Midterm Review
EE 226A

Ahmed Shakil

Contents

1 Tools and Tricks **2**

2 Limit Theorems and Modes of Convergence **3**
2.1 Asymptotic Behavior of the Empirical Mean 3

3 Time Series Analysis **3**
3.1 Second-Order Processes 3
3.2 Spectral Theory of Second-Order Processes 4
3.2.1 Fourier transform speedrun 4
3.3 Linear Estimation from WSS Observations 7
3.4 The Noncasual Wiener Filter 8
3.5 Linear Filtering of WSS process 8

1 Tools and Tricks

Remark 1.1: Convergence in second moment

Convergence in second moment implies convergence in distribution. (via hw 5.3 solution)

2 Limit Theorems and Modes of Convergence

2.1 Asymptotic Behavior of the Empirical Mean

Theorem 2.1: Strong Law of Large Numbers

Let $(X_n)_{n\geqslant 1}$ be i.i.d. integrable random variables. For $S_n := X_1 + \cdots + X_n,\ n \geqslant 1$, it holds that

$$\Pr\left\{\lim_{n\to\infty} \frac{1}{n} S_n = \mathbb{E}[X_1]\right\} = 1.$$

Remark 2.2

This conclusion continues to hold when (i) the X_n's are identically distributed and pairwise independent; and/or (ii) the expectation $\mathbb{E}[X_1]$ exists, but is not necessarily finite. The first follows from the proof given, the second is left as an exercise for the reader.

3 Time Series Analysis

3.1 Second-Order Processes

Definition 3.1: Second-Order Process

Let $X = (X_n)_{n\in\mathbb{Z}}$ be a stochastic process on a probability space $(\Omega, \mathcal{F}, P)$. The process X is said to be a (discrete-time) second-order process if it has finite second moments $\mathbb{E}|X_n|^2 < \infty$ for all $n \in \mathbb{Z}$. Since all X_n are elements of $L^2(\Omega, \mathcal{F}, P)$, it follows that second-order processes also form a vector space.

Example 3.2

Gaussian processes are second-order processes, and the collection of Gaussian processes is a subspace of second-order processes.

Definition 3.3: Second-Order Statistics

For second-order processes X and Y the second-order statistics are summarized by the mean function $\mu_X(n) := \mathbb{E}[X_n]$ and covariance function

$$R_{XY}(m, n) := \operatorname{Cov}(X_m, Y_n), \quad m, n \in \mathbb{Z}.$$

For second-order processes, the mean and covariance functions are finite everywhere.

Example 3.4

If X is a Gaussian process, then all finite-dimensional marginals are characterized by the functions μ_X and R_{XX}.

Definition 3.5: Wide Sense Stationary

A second-order process $X = (X_n)_{n\in\mathbb{Z}}$ is wide sense stationary (WSS), if $\mu_x(n) = \mu_X(0)$ for all n, and $R_{XX}(m,n)$ is a function of only the difference $(m-n)$. In this case, we often abbreviate $R_{XX}(m,n)$ as $R_{XX}(m-n)$ to denote parametrization of the covariance function by the difference $(m-n)$.

Remark 3.6

For WSS processes the covariance enjoys the follow symmetries:

$$R_{XX}(n, n+k) = R_{XX}(0,k) = R_{XX}(k,0) = R_{XX}(0,-k).$$

In our compact notation, $R_{XX}(k) = R_{XX}(-k)$, so that R_{XX} is a symmetric function of k.

Definition 3.7: Jointly Wide Senese Stationary

Processes $X = (X_n)_{n\in\mathbb{Z}}$ and $Y = (Y_n)_{n\in\mathbb{Z}}$ are jointly wide sense stationary (JWSS) if each are WSS and the covariance function $R_{XY}(m,n) = \mathrm{Cov}(X_m, Y_n)$ depends only on the difference $m-n$. In this case, we abbreviate $R_{XY}(m,n)$ as $R_{XY}(m-n)$.

