of AR)

Alternatively, the AR model ca be represented as

$$Y_t = \phi Y_{t-1} + \epsilon_t, \ \forall t$$

Proof 1.1

$$Y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} = \epsilon_t + \sum_{i=1}^{\infty} \psi_i \epsilon_{t-i} = \epsilon_t + \sum_{i=0}^{\infty} \psi_{i+1} \epsilon_{t-i-1} = \epsilon_t + \phi \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i-1} = \epsilon_t + \phi Y_{t-1}$$

The natural estimator of ϕ (OLS) is

$$\hat{\phi} = \frac{\sum_{t=2}^{T} Y_{t-1}Y_t}{\sum_{t=2}^{T} Y_{t-1}^2}$$

Definition 1.10 (Model for Finite AR)

$$Y_t = \phi Y_{t-1} + \epsilon_t, \; 2 \leq t \leq T$$

where

- $\epsilon_t \sim \mathrm{WN}(0, \sigma^2)$;
- $|\phi| < 1$;
- $Y_1 = \sum_{i=0}^{\infty} \phi^i \epsilon_{1-i}$

More generally, consider an AR with a drift,

$$Y_t = c + \phi Y_{t-1} + \epsilon_t, \; \forall t$$

where $c = \mu(1 - \phi)$.

Definition 1.11 ($AR(1)$)

$\{Y_t : 1 \leq t \leq T\}$ is an **autoregreessive process** of order 1, $Y_t \sim AR(1)$, if

$$Y_t = c + \phi Y_{t-1} + \epsilon_t, \; 2 \leq t \leq T$$

where $\epsilon_t \sim \mathrm{WN}(0, \sigma^2)$.

Note *$|\phi| < 1$ is not assumed (yet) and $Y_1 = \mu + \sum_{i=0}^{\infty} \phi^i \epsilon_{1-i}$ is not assumed.*

We call the AR(1) model is **stable** iff $|\phi| < 1$.

- If $|\phi| < 1$ and $Y_1 = \mu + \sum_{i=0}^{\infty} \phi^i \epsilon_{1-i}$,

$$Y_t = \mu + \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i},$$

  where $\mu = \frac{c}{1-\phi}$.
- OLS "works" when $|\phi| < 1$.
- The $AR(1)$ model admits and $MA(\infty)$ solution

$$Y_t = \mu + \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}, \quad \sum_{i=0}^{\infty} \psi_i^2 < \infty$$

  iff $|\phi| < 1$.
- The AR(1) model admits a covariance stationary solution iff $|\phi| \neq 1$.

Note *Consider the case that $\phi > 1$, the intuition is*

$$Y_t = \phi Y_{t-1} + \epsilon_t \Leftrightarrow Y_{t-1} = \phi^{-1}(Y_t - \epsilon_t)$$

1.4 $AR(p)$

Definition 1.12

$\{Y_t : t \in \mathbb{N}\}$ is a p^{th}**-order autoregressive process**, $Y_t \sim AR(p)$, iff

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \epsilon_t,\ t \geq p+1$$

where $\epsilon_t \sim \text{WN}(0, \sigma^2)$.

Claim 1.4

OLS ``works" when the $AR(p)$ model is stable.

Vector Notation We write

$$Y_t = \beta' X_t + \epsilon_t,\ t \geq p+1$$

where $\beta = (c, \phi_1, \phi_2, \cdots, \phi_p)'$ and $X_t = (1, Y_{t-1}, Y_{t-2}, \cdots, Y_{t-p})'$.

Then the *OLS estimator* is given by

$$\hat{\beta} = (\sum_{t=p+1}^{T} X_t' X_t)^{-1} (\sum_{t=p+1}^{T} X_t' Y_t)$$

Lag Operator Notation For a time series $\{X_t\}$, let

$$LX_t = X_{t-1}$$
$$\vdots$$
$$L^k X_t = X_{t-k},\ \forall t \in \mathbb{Z}$$

which is called *lag operator notation*.

Then, in this notation, the $AR(p)$ model can be written as

$$\phi(L) Y_t = c + q_t,\ t \geq p+1$$

where $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$.

Definition 1.13 (Stability of $AR(p)$)

The $AR(p)$ model is **stable** if

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = 0 \Rightarrow |z| > 1$$

(All solutions are greater than 1).

- The $AR(p)$ model admits an $MA(\infty)$ solution

$$Y_t = \mu + \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}, \quad \sum_{i=0}^{\infty} \psi_i^2 < \infty$$

iff it is *stable*. The $MA(\infty)$ solution has

$$\mu = \frac{c}{1 - \phi_1 - \cdots - \phi_p} = \frac{c}{\phi(1)}$$

and (computable) ψ_i's satisfy

$$|\psi_i| \leq M\lambda^i, \ \forall i,$$

where $M < \infty$ and $|\lambda| < 1$.

1.5

$MA(q)$ **model in lag operator notation** :

$$\begin{aligned} Y_t &= \mu + \underbrace{\epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}}_{:=\theta(L)\epsilon_t} \\ &= \mu + \theta(L)\epsilon_t, \end{aligned}$$

where $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$.

Definition 1.14 (Invertibility of $MA(q)$)

The $MA(q)$ model is **invertible** if

$$\theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q = 0 \Rightarrow |z| > 1$$

(All solutions are greater than 1).

Note *If the $MA(q)$ model is invertible, then*

$$\epsilon_t = \pi(L)(Y_t - \mu),$$

where $\Pi(L) = \sum_{i=0}^{\infty} \pi_i L^i$ *with* $\sum_{i=0}^{\infty} |\pi_i| < \infty$.

Technicalities

- If $\sum_{i=0}^{\infty} |\pi_i| < \infty$, then $\sum_{i=0}^{\infty} \pi_i^2 < \infty$.
- If

$$|\pi_i| \leq M\lambda^i, \ \forall i \text{ (some } M < \infty \text{ and } |\lambda| < 1), \tag{*}$$

  then

$$\sum_{i=0}^{\infty} i^r |\pi_i|^s < \infty, \ \forall r \geq 0, s > 0$$

- Invertibility $\Rightarrow$ (*).
- If $X_0, X_1, \ldots$ are random variables with $\sup_i \mathbb{E}X_i^2 < \infty$, then $\sum_{i=0}^{\infty} \pi_i X_i$ exists (as a limit in mean

squared) if $\sum_{i=0}^{\infty} |\pi_i| < \infty$.

1.5.1

Lemma 1.5

If $\{Y_t\}$ is covariance stationary, then $\gamma(j) = 0$, $\forall j > q$ iff $Y_t \sim MA(q)$.

Question: Is there a "$q = \infty$" analog?

Example 1.3

Suppose $Y_t = Z \sim \mathcal{N}(0, 1), \forall t$. Then, $\text{Cov}(Y_t, Y_{t-1}) = 1, \forall j$.

1. It is not a $MA(\infty)$.
2. Y_t can be predicted without error using $\{Y_s : s \leq t - 1\}$.
3. Y_t is "deterministic".