

Review Session 3

Sangmyung Ha

Fall 2026

1 Linear Filtering

Let $\mathbb{T} = [-\pi, \pi)$ denote the frequency domain, with frequencies identified modulo 2π , and let $\mathcal{B}(\mathbb{T})$ be its Borel σ -field. All stochastic integrals below are understood in the mean-square sense. We use the negative-exponential convention for spectral representations. When a spectral measure F_X is absolutely continuous with respect to Lebesgue measure, we define its spectral density f_X by

$$F_X(d\lambda) = \frac{f_X(\lambda)}{2\pi} d\lambda.$$

Thus, for a mean-zero process, the autocovariance function satisfies

$$\gamma_X(h) := \mathbb{E}[X_{t+h}\overline{X_t}] = \int_{\mathbb{T}} e^{-ih\lambda} F_X(d\lambda) = \frac{1}{2\pi} \int_{\mathbb{T}} e^{-ih\lambda} f_X(\lambda) d\lambda.$$

Theorem 1 (Linear filtering). *Let $(X_t)_{t \in \mathbb{Z}}$ be a mean-zero, covariance-stationary scalar process with $\mathbb{E}|X_t|^2 < \infty$ and spectral representation*

$$X_t = \int_{\mathbb{T}} e^{-it\lambda} dZ_X(\lambda),$$

where Z_X is a mean-zero orthogonal random measure with control measure F_X . Let $(\alpha_k)_{k \in \mathbb{Z}}$ satisfy

$$\sum_{k \in \mathbb{Z}} |\alpha_k| < \infty,$$

and define the transfer function

$$\alpha(e^{i\lambda}) := \sum_{k \in \mathbb{Z}} \alpha_k e^{ik\lambda}, \quad \lambda \in \mathbb{T}.$$

Then the filtered process

$$Y_t := \sum_{k \in \mathbb{Z}} \alpha_k X_{t-k} = \alpha(L)X_t, \quad \alpha(L) := \sum_{k \in \mathbb{Z}} \alpha_k L^k,$$

is well defined as a mean-square limit and is mean-zero and covariance stationary. Its spectral representation is

$$Y_t = \int_{\mathbb{T}} e^{-it\lambda} dZ_Y(\lambda) = \int_{\mathbb{T}} e^{-it\lambda} \alpha(e^{i\lambda}) dZ_X(\lambda), \quad (1)$$

where

$$Z_Y(\Lambda) = \int_{\Lambda} \alpha(e^{i\lambda}) dZ_X(\lambda), \quad \Lambda \in \mathcal{B}(\mathbb{T}). \quad (2)$$

The spectral measure of Y_t is

$$F_Y(\Lambda) = \int_{\Lambda} |\alpha(e^{i\lambda})|^2 F_X(d\lambda), \quad \Lambda \in \mathcal{B}(\mathbb{T}). \quad (3)$$

In particular, if F_X is absolutely continuous with respect to Lebesgue measure, then so is F_Y , and

$$f_Y(\lambda) = |\alpha(e^{i\lambda})|^2 f_X(\lambda) \quad (4)$$

for Lebesgue-almost every $\lambda \in \mathbb{T}$.

Proof. First, we establish the existence of the filtered process. Let

$$Y_t^{(n)} := \sum_{k=-n}^n \alpha_k X_{t-k}.$$

For $m > n$, the triangle inequality in $\mathcal{L}^2(\Omega)$ and covariance stationarity imply

$$\begin{aligned} \|Y_t^{(m)} - Y_t^{(n)}\|_{\mathcal{L}^2(\Omega)} &\leq \sum_{n < |k| \leq m} |\alpha_k| \|X_{t-k}\|_{\mathcal{L}^2(\Omega)} \\ &= \sqrt{\gamma_X(0)} \sum_{n < |k| \leq m} |\alpha_k| \\ &\leq \sqrt{\gamma_X(0)} \sum_{|k| > n} |\alpha_k| \longrightarrow 0. \end{aligned}$$

Hence $(Y_t^{(n)})$ is Cauchy in $\mathcal{L}^2(\Omega)$ and converges in mean square to a random variable Y_t . Since $\mathbb{E}[Y_t^{(n)}] = 0$, mean-square convergence also gives $\mathbb{E}[Y_t] = 0$.

