Lessons 3-4: Estimating parameters of weakly stationary time series (Ergodicity, Periodograms, Spectra of Spectral Densities)

Marianna Bolla and Tamás Szabados, BME Math. Inst.

October 8, 2020

1 Estimation of the mean

It is important in practice to estimate parameters of a stationary time series by observing a single trajectory of the process for a long enough time. The first thing to estimate is the mean $\mu \in \mathbb{C}$ of a process, which could differ from 0 now. It suffices to consider a one-dimensional time series $\{X_t\}_{t\in\mathbb{Z}}$, because expectation can be taken componentwise. If

$$X_t^\mu := X_t + \mu, \quad \mathbb{E}X_t = 0, \quad \mu \in \mathbb{C} \quad (t \in \mathbb{Z}),$$

then one gets a natural approximation of μ by taking a positive integer T and computing the *empirical mean*, that is, the average of a single trajectory for $t = 0, 1, \dots, T-1$:

$$\tilde{X}_T^\mu := \frac{1}{T}\sum_{t=0}^{T-1} X_t^\mu = \mu + \frac{1}{T}\sum_{t=0}^{T-1} X_t = \mu + \tilde{X}_T.$$

If we have convergence of the time average $\tilde{X}_T^\mu$ to the theoretical expectation μ in mean square then it is called *ergodicity for the mean*; it is a "law of large numbers". Obviously, for this it is necessary and sufficient that $\tilde{X}_T \to 0$ in mean square; this is the case that we are going to investigate in the sequel.

Let $H_X \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ be the Hilbert space defined in Lesson 1. In this section each random variable is considered as an element (vector) in H_X. So equality of two random variables means that they are $\mathbb{P}$-a.s. equal. Also, convergence of random variables is always understood in this Hilbert space, that is, as convergence in mean square.

Let S denote the unitary operator of forward shift in H_X and

$$\mathcal{I} := \{\xi \in H_X : S\xi = \xi\}$$

the subspace of translation invariant random variables in H_X. It is clear that $\mathcal{I}$ is a closed subspace in H_X.

Theorem 1. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be a stationary time series (with mean 0). Then the time average $\tilde{X}_T$ converges to a random variable Y in mean square as $T \to \infty$:*

$$\lim_{T\to\infty} \mathbb{E}|\tilde{X}_T - Y|^2 = 0,$$

where Y is the orthogonal projection of X_0 onto the subspace $\mathcal{I}$, denoted as $Y = P_{\mathcal{I}} X_0$; moreover, $\mathbb{E}Y = 0$.

Ergodicity means that $Y = 0$ a.s., that is only the 0 is invariant for the shift.

The next theorem gives a necessary and sufficient condition of ergodicity for the mean.

Theorem 2. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be a stationary time series with mean 0 and with covariance function $c(j)$ $(j \in \mathbb{Z})$. Then the time average $\tilde{X}_T$ converges to $Y = 0$ in mean square as $T \to \infty$ if and only if*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{j=0}^{T-1} c(j) = 0. \tag{1}$$

Corollary 1. *It is an elementary analysis fact that $\lim_{j\to\infty} c(j) = 0$ implies (1), thus it is a sufficient (but not necessary) condition of ergodicity for the mean. An even stronger sufficient condition is that $\sum_{j=-\infty}^{\infty} |c(j)| < \infty$ holds.*

Another approach to ergodicity for the mean is to use spectral representation of the time series.

Theorem 3. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be a stationary time series with mean 0, with stochastic increment process dZ_ω, and spectral measure $dF(\omega)$, for $\omega \in [-\pi, \pi]$.*

(a) Then the time average $\tilde{X}_T$ converges in mean square to $Y = Z_{\{0\}}$, the atom (point mass) of the random spectral measure at $\{0\}$, as $T \to \infty$.

(b) Ergodicity for the mean holds if and only if $dF(\{0\}) = 0$, that is, the spectral measure has no atom at $\{0\}$.