Remark 3.8

Unlike R_{XX}, the function R_{XY} is not symmetric in its argument. However, if X and Y are JWSS, then we do have the following identities

$$R_{XY}(n+k, n) = \mathrm{Cov}(X_k, Y_0) = \mathrm{Cov}(X_0, Y_{-k}) = R_{XY}(k,0) = R_{XY}(0,-k).$$

In particular, noting the order of subscripts, we have $R_{XY}(k) = R_{YX}(-k)$.

3.2 Spectral Theory of Second-Order Processes

3.2.1 Fourier transform speedrun

To start, we note some info and results about Fourier transforms.

3.9: About l^p spaces

We write $x \in l^p(\mathbb{Z})$ if x is a real-valued sequence $(x(n))_{n\in\mathbb{Z}} \subset \mathbb{R}$ satisfying $\|x\|_p := \left(\sum_n |x(n)^p|\right)^{\frac{1}{p}} < \infty$, with $\|x\|_\infty := \sup_n |x(n)|$. For $1 \leqslant p \leqslant q \leqslant \infty$, $l^p(\mathbb{Z})$ is complete with respect to convergence in its norm $\|\cdot\|$, and $l^p(\mathbb{Z}) \subset l^q(\mathbb{Z})$ on account of $\|x\|_q \leqslant \|x\|_p$. The spaces $l^p(\mathbb{Z})$ are equal to the closure of $l^1(\mathbb{Z})$. Of particular note, $l^2(\mathbb{X})$ is a Hilbert space when equipped with the inner product

$$(x,y) \mapsto \sum_n x(n)y(n), \quad x,y \in l^2(\mathbb{Z}).$$

Definition 3.10: Discrete-time Fourier Transform

For a sequence $x \in l^1(\mathbb{Z})$, its discrete-time Fourier transform is defined as the complex-valued function

$$\hat{x}(\omega) = \sum_n x(n)e^{-i\omega n}, \quad \omega \in [-\pi, \pi).$$

Note that the mapping $x \mapsto \hat{x}$ is a linear transformation from $l^1(\mathbb{Z})$ to the function space

$$L^\infty([-\pi,\pi)) \coloneqq \left\{ f : [-\pi,\pi) \to \mathbb{C};\ \sup_\omega |f(\omega)| < \infty \right\}.$$

This leads to the Fourier inversion identity

$$x(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{x}(\omega)e^{i\omega n}\, d\omega = \sum_k x(k)\delta(n-k), \quad n \in \mathbb{Z}.$$

This formula holds if $\hat{x}$ is the Fourier transform of $x \in l^2(\mathbb{Z})$.

Theorem 3.11: Convolution theorem for Fourier transforms

If $x, y \in l^2(\mathbb{Z})$ and $\operatorname{ess\,sup}(|\hat{y}|) < \infty$, then their convolution $z = x * y$ is in $l^2(\mathbb{Z})$. In particular, all Fourier transforms exist and satisfy

$$\hat{z} = \hat{x}\hat{y}.$$

3.12: Parseval identity

If $x, y \in l^1(\mathbb{Z})$, we have the easily verified Parseval identity

$$\sum_n x(n)y(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{x}(\omega)\hat{y}^*(\omega)\, d\omega.$$

In particular, this implies that the mapping $x \mapsto \hat{x}$ is a linear isometry from $l^2(\mathbb{Z}) \cap l^1(\mathbb{Z})$ into $L^2([-\pi,\pi))$; i.e.,

$$\|x\|_2 = \|\hat{x}\|_{L^2} \text{ for all } x \in l^1(\mathbb{Z}),$$

where $\|\cdot\|_{L^2}$ denotes the norm on $L^2([-\pi,\pi))$ induced by its inner product.

Definition 3.13: Impulse response

If x is input to a LTI system with impulse response g, then the output sequence y is defined by the convolution

$$y(n) = (x * g)(n) = \sum_k x(n-k)g(k), \quad n \in \mathbb{Z},$$

provided the series converges.