Next, define

$$A_n(\lambda) := \sum_{k=-n}^n \alpha_k e^{ik\lambda}, \quad A(\lambda) := \alpha(e^{i\lambda}).$$

Absolute summability gives uniform convergence:

$$\sup_{\lambda \in \mathbb{T}} |A_n(\lambda) - A(\lambda)| \leq \sum_{|k| > n} |\alpha_k| \longrightarrow 0.$$

For each n , linearity of the stochastic integral yields

$$\begin{aligned} Y_t^{(n)} &= \sum_{k=-n}^n \alpha_k \int_{\mathbb{T}} e^{-i(t-k)\lambda} dZ_X(\lambda) \\ &= \int_{\mathbb{T}} e^{-it\lambda} \left(\sum_{k=-n}^n \alpha_k e^{ik\lambda} \right) dZ_X(\lambda) \\ &= \int_{\mathbb{T}} e^{-it\lambda} A_n(\lambda) dZ_X(\lambda). \end{aligned}$$

By the isometry of the stochastic integral,

$$\begin{aligned}\mathbb{E} \left| \int_{\mathbb{T}} e^{-it\lambda} (A_n(\lambda) - A(\lambda)) dZ_X(\lambda) \right|^2 &= \int_{\mathbb{T}} |A_n(\lambda) - A(\lambda)|^2 F_X(d\lambda) \\ &\leq F_X(\mathbb{T}) \left(\sum_{|k|>n} |\alpha_k| \right)^2 \\ &\longrightarrow 0.\end{aligned}$$

Taking mean-square limits therefore gives

$$Y_t = \int_{\mathbb{T}} e^{-it\lambda} A(\lambda) dZ_X(\lambda).$$

Define

$$Z_Y(\Lambda) := \int_{\Lambda} A(\lambda) dZ_X(\lambda), \quad \Lambda \in \mathcal{B}(\mathbb{T}).$$

These integrals are well defined because A is bounded and F_X is finite. Moreover, $\mathbb{E}[Z_Y(\Lambda)] = 0$, and the inner-product version of the isometry gives

$$\begin{aligned}\mathbb{E}[Z_Y(\Lambda_1) \overline{Z_Y(\Lambda_2)}] &= \int_{\mathbb{T}} \mathbf{1}_{\Lambda_1}(\lambda) \mathbf{1}_{\Lambda_2}(\lambda) |A(\lambda)|^2 F_X(d\lambda) \\ &= \int_{\Lambda_1 \cap \Lambda_2} |A(\lambda)|^2 F_X(d\lambda).\end{aligned}$$

Thus $Z_Y(\Lambda_1)$ and $Z_Y(\Lambda_2)$ are orthogonal whenever Λ_1 and Λ_2 are disjoint.

To verify countable additivity in mean square, let $(\Lambda_j)_{j \geq 1}$ be pairwise disjoint Borel sets and put $\Lambda = \bigcup_{j \geq 1} \Lambda_j$. Then

$$\mathbb{E} \left| Z_Y(\Lambda) - \sum_{j=1}^m Z_Y(\Lambda_j) \right|^2 = \int_{\Lambda \setminus \bigcup_{j=1}^m \Lambda_j} |A(\lambda)|^2 F_X(d\lambda) \longrightarrow 0.$$

Consequently, Z_Y is a mean-zero orthogonal random measure with control measure

$$F_Y(\Lambda) := \mathbb{E}|Z_Y(\Lambda)|^2 = \int_{\Lambda} |A(\lambda)|^2 F_X(d\lambda).$$

The identity

$$\int_{\mathbb{T}} g(\lambda) dZ_Y(\lambda) = \int_{\mathbb{T}} g(\lambda) A(\lambda) dZ_X(\lambda)$$

holds first for simple functions and then, by the isometry, for every $g \in \mathcal{L}^2(\mathbb{T}, F_Y)$. Taking $g(\lambda) = e^{-it\lambda}$ establishes (1).

Furthermore,

$$\begin{aligned}\mathbb{E}[Y_{t+h} \overline{Y_t}] &= \int_{\mathbb{T}} e^{-i(t+h)\lambda} A(\lambda) \overline{e^{-it\lambda} A(\lambda)} F_X(d\lambda) \\ &= \int_{\mathbb{T}} e^{-ih\lambda} |A(\lambda)|^2 F_X(d\lambda) \\ &= \int_{\mathbb{T}} e^{-ih\lambda} F_Y(d\lambda).\end{aligned}$$

This expression depends only on h , so (Y_t) is covariance stationary and F_Y is its spectral measure.

Finally, (3) implies $F_Y \ll F_X$. Hence, if $F_X \ll d\lambda$, then $F_Y \ll d\lambda$, and

$$F_Y(d\lambda) = |\alpha(e^{i\lambda})|^2 F_X(d\lambda) = \frac{|\alpha(e^{i\lambda})|^2 f_X(\lambda)}{2\pi} d\lambda.$$

Comparing this with $F_Y(d\lambda) = f_Y(\lambda) d\lambda / (2\pi)$ proves (4). $\square$

Nonzero means. If X_t has a nonzero constant mean μ_X , the theorem applies to the centered process $X_t - \mu_X$. The filtered process has mean

$$\mu_Y = \alpha(1)\mu_X, \quad \alpha(1) = \sum_{k \in \mathbb{Z}} \alpha_k,$$

and

$$Y_t - \mu_Y = \sum_{k \in \mathbb{Z}} \alpha_k (X_{t-k} - \mu_X).$$

Thus the same spectral-measure and spectral-density formulas hold for the centered processes.