Remark 1. *By Theorem 3(b), any stationary time series whose spectral measure dF is absolutely continuous w.r.t. Lebesgue measure (that is, has a spectral density f) is ergodic for the mean, since then the spectral measure dF does not have atoms.*

Example 1. *A simple example for a time series which is* not *ergodic for the mean:*

$$X_k = \sum_{j=1}^{n} A_j e^{ik\omega_j} \quad (k \in \mathbb{Z}),$$

where $0 = \omega_1 < \omega_2 < \cdots < \omega_n < 2\pi$; $A_1, \ldots, A_n$ are uncorrelated random variables with mean 0 and variance $\sigma_j^2 > 0$ $(j = 1, \ldots, n)$. (The A_j's can be e.g. Gaussian random variables.)

This process is weakly stationary with

$$c(k) = \mathbb{E}(X_{m+k}\overline{X_m}) = \sum_{j=1}^{n} \mathbb{E}(|A_j|^2) e^{ik\omega_j} = \sum_{j=1}^{n} \sigma_j^2 e^{ik\omega_j} = \int_0^{2\pi} e^{ik\omega}\, dF(\omega),$$

where

$$F(\omega) = \sum_{\omega_j \le \omega} \sigma_j^2 \quad (\omega \in [0, 2\pi]).$$

Here $e^{ik\omega}$ is integrated with respect to a discrete stochastic measure. By Theorem 3(b), this process is not ergodic for the mean, because dF has an atom σ_1^2 at $\{0\}$.

Theorem 3(b) shows that a time series is not ergodic for the mean if and only if it contains a component which is a nonzero time-constant random variable, like A_1 in the present example.

2 Estimation of the covariances

Let $\{\mathbf{X}_t\}_{t\in\mathbb{Z}}$ be a d-dimensional stationary time series with mean $\mathbf{0}$ and with covariance matrix function $\boldsymbol{C}(h) = [c_{jk}(h)]_{d\times d}$ $(h \in \mathbb{Z})$. For any positive integer T and for any $j, k = 1, \dots, d$, a natural estimator of a covariance $c_{jk}(h)$ is

$$\tilde{c}_{jk}(h) := \frac{1}{T-h} \sum_{t=0}^{T-1-h} X_{t+h}^j \overline{X_t^k} \quad (0 \le h \le T-1).$$

Clearly, this is an unbiased estimator:

$$\mathbb{E}(\tilde{c}_{jk}(h)) = \frac{1}{T-h} \sum_{t=0}^{T-h-1} \mathbb{E}(X_{t+h}^j \overline{X_t^k}) = c_{jk}(h) \quad (0 \le h \le T-1).$$

Another useful estimator is the *empirical covariance*

$$\hat{c}_{jk}(h) = \hat{c}_{jk}^{(T)}(h) := \frac{1}{T} \sum_{t=0}^{T-1-h} X_{t+h}^j \overline{X_t^k} \quad (0 \le h \le T-1). \tag{2}$$

This is only asymptotically unbiased if $0 \le h \le h_T = o(T)$ as $T \to \infty$, that is, $h_T/T \to 0$. However, the *empirical covariance matrix function*

$$\hat{\boldsymbol{C}}(h) := [\hat{c}_{jk}(h)]_{d \times d} = \frac{1}{T} \sum_{t=0}^{T-1-h} \mathbf{X}_{t+h} \mathbf{X}_t^* \quad (0 \le h \le T-1) \tag{3}$$

has the desirable property that the following $Td \times Td$ block Toeplitz matrix

$$\hat{\mathfrak{C}}_T := \begin{bmatrix} \hat{\boldsymbol{C}}(0) & \hat{\boldsymbol{C}}(1) & \cdots & \hat{\boldsymbol{C}}(T-1) \\ \hat{\boldsymbol{C}}(1)^* & \hat{\boldsymbol{C}}(0) & \cdots & \hat{\boldsymbol{C}}(T-2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\boldsymbol{C}}(T-1)^* & \hat{\boldsymbol{C}}(T-2)^* & \cdots & \hat{\boldsymbol{C}}(0) \end{bmatrix}$$

