ST304 Week 2

Kaixin Liu[1]

[1]PhD Student in Statistics, LSE

Jan 29, 2026

Table of Contents

Table of Contents

About me

Education:

- Second year PhD student.
- Supervision under Prof. Clifford Lam and Dr. Yunxiao Chen.
- Research interests in: High-dimensional time series analysis; Latent factor model with vector, matrix and tensor time series.
- Obtained BEcon in major of Economic Statistics from Beihang University (formerly known as Beijing University of Aeronautics and Astronautics) in 2023, and MSc in major of Statistics (Financial Statistics) from LSE in 2024.

How to reach out:

- Email Address: K.Liu31@lse.ac.uk
- Office: Columbia House 5.02
- Office Hour: 12pm – 1pm, Thursdays (excluding reading week)
- Homepage: https://www.lse.ac.uk/people/kaixin-liu

Table of Contents

Strong/Strict stationarity

Definition (Strong/Strict stationarity)

A stochastic process $(X_t)_{t\in T}$ is **strongly (strictly) stationary** if for any $k \in \mathbb{N}$, the joint distributions of $(X_{t_1}, \dots, X_{t_k})$ and $(X_{t_1+h}, \dots, X_{t_k+h})$ are identical whenever t_j, $t_j + h \in T$ for all j, i.e.

$$\begin{aligned} &\mathbb{P}(X_{t_1} \leq x_1, \dots, X_{t_k} \leq x_k) \\ =&\mathbb{P}(X_{t_1+h} \leq x_1, \dots, X_{t_k+h} \leq x_k), \quad \forall x_1, \dots, x_k \in \mathbb{R}. \end{aligned}$$

- $k = 1$: all X_t have the same marginal distribution $\Rightarrow$ if $\mathbb{E}(X_t)$ exists, it is constant in t.
- $k = 2$: (X_s, X_t) and (X_{s+h}, X_{t+h}) have the same bivariate distribution $\Rightarrow$ covariance depends only on lag.
- $\dots$

Table of Contents

Weak (second-order) stationarity

Definition (Weak stationarity)

A stochastic process $(X_t)_{t\in T}$ is **weakly (second-order) stationary** if $\mathbb{E}(X_t^2) < \infty$ for all $t \in T$ and

(i) $\mathbb{E}(X_t)$ is the same for all t (write this constant as μ);

(ii) $\mathrm{Cov}(X_s, X_t) = \mathrm{Cov}(X_{s+h}, X_{t+h})$ whenever $s, t, s+h, t+h \in T$.

- In this course, "stationary" usually means **weakly stationary**.
- i.i.d. processes are strongly stationary $\Rightarrow$ (if $\mathbb{E}X_t^2 < \infty$) also weakly stationary.

When weak $\Leftrightarrow$ strong: Gaussian processes

Definition (Gaussian process)

$(X_t)_{t \in T}$ is a **Gaussian process** if for any $m \in \mathbb{N}$, any $a_1, \ldots, a_m \in \mathbb{R}$ and any $t_1, \ldots, t_m \in T$, the linear combination $\sum_{j=1}^{m} a_j X_{t_j}$ is normally distributed. Equivalently, $(X_{t_1}, \ldots, X_{t_m})$ is jointly multivariate normal for any finite set of indices.

Key fact

For Gaussian processes, the joint distribution is fully determined by the mean vector and covariance matrix. Hence, **weak stationarity and strong stationarity are equivalent** for Gaussian processes.

White noise (WN)

Definition (White noise)

$(\varepsilon_t)_{t \in T}$ is **white noise** if it is weakly stationary with $\mathbb{E}(\varepsilon_t) = 0$ and

$$\gamma(h) = \text{Cov}(\varepsilon_t, \varepsilon_{t+h}) = \begin{cases} \sigma^2, & h = 0, \\ 0, & h \neq 0. \end{cases}$$

We write $\varepsilon_t \in WN(0, \sigma^2)$.

- White noise does *not* require independence unless explicitly stated i.i.d.

Table of Contents

Autocovariance and autocorrelation

Autocovariance function (ACVF)

For a weakly stationary process (X_t), define for lag $h = t - s$:

$$\gamma(h) := \mathrm{Cov}(X_s, X_t), \qquad h \in \mathbb{Z}.$$

This is well-defined because the covariance depends only on the lag.

Autocorrelation function (ACF)

$$\rho(h) := \mathrm{Corr}(X_s, X_t) = \frac{\gamma(h)}{\gamma(0)}, \qquad \rho(0) = 1.$$

Properties of ACVF and ACF

Let $(X_t)_{t \in T}$ be weakly stationary, with ACVF $\gamma(h)$ and ACF $\rho(h)$.

Basic properties (Prop. 1.8)

- $\gamma(0) = \mathrm{Var}(X_t) \in [0, \infty)$ and $\rho(0) = 1$.
- Symmetry: $\gamma(h) = \gamma(-h)$ and $\rho(h) = \rho(-h)$ for all h.
- Bound: $|\rho(h)| \leq 1$ (equivalently $|\gamma(h)| \leq \gamma(0)$) for all h.

Non-negative definiteness (Prop. 1.9)

For any $m \in \mathbb{N}$, $t_1, \ldots, t_m \in T$ and $a_1, \ldots, a_m \in \mathbb{R}$,

$$\sum_{j=1}^{m} \sum_{k=1}^{m} a_j \, \gamma(t_j - t_k) \, a_k \ \geq \ 0.$$

Equivalently, the covariance matrix with entries $\gamma(t_j - t_k)$ is positive semidefinite. The same holds for $\rho(\cdot)$.

Sample autocovariance and autocorrelation

Suppose we observe $X_1, \ldots, X_n$ from a weakly stationary process. Let $\bar{X}_n := \frac{1}{n} \sum_{t=1}^{n} X_t$.

Sample autocovariance function (ACVF)

For $|h| \leq n-1$,

$$\hat{\gamma}_n(h) := \frac{1}{n} \sum_{t=1}^{n-|h|} (X_t - \bar{X}_n)(X_{t+|h|} - \bar{X}_n),$$

and set $\hat{\gamma}_n(h) := 0$ otherwise.

Sample autocorrelation function (ACF)

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

Note $\hat{\gamma}_n(0) = \frac{1}{n} \sum_{t=1}^{n} (X_t - \bar{X}_n)^2$ (sample variance).

Estimators: unbiased vs practical choice

- If the true mean μ is known, then for $|h| \le n-1$,

$$\tilde{\gamma}_n^*(h) := \frac{1}{n-|h|} \sum_{t=1}^{n-|h|} (X_t - \mu)(X_{t+|h|} - \mu)$$

  is an **unbiased** estimator of $\gamma(h)$.
- In practice μ is unknown, so we replace it by $\bar{X}_n$:

$$\tilde{\gamma}_n(h) := \frac{1}{n-|h|} \sum_{t=1}^{n-|h|} (X_t - \bar{X}_n)(X_{t+|h|} - \bar{X}_n),$$

  which is generally **biased**.
- The notes often prefer using $\hat{\gamma}_n(h)$ with denominator n: despite bias, it often has **smaller MSE** and yields a **non-negative definite** ACVF.