Corollary 1 (Spectrum of an ARMA process). *Let (X_t) be a mean-zero, covariance-stationary solution of*

$$\alpha(L)X_t = \beta(L)\varepsilon_t,$$

where α and β are polynomials and $(\varepsilon_t) \sim WN(0, \sigma^2)$ with $\sigma^2 > 0$. Then, as an equality of measures on $\mathbb{T}$,

$$|\alpha(e^{i\lambda})|^2 F_X(d\lambda) = |\beta(e^{i\lambda})|^2 F_\varepsilon(d\lambda) = \frac{\sigma^2}{2\pi} |\beta(e^{i\lambda})|^2 d\lambda. \quad (5)$$

If, in addition,

$$\alpha(z) \neq 0 \quad \text{for every } |z| = 1,$$

then F_X is absolutely continuous and

$$f_X(\lambda) = \sigma^2 \frac{|\beta(e^{i\lambda})|^2}{|\alpha(e^{i\lambda})|^2} \quad (6)$$

for Lebesgue-almost every $\lambda \in \mathbb{T}$.

Proof. Define the common filtered process

$$U_t := \alpha(L)X_t = \beta(L)\varepsilon_t.$$

Since polynomial filters have absolutely summable coefficients, Theorem 1 applies to both representations and gives

$$F_U(d\lambda) = |\alpha(e^{i\lambda})|^2 F_X(d\lambda) = |\beta(e^{i\lambda})|^2 F_\varepsilon(d\lambda).$$

For white noise,

$$\gamma_\varepsilon(h) = \sigma^2 \mathbf{1}\{h = 0\}, \quad f_\varepsilon(\lambda) = \sum_{h \in \mathbb{Z}} \gamma_\varepsilon(h) e^{ih\lambda} = \sigma^2.$$

Therefore $F_\varepsilon(d\lambda) = \sigma^2 d\lambda/(2\pi)$, which proves (5).

If α has no zeros on the unit circle, continuity implies that $|\alpha(e^{i\lambda})|$ is bounded away from zero. Dividing the measure identity by $|\alpha(e^{i\lambda})|^2$ yields

$$F_X(d\lambda) = \frac{\sigma^2}{2\pi} \frac{|\beta(e^{i\lambda})|^2}{|\alpha(e^{i\lambda})|^2} d\lambda,$$

which proves absolute continuity and (6). $\square$

Causality and invertibility. The condition that α has no zeros on the unit circle does not itself impose causality. The stronger condition $\alpha(z) \neq 0$ for every $|z| \leq 1$ yields the usual causal stationary solution. No invertibility condition on the moving-average polynomial β is needed for the spectral-density formula.

1.1 Example: ARMA(1,1)

Let $\alpha, \beta \in \mathbb{R}$, suppose $|\alpha| < 1$, and let $(\varepsilon_t) \sim WN(0, \sigma^2)$. Consider the causal stationary solution of

$$\begin{aligned} X_t &= \alpha X_{t-1} + \varepsilon_t + \beta \varepsilon_{t-1}, \\ (1 - \alpha L)X_t &= (1 + \beta L)\varepsilon_t. \end{aligned}$$

Applying Corollary 1 gives

$$|1 - \alpha e^{i\lambda}|^2 f_X(\lambda) = |1 + \beta e^{i\lambda}|^2 \sigma^2.$$

Since $|\alpha| < 1$, the denominator below is strictly positive. Choosing the continuous version of the spectral density, we obtain

$$\begin{aligned} f_X(\lambda) &= \sigma^2 \frac{|1 + \beta e^{i\lambda}|^2}{|1 - \alpha e^{i\lambda}|^2} \\ &= \sigma^2 \frac{(1 + \beta e^{i\lambda})(1 + \beta e^{-i\lambda})}{(1 - \alpha e^{i\lambda})(1 - \alpha e^{-i\lambda})} \\ &= \sigma^2 \frac{1 + \beta^2 + \beta(e^{i\lambda} + e^{-i\lambda})}{1 + \alpha^2 - \alpha(e^{i\lambda} + e^{-i\lambda})} \\ &= \sigma^2 \frac{1 + \beta^2 + 2\beta \cos \lambda}{1 + \alpha^2 - 2\alpha \cos \lambda}. \end{aligned}$$

The additional restriction $|\beta| < 1$ imposes the usual moving-average invertibility condition, but it is not needed for this calculation.