is non-negative definite for arbitrary $T \ge 1$. This follows from a factorization of $\hat{\mathfrak{C}}_T$. Consider the $Td \times 2Td$ matrix

$$\mathfrak{X}_T := \begin{bmatrix} \mathbf{X}_0 & \mathbf{X}_1 & \cdots & \mathbf{X}_{T-1} & 0 & 0 & \cdots & 0 & 0 \\ 0 & \mathbf{X}_0 & \cdots & \mathbf{X}_{T-2} & \mathbf{X}_{T-1} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{X}_0 & \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_{T-1} & 0 \end{bmatrix}.$$

Then $\hat{\mathfrak{C}}_T = T^{-1} \mathfrak{X}_T \mathfrak{X}_T^*$. Hence, for any $\mathbf{a}_1, \dots, \mathbf{a}_T \in \mathbb{C}^d$, we have

$$\sum_{j,k=1}^{T} \mathbf{a}_j^* \hat{\boldsymbol{C}}(j-k) \mathbf{a}_k = \mathbf{a}^* \frac{\hat{\mathfrak{C}}_T}{T} \mathbf{a} = \frac{1}{T} (\mathbf{a}^* \mathfrak{X}_T)(\mathbf{a}^* \mathfrak{X}_T)^* \ge 0,$$

where

$$\mathbf{a} := \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_T \end{bmatrix}.$$

In Lesson 1 we saw that in case of a weakly stationary process the large block Toeplitz matrix $\mathfrak{C}_T$ is positive semidefinite.

We mention that in the case when the mean $\boldsymbol{\mu}$ of the time series is not $\mathbf{0}$, the empirical covariances are defined as

$$\hat{c}_{jk}(h) = \hat{c}_{jk}^{(T)}(h) := \frac{1}{T} \sum_{t=0}^{T-1-h} (X_{t+h}^j - \tilde{X}_T^j)(\overline{X_t^k} - \overline{\tilde{X}_T^k}) \quad (0 \le h \le T-1),$$

where $\tilde{X}_T^j$ and $\tilde{X}_T^k$ are the empirical means of $\{X_t^j\}$ and $\{X_t^k\}$, respectively.

The weakly stationary time series $\{\mathbf{X}_t\}_{t\in\mathbb{Z}}$ is called *ergodic for the covariance* if it is ergodic for the mean, plus for each $j,k = 1,\dots,d$, the empirical covariance $\hat{c}_{jk}(h)$ converges in mean square to the covariance $c_{jk}(h)$ for any $0 \le h \le h_T$ as $T \to \infty$, where $h_T = o(T)$.

Thus fix $j,k \in \{1,\dots,d\}$ and an $h \in \mathbb{Z}$ from now on. Define the following subspace of $L^2(\Omega,\mathcal{F},\mathbb{P})$:

$$H_{X(j,k,h)} := \overline{\text{span}}\{X_{t+h}^j \bar{X}_t^k - c_{jk}(h) : t \in \mathbb{Z}\}.$$

So by defining the process

$$Y_t := X_{t+h}^j \bar{X}_t^k - c_{jk}(h), \qquad t \in \mathbb{Z}, \tag{4}$$

it follows that $\mathbb{E}(Y_t) = 0$ for any $t \in \mathbb{Z}$. If we also assume that $\mathbb{E}(Y_{t+s}\bar{Y}_t) = \mathbb{E}(Y_s\bar{Y}_0)$, equivalently,

$$\mathbb{E}\left(X_{t+s+h}^j \bar{X}_{t+s}^k \overline{X_{t+h}^j \bar{X}_t^k}\right) = \mathbb{E}\left(X_{s+h}^j \bar{X}_s^k \overline{X_h^j \bar{X}_0^k}\right) \quad \forall t,s \in \mathbb{Z}, \tag{5}$$

then $\{Y_t\}_{t\in\mathbb{Z}}$ becomes a weakly stationary time series with expectation 0. The right shift (forward shift) operator S is defined and unitary in the Hilbert subspace $H_{X(j,k,h)} = H_Y = \overline{\text{span}}\{Y_t : t \in \mathbb{Z}\}$ as well:

$$SY_0 = Y_t, \quad \text{that is,} \quad S(X_h^j \bar{X}_0^k) = X_{t+h}^j \bar{X}_t^k \quad \text{for all} \quad t \in \mathbb{Z}.$$