Definition 3.14: Frequence response

An LTI system with impulse response $g \in l^2(\mathbb{Z})$ is equivalently characterized by its frequency response G, which is simply the Fourier transform of the impulse response:

$$G(\omega) = \sum_n g(n)e^{-i\omega n}, \quad \omega \in [-\pi,\pi).$$

Remark 3.15

By convolution theorem, if a finite-energy sequence x (i.e., $x \in l^2(\mathbb{Z})$) is input to a stable (i.e. BIBO stability) LTI system with impulse response g, the output y will also have finite energy, and is characterized by its Fourier transform

$$\hat{y} = G\hat{x},$$

where $\hat{x}$ denotes the Fourier transform of the input x.

Now back to section 5.2.

Definition 3.16: Energy Spectral Density

Given $x \in l^1(\mathbb{Z})$, we can define a sequence $a \in l^1(\mathbb{Z})$ via the self-convolution

$$a(n) = \sum_k x(k)x(n-k), \quad n \in \mathbb{Z}.$$

By the convolution theorem and time-reversal property of Fourier transforms, the discrete-time Fourier transform of a is equal to

$$\hat{a}(\omega) = \hat{x}(\omega)\hat{x}^*(\omega) = |\hat{x}(\omega)|^2 \geqslant 0.$$

The function $\hat{a}$ is called the energy spectral density of x, since it is a nonnegative function with the property that its integral over any subset of frequencies in $[-\pi, \pi)$ is equal to the energy of the sequence x restricted to those frequencies.

Definition 3.17: Power Spectral Density

The average energy normalized by time is called power. Assume we are working with a zero-mean WSS random process.

$$\frac{1}{2N+1}\mathbb{E}[A_N(\omega)] = \frac{1}{2N+1}\sum_{-N \leqslant m,n \leqslant N} \mathbb{E}[X_n X_m] e^{-i\omega(n-m)} = \sum_{k=-2N}^{2N} R_{XX}(k)e^{-i\omega k}\left(1 - \frac{|k|}{2N+1}\right).$$

Now, if $R_{XX} \in l^1(\mathbb{Z})$, then the limit as $N \to \infty$ exists on the right by dominated convergence, so that

$$S_{XX}(\omega) := \lim_{N\to\infty} \frac{1}{2N+1}\mathbb{E}[A_N(\omega)] = \sum_k R_{XX}(k)e^{-i\omega k}, \quad \omega \in [-\pi, \pi).$$

The function S_{XX} is called the power spectral density of the process X, and is a real, non-negative function. The definition of power spectral density can be extended to WSS processes X with $R_{XX} \in l^2(\mathbb{Z})$ using the mean-square convergence of the Fourier transform. In this case, S_{XX} continues to be real and non-negative.

Definition 3.18: Regular covariance

We say that X admits a regular covariance if: (i) $R_{XX} \in l^2(\mathbb{Z})$; and (ii) there exists $\lambda > 0$ such that the power spectral density satisfies

$$\lambda \leqslant \operatorname{ess\,inf}(S_{XX}) \leqslant \operatorname{ess\,sup}(S_{XX}) \leqslant \lambda^{-1}.$$

Definition 3.19: Cross-power spectrum

If X is a random variable with finite variance and $Y = (Y_n)_{n\in\mathbb{Z}}$ is zero-mean WSS process, then the cross-power spectrum is defined via the discrete-time Fourier transform

$$S_{YX}(\omega) \coloneqq \sum_n \mathbb{E}[XY_n]e^{-i\omega n}, \quad \omega \in [-\pi, \pi),$$

provided the series converges in a suitable sense (e.g., if $n \mapsto \mathbb{E}[XY_n]$ is in $l^2(\mathbb{Z})$, then series converges in the mean-square sense).
Note the order of subscripts. If $X = (X_n)_{n\in\mathbb{Z}}$ is JWSS with Y, we define

$$S_{YX}(\omega) \coloneqq \sum_n \mathbb{E}[X_0Y_n]e^{-i\omega n};$$

i.e., S_{YX} is the Fourier transform of R_{YX}, consistent with the definition of power spectral density. In this case, the quantity S_{XY} is also well-defined (as the Fourier transform of R_{XY}), and enjoys the conjugate symmetry $S_{XY} = S_{YX}^*$.