2 Root Flipping in ARMA Processes

Let $(X_t)_{t \in \mathbb{Z}}$ be a real-valued, mean-zero, covariance-stationary process satisfying

$$X_t = \sum_{\ell=1}^p \alpha_\ell X_{t-\ell} + \sum_{\ell=0}^q \theta_\ell \varepsilon_{t-\ell}, \quad \theta_0 = 1, \tag{7}$$

where $(\varepsilon_t) \sim WN(0, \sigma^2)$ and $\sigma^2 > 0$. Define the real-coefficient polynomials

$$\alpha(z) := 1 - \sum_{\ell=1}^p \alpha_\ell z^\ell, \quad \theta(z) := 1 + \sum_{\ell=1}^q \theta_\ell z^\ell,$$

so that (7) becomes

$$\alpha(L)X_t = \theta(L)\varepsilon_t.$$

We take p and q to be the degrees of α and θ , respectively, and assume

$$\alpha(z)\theta(z) \neq 0 \quad \text{for every } |z| = 1. \quad (8)$$

Empty sums and products are interpreted as zero and one, respectively. All infinite filters below are understood in the mean-square sense.

The original representation need not be causal or invertible. We will show that reflecting every root inside the unit circle to its reciprocal conjugate produces a causal and invertible model with the same autocovariance function, after an appropriate adjustment to the white-noise variance. We will also show that the original process X_t itself admits a causal and invertible representation with respect to a suitably constructed weak white-noise sequence.

Two-sided inverse filters. Under (8), there exists an annulus $r_0 < |z| < R_0$, with $r_0 < 1 < R_0$, on which

$$\begin{aligned} \frac{1}{\alpha(z)} &= \sum_{n \in \mathbb{Z}} a_n z^n, \\ \frac{1}{\theta(z)} &= \sum_{n \in \mathbb{Z}} b_n z^n. \end{aligned} \quad (9)$$

Moreover,

$$\sum_{n \in \mathbb{Z}} |a_n| < \infty, \quad \sum_{n \in \mathbb{Z}} |b_n| < \infty.$$

These Laurent expansions provide two-sided inverse filters, which may involve future as well as past observations.

Reflecting the roots. Because $\alpha(0) = \theta(0) = 1$, the fundamental theorem of algebra gives

$$\alpha(z) = \prod_{\ell=1}^p \left(1 - \frac{z}{\alpha_\ell^*}\right), \quad \theta(z) = \prod_{\ell=1}^q \left(1 - \frac{z}{\theta_\ell^*}\right),$$

where roots are counted with multiplicity. Order the roots so that

$$\begin{aligned} |\alpha_\ell^*| &> 1 & (\ell = 1, \dots, r), & & |\alpha_\ell^*| < 1 & (\ell = r+1, \dots, p), \\ |\theta_\ell^*| &> 1 & (\ell = 1, \dots, s), & & |\theta_\ell^*| < 1 & (\ell = s+1, \dots, q). \end{aligned}$$

For each root a inside the unit circle, replace a by $1/\bar{a}$. This replaces the factor $1 - z/a$ by $1 - \bar{a}z$. Define

$$\begin{aligned} \tilde{\alpha}(z) &:= \prod_{\ell=1}^r \left(1 - \frac{z}{\alpha_\ell^*}\right) \prod_{\ell=r+1}^p (1 - \bar{\alpha}_\ell^* z), \\ \tilde{\theta}(z) &:= \prod_{\ell=1}^s \left(1 - \frac{z}{\theta_\ell^*}\right) \prod_{\ell=s+1}^q (1 - \bar{\theta}_\ell^* z). \end{aligned} \quad (10)$$

All roots of $\tilde{\alpha}$ and $\tilde{\theta}$ lie outside the unit circle. Since nonreal roots of a real-coefficient polynomial occur in conjugate pairs, reflecting all interior roots preserves real coefficients. Also, $\tilde{\alpha}(0) = \tilde{\theta}(0) = 1$.

Define the positive constants

$$C_\alpha := \prod_{\ell=r+1}^p |\alpha_\ell^*|^2, \quad C_\theta := \prod_{\ell=s+1}^q |\theta_\ell^*|^2. \quad (11)$$

Theorem 2 (Root flipping). *Under the preceding assumptions, let*

$$\tilde{\sigma}^2 := \frac{C_\alpha}{C_\theta} \sigma^2. \quad (12)$$

For any weak white-noise sequence $(\tilde{\varepsilon}_t) \sim WN(0, \tilde{\sigma}^2)$, the causal stationary solution of

$$\tilde{\alpha}(L)\tilde{X}_t = \tilde{\theta}(L)\tilde{\varepsilon}_t \quad (13)$$

is also invertible and has the same autocovariance function as X_t .

Furthermore, the original process itself satisfies

$$\tilde{\alpha}(L)X_t = \tilde{\theta}(L)\eta_t, \quad (14)$$

where

$$\eta_t := \tilde{\theta}(L)^{-1} \tilde{\alpha}(L)X_t \quad (15)$$

is weak white noise with variance $\tilde{\sigma}^2$. This representation is causal and invertible, and η_t is the one-step linear innovation of X_t .