Similarly as in Subsection 1, we define

$$\mathcal{I}(j,k,h) = \{\xi \in H_{X(j,k,h)} : S\xi = \xi\}.$$

Theorem 4. *Assume that* $\{\mathbf{X}_t\}$ *is a* d*-dimensional complex weakly stationary time series with expectation* $\mathbf{0}$ *and property* (5) *holds for each* $j,k=1,\dots,d$ *and* $0 \le h \le h_T$*, where* $h_T = o(T)$*. Take the empirical covariance matrix function* $\hat{\boldsymbol{C}}(h) = [\hat{c}_{jk}^{(T)}(h)]_{d\times d}$ *defined by* (3) *for* $0 \le h \le h_T$*.*

(a) Then for each $j,k=1,\dots,d$ *and* $0 \le h \le h_T$*,*

$$\lim_{T\to\infty} \hat{c}_{jk}^{(T)}(h) - c_{jk}(h) = \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1-h} X_{t+h}^j \bar{X}_t^k - c_{jk}(h) = Y(j,k,h), \qquad (6)$$

where lim *denotes here limit in mean square and the random variable* $Y(j,k,h)$ *is the orthogonal projection of* $X_h^j \bar{X}_0^k - c_{jk}(h)$ *to the subspace* $\mathcal{I}(j,k,h)$*;* $\mathbb{E}Y(j,k,h) = 0$*.*

(b) The time series $\{\mathbf{X}_t\}$ *is ergodic for the covariance if and only if*

$$\lim_{T\to\infty} \frac{1}{T}\sum_{s=0}^{T-1-h} \mathbb{E}\left(X_{s+h}^j \bar{X}_s^k \overline{X_h^j \bar{X}_0^k}\right) = |c_{jk}(h)|^2, \qquad (7)$$

i.e., if $Y(j,k,h) = 0$ *for each* $j,k=1,\dots,d$ *and* $h \ge 0$*.*

3 Periodograms

First we extend the Discrete Fourier Transform (DFT) and its inverse IDFT to a random sample of a d-dimensional complex stationary time series $\{\mathbf{X}_t\}$ with $\mathbf{0}$ expectation. Assume that for a positive integer N, a *random sample* $\{\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_{T-1}\}$ is given. Then we define $\Delta\omega := \frac{2\pi}{T}$ and the Fourier frequencies $\omega_j = j\Delta\omega \in [0, 2\pi]$ $(j = 0, \dots, T-1)$. (In this subsection it is simpler to work with the frequency interval $[0, 2\pi]$ than with $[-\pi, \pi]$.)

By the DFT, one may assign d-dimensional *spectral amplitudes* $\hat{\mathbf{Z}}_{\omega_j}$ to the sample:

$$\hat{\mathbf{Z}}_{\omega_j} := T^{-\frac{1}{2}} \sum_{t=0}^{T-1} \mathbf{X}_t e^{-it\omega_j}, \quad j = 0, \dots, T-1. \qquad (8)$$

Then $\{\mathbf{X}_t\}$ can be obtained by IDFT from the spectral amplitudes:

$$\mathbf{X}_t = T^{-\frac{1}{2}} \sum_{j=0}^{T-1} \hat{\mathbf{Z}}_{\omega_j} e^{it\omega_j}, \quad t = 0, 1, \dots, T-1.$$

This is a discrete version of the spectral representation of the time series $\{\mathbf{X}_t\}$.