3.3 Linear Estimation from WSS Observations

Theorem 3.20

Fix $I \subset \mathbb{Z}$ and let $Y = (Y_n)_{n\in\mathbb{Z}}$ be a zero-mean WSS stationary process with regular covariance. For any zero-mean random variable X with finite variance, there exists $h \in l^2(\mathbb{Z})$ such that

$$\mathbb{L}[X|Y_I] = \sum_{n\in I} h(n)Y_n \quad \text{in } L^2.$$

Moreover, the sequence h is unique on the indices in I.

Theorem 3.21: Weiner-Hopf equations

Fix $I \subset \mathbb{Z}$ and let $Y = (Y_n)_{n\in\mathbb{Z}}$ be a zero-mean WSS stationary process with regular covariance. The sequence $h \in l^2(\mathbb{Z})$ defining the best linear estimator from the previous theorem uniquely solves the system of equations

$$\begin{aligned} \mathbb{E}[XY_n] &= (R_{YY} * h)(n), \quad n \in I. \\ h(n) &= 0, \quad n \notin I. \end{aligned}$$

Corollary 3.22

Let $Y = (Y_n)_{n\in\mathbb{Z}}$ be a zero-mean WSS stationary process with regular covariance. The sequence $h \in l^2(\mathbb{Z})$ defining the best linear estimator $\mathbb{L}[X|Y]$ via (3.20) (for $I = \mathbb{Z}$) has Fourier transform

$$H(\omega) = \frac{S_{YX}(\omega)}{S_{YY(\omega)}}, \quad \omega \in [-\pi, \pi).$$

Moreover, the resulting estimation error is

$$\mathbb{E}[|X - \mathbb{L}[X|Y]|^2] = \mathrm{Var}(X) - \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|S_{YX}(\omega)|^2}{S_{YY}(\omega)}\, d\omega.$$

3.4 The Noncasual Wiener Filter

Theorem 3.23: The noncasual Wiener filter

Let $X = (X_n)_{n\in\mathbb{Z}}$ and $Y = (Y_n)_{n\in\mathbb{Z}}$ be zero-mean JWSS process, and assume that Y has regular covariance. The process $(\mathbb{L}[X_n|Y])_{n\in\mathbb{Z}}$ can be realized by passing Y through a LTI system with frequence response

$$H(\omega) = \frac{S_{YX}^*(\omega)}{S_{YY}(\omega)}, \quad \omega \in [-\pi, \pi).$$

The resulting estimation error is

$$\mathbb{E}[|X_n - \mathbb{L}[X_n|Y]|^2] = R_{XX}(0) - \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|S_{YX}(\omega)|^2}{S_{YY}(\omega)}\, d\omega.$$

3.5 Linear Filtering of WSS process

Theorem 3.24: Effect of LTI systems on WSS processes

Let $X = (X_n)_{n\in\mathbb{Z}}$ be a zero-mean WSS process and consider a stable LTI system with impulse response g (frequency response G). There exists a zero-mean process $Y = (Y_n)_{n\in\mathbb{Z}}$, JWSS with X, such that

$$Y_n = \sum_k g(k) X_{n-k} \text{ a.s. and in } L^2, \text{ for all } n \in \mathbb{Z}.$$

Moreover, if $R_{XX} \in l^2(\mathbb{Z})$, then $R_{YY} \in l^2(\mathbb{Z})$ and

$$S_{YY}(\omega) = S_{XX}(\omega)|G(\omega)|^2.$$

The cross-power spectrum S_{YX} also exists in this case, and is equal to

$$S_{YX}(\omega) = S_{XX} G(\omega).$$