Proof. For any nonzero a and any z on the unit circle,

$$|1 - \bar{a}z| = |1 - a\bar{z}| = |a| \left| 1 - \frac{z}{a} \right|.$$

Applying this identity to every reflected root yields

$$\begin{aligned} |\tilde{\alpha}(e^{i\lambda})|^2 &= C_\alpha |\alpha(e^{i\lambda})|^2, \\ |\tilde{\theta}(e^{i\lambda})|^2 &= C_\theta |\theta(e^{i\lambda})|^2. \end{aligned} \quad (16)$$

Since all roots of the reflected polynomials lie outside the unit circle, the rational functions

$$\psi(z) := \frac{\tilde{\theta}(z)}{\tilde{\alpha}(z)}, \quad \pi(z) := \frac{\tilde{\alpha}(z)}{\tilde{\theta}(z)}$$

have power-series expansions

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j, \quad \pi(z) = \sum_{j=0}^{\infty} \pi_j z^j,$$

with

$$\psi_0 = \pi_0 = 1, \quad \sum_{j=0}^{\infty} |\psi_j| < \infty, \quad \sum_{j=0}^{\infty} |\pi_j| < \infty.$$

Thus

$$\tilde{X}_t = \sum_{j=0}^{\infty} \psi_j \tilde{\varepsilon}_{t-j}$$

defines a causal stationary solution of (13), and

$$\tilde{\varepsilon}_t = \sum_{j=0}^{\infty} \pi_j \tilde{X}_{t-j}$$

establishes invertibility.

Under the spectral-density normalization used in the preceding section, white noise with variance σ^2 has spectral density σ^2 . By Corollary 1, we may therefore take

$$f_X(\lambda) = \sigma^2 \frac{|\theta(e^{i\lambda})|^2}{|\alpha(e^{i\lambda})|^2}.$$

Using (16) and (12), the spectral density of $\tilde{X}_t$ is

$$\begin{aligned} f_{\tilde{X}}(\lambda) &= \tilde{\sigma}^2 \frac{|\tilde{\theta}(e^{i\lambda})|^2}{|\tilde{\alpha}(e^{i\lambda})|^2} \\ &= \left(\frac{C_\alpha}{C_\theta} \sigma^2 \right) \frac{C_\theta |\theta(e^{i\lambda})|^2}{C_\alpha |\alpha(e^{i\lambda})|^2} \\ &= \sigma^2 \frac{|\theta(e^{i\lambda})|^2}{|\alpha(e^{i\lambda})|^2} \\ &= f_X(\lambda). \end{aligned}$$

Consequently,

$$\gamma_{\tilde{X}}(h) = \frac{1}{2\pi} \int_{\mathbb{T}} e^{-ih\lambda} f_{\tilde{X}}(\lambda) d\lambda = \frac{1}{2\pi} \int_{\mathbb{T}} e^{-ih\lambda} f_X(\lambda) d\lambda = \gamma_X(h)$$

for every $h \in \mathbb{Z}$.

To obtain a new representation of the same process X_t , define

$$\eta_t := \pi(L)X_t = \tilde{\theta}(L)^{-1}\tilde{\alpha}(L)X_t.$$

This is a well-defined filter of current and past observations. By Theorem 1,

$$\begin{aligned} f_\eta(\lambda) &= \frac{|\tilde{\alpha}(e^{i\lambda})|^2}{|\tilde{\theta}(e^{i\lambda})|^2} f_X(\lambda) \\ &= \frac{C_\alpha |\alpha(e^{i\lambda})|^2}{C_\theta |\theta(e^{i\lambda})|^2} \left(\sigma^2 \frac{|\theta(e^{i\lambda})|^2}{|\alpha(e^{i\lambda})|^2} \right) \\ &= \frac{C_\alpha}{C_\theta} \sigma^2 \\ &= \tilde{\sigma}^2. \end{aligned}$$

Since η_t is centered,

$$\mathbb{E}[\eta_{t+h}\eta_t] = \frac{1}{2\pi} \int_{\mathbb{T}} e^{-ih\lambda} \tilde{\sigma}^2 d\lambda = \tilde{\sigma}^2 \mathbf{1}\{h = 0\}.$$

Hence $(\eta_t) \sim WN(0, \tilde{\sigma}^2)$. Its definition gives (14), and the absolutely summable inverse filters imply

$$X_t = \sum_{j=0}^{\infty} \psi_j \eta_{t-j}, \quad \eta_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}.$$

Thus the representation of the original process is both causal and invertible.