By definition, the *periodogram* of the sample is the sequence

$$\{\boldsymbol{I}_T(\omega_0), \boldsymbol{I}_T(\omega_1), \ldots, \boldsymbol{I}_T(\omega_{T-1})\}$$

of $d \times d$ *intensity matrices*:

$$\boldsymbol{I}_T(\omega_j) := \hat{\mathbf{Z}}_{\omega_j}\hat{\mathbf{Z}}^*_{\omega_j} = \frac{1}{T}\sum_{t,s=0}^{T-1} \mathbf{X}_t\mathbf{X}_s^* e^{-i(t-s)\omega_j}. \tag{9}$$

Substituting $t = s + h$ into (9) and rearranging the terms, it follows that

$$\boldsymbol{I}_T(\omega_j) = \sum_{h=-T+1}^{T-1} \frac{1}{T} \sum_{s=0}^{T-1-|h|} \mathbf{X}_{s+h}\mathbf{X}_s^* \, e^{-ih\omega_j} = \sum_{h=-T+1}^{T-1} \hat{\boldsymbol{C}}(h) \, e^{-ih\omega_j}, \tag{10}$$

where

$$\hat{\boldsymbol{C}}(h) = \frac{1}{T}\sum_{s=0}^{T-1-|h|} \mathbf{X}_{s+h}\mathbf{X}_s^* \qquad (h = -T+1, \ldots, T-1), \tag{11}$$

is the empirical covariance matrix function of the sample.

Assume that the spectral measure of the time series $\{\mathbf{X}_t\}$ is absolutely continuous with density matrix $\boldsymbol{f}$. Then comparing the DFT with (10) shows that

$$\hat{\boldsymbol{f}}(\omega_j) := \frac{1}{2\pi}\boldsymbol{I}_T(\omega_j), \quad \omega_j = j\frac{2\pi}{T}, \quad j = 0, 1, \ldots, T-1, \tag{12}$$

is a discrete estimate of the spectral density $\boldsymbol{f}(\omega)$, $\omega \in [0, 2\pi)$. The next proposition shows that estimate (12) is *asymptotically unbiased.*

Proposition 1. *Suppose that $\boldsymbol{C}(h)$, $h \in \mathbb{Z}$, is absolutely summable. For $\omega \in [0, 2\pi)$ set $\omega^{(T)} := \lfloor \omega/\Delta\omega \rfloor \Delta\omega$, where $\Delta\omega := \frac{2\pi}{T}$. Then for the estimate* (12) *we have*

$$\lim_{T\to\infty} \frac{1}{2\pi}\mathbb{E}\boldsymbol{I}_T(\omega^{(T)}) = \boldsymbol{f}(\omega),$$

uniformly in $\omega \in [0, 2\pi)$.

4 Spectra of spectra

Let $\{\mathbf{X}_t\}$ be a d-dimensional, weakly stationary time series with real components and autocovariance matrices $\boldsymbol{C}(h)$, $h \in \mathbb{Z}$, $\boldsymbol{C}(-h) = \boldsymbol{C}^T(h)$. Consider the finite segment $\mathbf{X}_1, \dots, \mathbf{X}_n \in \mathbb{R}^d$ of it and the $nd \times nd$ covariance matrix $\mathfrak{C}_n$ of the compounded random vector $[\mathbf{X}_1^T, \dots, \mathbf{X}_n^T]^T \in \mathbb{R}^{nd}$, as introduced in Lesson 1. This is a symmetric, positive semidefinite block-Toeplitz matrix, the (i, j) block of which is $\boldsymbol{C}(j-i)$. The symmetry comes from the fact, that the (j, i) entry is $\boldsymbol{C}(i-j) = \boldsymbol{C}^T(j-i)$. To characterize its eigenvalues, first we need the symmetric block circulant matrix $\mathfrak{C}_n^{(s)}$