Finally, let

$$\mathcal{H}_{t-1}^X := \overline{\text{span}}\{X_s : s \leq t-1\}, \quad \mathcal{H}_{t-1}^\eta := \overline{\text{span}}\{\eta_s : s \leq t-1\},$$

where the closures are taken in $\mathcal{L}^2(\Omega)$. The two one-sided representations imply

$$\mathcal{H}_{t-1}^X = \mathcal{H}_{t-1}^\eta.$$

Since $\psi_0 = 1$,

$$X_t = \eta_t + \sum_{j=1}^{\infty} \psi_j \eta_{t-j}.$$

The sum belongs to $\mathcal{H}_{t-1}^\eta = \mathcal{H}_{t-1}^X$, whereas η_t is orthogonal to this space because (η_t) is white noise. Therefore, if P_{t-1} denotes the orthogonal projection onto $\mathcal{H}_{t-1}^X$,

$$P_{t-1}X_t = \sum_{j=1}^{\infty} \psi_j \eta_{t-j}, \quad \eta_t = X_t - P_{t-1}X_t.$$

Thus η_t is the one-step linear innovation. □

A new process versus a new representation. One possible choice in (13) is

$$\tilde{\varepsilon}_t = \sqrt{\frac{C_\alpha}{C_\theta}} \varepsilon_t.$$

This constructs a process $\tilde{X}_t$ with the same autocovariance function as X_t , but generally $\tilde{X}_t$ and X_t are different random variables. In contrast, (14) leaves the original process unchanged and replaces its driving noise by η_t . In general, η_t is not a scalar rescaling of ε_t .

The sequence (η_t) need not be independent over time, even when (ε_t) is independent. Accordingly, root flipping preserves the second-order structure but does not generally preserve an independent-noise specification or the finite-dimensional distributions. If both X_t and $\tilde{X}_t$ are Gaussian processes, their equal means and autocovariance functions do imply equality of their finite-dimensional distributions.

The resulting ARMA orders. The reflected polynomials retain degrees p and q , but they may have common factors, even if the original polynomials were coprime. Canceling common factors gives a causal and invertible representation of orders at most p and q , respectively. For example,

$$(1 - 2L)X_t = (1 - \tfrac{1}{2}L)\varepsilon_t$$

has reflected polynomials

$$\tilde{\alpha}(z) = \tilde{\theta}(z) = 1 - \tfrac{1}{2}z$$

and adjusted noise variance $\tilde{\sigma}^2 = \sigma^2/4$. The reflected representation is

$$(1 - \tfrac{1}{2}L)X_t = (1 - \tfrac{1}{2}L)\eta_t, \quad (\eta_t) \sim WN(0, \sigma^2/4).$$

The common factor cancels, leaving $X_t = \eta_t$. Thus the minimal causal and invertible representation in this example is white noise rather than ARMA(1, 1).

2.1 Example: MA(1)

Consider

$$X_t = \varepsilon_t + \theta \varepsilon_{t-1}, \quad (\varepsilon_t) \sim WN(0, \sigma^2),$$

where $\theta \in \mathbb{R}$ and $|\theta| > 1$. Its autocovariance function is

$$\gamma_X(h) = \begin{cases} (1 + \theta^2)\sigma^2, & h = 0, \\ \theta\sigma^2, & |h| = 1, \\ 0, & |h| > 1. \end{cases}$$

The moving-average polynomial

$$\theta(z) = 1 + \theta z$$

has root $z^* = -1/\theta$, which lies inside the unit circle. Reflecting this root gives

$$\frac{1}{z^*} = -\theta,$$

so the reflected polynomial is

$$\tilde{\theta}(z) = 1 + \theta^{-1}z.$$

Since $|\theta^{-1}| < 1$, the reflected representation is invertible. Here

$$C_\alpha = 1, \quad C_\theta = \frac{1}{\theta^2}, \quad \tilde{\sigma}^2 = \theta^2 \sigma^2.$$

To obtain an invertible representation of the original process, define

$$\eta_t := (1 + \theta^{-1}L)^{-1}X_t = \sum_{j=0}^{\infty} (-\theta^{-1})^j X_{t-j}.$$

The series is well defined because $|\theta^{-1}| < 1$. By Theorem 2,

$$(\eta_t) \sim WN(0, \theta^2 \sigma^2),$$

and the same process X_t satisfies

$$X_t = \eta_t + \theta^{-1}\eta_{t-1}.$$

This representation is both causal and invertible, and η_t is the one-step linear innovation.

Alternatively, define a rescaled copy of the original noise by

$$\tilde{\varepsilon}_t := |\theta|\varepsilon_t,$$

and construct the new process

$$\begin{aligned} \tilde{X}_t &= \tilde{\varepsilon}_t + \theta^{-1}\tilde{\varepsilon}_{t-1} \\ &= |\theta|\varepsilon_t + \frac{|\theta|}{\theta}\varepsilon_{t-1}. \end{aligned}$$

Since $\text{Var}(\tilde{\varepsilon}_t) = \theta^2 \sigma^2$,

$$\begin{aligned} \gamma_{\tilde{X}}(0) &= (1 + \theta^{-2})\theta^2 \sigma^2 = (1 + \theta^2)\sigma^2, \\ \gamma_{\tilde{X}}(1) &= \theta^{-1}\theta^2 \sigma^2 = \theta\sigma^2, \end{aligned}$$

and $\gamma_{\tilde{X}}(h) = 0$ for $|h| > 1$. Therefore $\tilde{X}_t$ and X_t have the same autocovariance function, although the processes generally differ.

When $|\theta| < 1$, the original MA(1) representation is already invertible and no root needs to be reflected. The case $\theta = 0$ is white noise, while $|\theta| = 1$ is excluded by the no-unit-circle-root assumption.

2.2 Example: AR(1)

Consider a stationary solution of

$$X_t = \alpha X_{t-1} + \varepsilon_t, \quad (\varepsilon_t) \sim WN(0, \sigma^2),$$

where $\alpha \in \mathbb{R} \setminus \{0\}$ and $|\alpha| \neq 1$. Backward iteration when $|\alpha| < 1$ and forward iteration when $|\alpha| > 1$ give the mean-square representations

$$X_t = \begin{cases} \sum_{j=0}^{\infty} \alpha^j \varepsilon_{t-j}, & |\alpha| < 1, \\ -\sum_{j=1}^{\infty} \alpha^{-j} \varepsilon_{t+j}, & |\alpha| > 1. \end{cases}$$

Thus the stationary solution is causal when $|\alpha| < 1$ and noncausal when $|\alpha| > 1$. Its autocovariance function is

$$\gamma_X(h) = \begin{cases} \frac{\sigma^2 \alpha^{|h|}}{1 - \alpha^2}, & |\alpha| < 1, \\ \frac{\sigma^2 \alpha^{-|h|}}{\alpha^2 - 1}, & |\alpha| > 1. \end{cases}$$

Now suppose $|\alpha| > 1$, so that the original representation is noncausal. The autoregressive polynomial

$$\alpha(z) = 1 - \alpha z$$

has root $z^* = 1/\alpha$ inside the unit circle. Reflecting this root gives α , and the reflected polynomial is

$$\tilde{\alpha}(z) = 1 - \alpha^{-1}z.$$

Here

$$C_\alpha = \frac{1}{\alpha^2}, \quad C_\theta = 1, \quad \tilde{\sigma}^2 = \frac{\sigma^2}{\alpha^2}.$$

For a causal representation of the original process, define

$$\eta_t := X_t - \alpha^{-1}X_{t-1}.$$

Theorem 2 gives

$$(\eta_t) \sim WN(0, \sigma^2/\alpha^2),$$

and

$$X_t = \alpha^{-1}X_{t-1} + \eta_t = \sum_{j=0}^{\infty} \alpha^{-j} \eta_{t-j}.$$

Since $|\alpha^{-1}| < 1$, this is a causal representation, and η_t is the one-step linear innovation. The original driving noise ε_t is not the one-step innovation in the noncausal representation. Indeed, the forward-noise expansion gives

$$\mathbb{E}[\varepsilon_t X_{t-1}] = -\frac{\sigma^2}{\alpha} \neq 0.$$

Alternatively, define

$$\tilde{\varepsilon}_t := \frac{1}{|\alpha|} \varepsilon_t$$

and let $\tilde{X}_t$ be the causal stationary solution of

$$\tilde{X}_t = \alpha^{-1} \tilde{X}_{t-1} + \tilde{\varepsilon}_t.$$

Its autocovariance function is

$$\begin{aligned} \gamma_{\tilde{X}}(h) &= \frac{(\sigma^2/\alpha^2)\alpha^{-|h|}}{1 - \alpha^{-2}} \\ &= \frac{\sigma^2\alpha^{-|h|}}{\alpha^2 - 1} \\ &= \gamma_X(h). \end{aligned}$$

Thus $\tilde{X}_t$ is a causal process with the same autocovariance function as the original noncausal process.

In this example, root flipping changes causality, not invertibility. Both representations of X_t are invertible because their driving noises can be recovered from current and past observations:

$$\varepsilon_t = X_t - \alpha X_{t-1}, \quad \eta_t = X_t - \alpha^{-1} X_{t-1}.$$

When $|\alpha| < 1$, the original representation is already causal and no AR root needs to be reflected. The case $\alpha = 0$ is white noise, while $|\alpha| = 1$ is excluded by the no-unit-circle-root assumption.

3 Estimation of the Spectral Density

Let (X_t) be a real-valued, mean-zero, covariance-stationary process with autocovariance function $\gamma(k)$ satisfying $\sum_{k \in \mathbb{Z}} |\gamma(k)| < \infty$. Under our normalization, its spectral density is

$$f(\lambda) = \sum_{k \in \mathbb{Z}} \gamma(k) e^{ik\lambda}, \quad \lambda \in \mathbb{T} = [-\pi, \pi).$$

The periodogram. Define the normalized finite Fourier transform and the periodogram by

$$d_n(\lambda) := \frac{1}{\sqrt{n}} \sum_{t=1}^n X_t e^{it\lambda}, \quad I_n(\lambda) := |d_n(\lambda)|^2 = \frac{1}{n} \left| \sum_{t=1}^n X_t e^{it\lambda} \right|^2.$$

Let

$$\hat{\gamma}_n(k) := \frac{1}{n} \sum_{t=k+1}^n X_t X_{t-k}, \quad k = 0, \dots, n-1,$$

and define $\hat{\gamma}_n(-k) := \hat{\gamma}_n(k)$. Grouping terms according to $k = t - s$ gives

$$I_n(\lambda) = \frac{1}{n} \sum_{t=1}^n \sum_{s=1}^n X_t X_s e^{i(t-s)\lambda} = \sum_{|k| < n} \hat{\gamma}_n(k) e^{ik\lambda}.$$

Thus the periodogram is the Fourier transform of the sample autocovariances. The denominator n , rather than $n - |k|$, is essential for this identity.

Why smoothing is needed. Since $\mathbb{E}[\hat{\gamma}_n(k)] = (1 - |k|/n)\gamma(k)$,

$$\mathbb{E}[I_n(\lambda)] = \sum_{|k| < n} \left(1 - \frac{|k|}{n}\right) \gamma(k) e^{ik\lambda} \longrightarrow f(\lambda),$$

where convergence follows from absolute summability. Nevertheless, the periodogram is not generally consistent because its random fluctuations need not vanish.

For example, if $X_t \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma^2)$, then $f(\lambda) = \sigma^2$, and at Fourier frequencies $\lambda_j = 2\pi j/n$ with $1 \leq j < n/2$,

$$\frac{I_n(\lambda_j)}{\sigma^2} \sim \frac{\chi_2^2}{2}, \quad \text{Var}(I_n(\lambda_j)) = \sigma^4.$$

This nondegenerate distribution persists as the sample grows. In particular, it holds at the fixed frequency $\lambda = \pi/2$ whenever n is divisible by four, ruling out consistency there. The problem is accumulated estimation error across many lags.

Lag-window estimation. To control this accumulated error, use

$$\hat{f}_n(\lambda) := \sum_{|k| < l_n} w\left(\frac{k}{l_n}\right) \hat{\gamma}_n(k) e^{ik\lambda}, \quad (17)$$

where $l_n \leq n$ is a positive integer and w is a lag-window kernel. Assume that w is real, even, bounded, continuous at zero, satisfies $w(0) = 1$, and vanishes for $|u| \geq 1$. Choose

$$l_n \longrightarrow \infty, \quad \frac{l_n}{n} \longrightarrow 0. \quad (18)$$

The first condition removes truncation and weighting bias; the second controls variance.

These bandwidth conditions require accompanying moment and dependence assumptions. For example, if (X_t) is Gaussian and its autocovariances are absolutely summable as assumed above, then, for every fixed λ ,

$$\mathbb{E}[\hat{f}_n(\lambda)] - f(\lambda) \longrightarrow 0, \quad \text{Var}(\hat{f}_n(\lambda)) = O\left(\frac{l_n}{n}\right) \longrightarrow 0.$$

Consequently, $\hat{f}_n(\lambda) \rightarrow f(\lambda)$ in mean square and in probability. Gaussianity can be replaced by suitable fourth-moment and weak-dependence conditions.

Choosing the lag window and truncation parameter. A larger l_n retains more autocovariances and typically reduces bias but increases variance. A smaller l_n gives stronger smoothing and lower variance but may obscure spectral peaks. For standard windows, the frequency-smoothing width is of order $1/l_n$, containing approximately n/l_n Fourier frequencies. Thus (18) makes the smoothing window narrower while allowing an increasing number of frequencies to contribute.

Kernel choice affects bias, variance, and nonnegativity. A common choice is the Bartlett lag window,

$$w(u) = (1 - |u|)\mathbf{1}\{|u| \leq 1\},$$

whose nonnegative spectral window guarantees $\hat{f}_n(\lambda) \geq 0$. Arbitrary lag windows, including hard lag truncation, need not produce nonnegative estimates.

Unknown mean. When the mean is unknown, replace X_t by $X_t - \bar{X}_n$ in the sample autocovariances and periodogram. Under the preceding sufficient assumptions, this preserves pointwise consistency of the lag-window estimator. Although the centered raw periodogram equals zero at frequency zero, its smoothed version need not.