$$\mathfrak{C}_n^{(s)} := \begin{bmatrix} \boldsymbol{C}(0) & \boldsymbol{C}(1) & \boldsymbol{C}(2) & \dots & \boldsymbol{C}(2) & \boldsymbol{C}(1) \\ \boldsymbol{C}^T(1) & \boldsymbol{C}(0) & \boldsymbol{C}(1) & \dots & \boldsymbol{C}(3) & \boldsymbol{C}(2) \\ \boldsymbol{C}^T(2) & \boldsymbol{C}^T(1) & \boldsymbol{C}(0) & \dots & \boldsymbol{C}(4) & \boldsymbol{C}(3) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{C}^T(1) & \boldsymbol{C}^T(2) & \boldsymbol{C}^T(3) & \dots & \boldsymbol{C}(1) & \boldsymbol{C}(0) \end{bmatrix}. \tag{13}$$

that we consider now for odd n, say $n = 2k+1$. The (i, j) block of $\mathfrak{C}_n^{(s)}$ for $1 \le i \le j \le n$ is

$$\mathfrak{C}_n^{(s)}(\text{block}_i, \text{block}_j) = \begin{cases} \boldsymbol{C}(j-i) & j-i \le k \\ \boldsymbol{C}(n-(j-i)), & j-i > k. \end{cases}$$

For $i > j$, it is

$$\mathfrak{C}_n^{(s)}(\text{block}_i, \text{block}_j) = \begin{cases} \boldsymbol{C}^T(i-j) & i-j \le k \\ \boldsymbol{C}^T(n-(i-j)), & i-j > k. \end{cases}$$

In this way, $\mathfrak{C}_n^{(s)}$ is a symmetric block Toeplitz matrix again, and it is the same as $\mathfrak{C}_n$ within the blocks (i, j)s for which $|j-i| \le k$ holds. For example, if $n = 7$ and $k = 3$, then we have

$$\mathfrak{C}_7^{(s)} := \begin{bmatrix} \boldsymbol{C}(0) & \boldsymbol{C}(1) & \boldsymbol{C}(2) & \boldsymbol{C}(3) & \boldsymbol{C}(3) & \boldsymbol{C}(2) & \boldsymbol{C}(1) \\ \boldsymbol{C}^T(1) & \boldsymbol{C}(0) & \boldsymbol{C}(1) & \boldsymbol{C}(2) & \boldsymbol{C}(3) & \boldsymbol{C}(3) & \boldsymbol{C}(2) \\ \boldsymbol{C}^T(2) & \boldsymbol{C}^T(1) & \boldsymbol{C}(0) & \boldsymbol{C}(1) & \boldsymbol{C}(2) & \boldsymbol{C}(3) & \boldsymbol{C}(3) \\ \boldsymbol{C}^T(3) & \boldsymbol{C}^T(2) & \boldsymbol{C}^T(1) & \boldsymbol{C}(0) & \boldsymbol{C}(1) & \boldsymbol{C}(2) & \boldsymbol{C}(3) \\ \boldsymbol{C}^T(3) & \boldsymbol{C}^T(3) & \boldsymbol{C}^T(2) & \boldsymbol{C}^T(1) & \boldsymbol{C}(0) & \boldsymbol{C}(1) & \boldsymbol{C}(2) \\ \boldsymbol{C}^T(2) & \boldsymbol{C}^T(3) & \boldsymbol{C}^T(3) & \boldsymbol{C}^T(2) & \boldsymbol{C}^T(1) & \boldsymbol{C}(0) & \boldsymbol{C}(1) \\ \boldsymbol{C}^T(1) & \boldsymbol{C}^T(2) & \boldsymbol{C}^T(3) & \boldsymbol{C}^T(3) & \boldsymbol{C}^T(2) & \boldsymbol{C}^T(1) & \boldsymbol{C}(0) \end{bmatrix}.$$

In the 1D case, we simply have the $n \times n$ positive semidefinite matrix $\boldsymbol{C}_n$ and the symmetric circulant matrix $\boldsymbol{C}_n^{(s)}$ with the autocovariances $c(h)$s, $h \in \mathbb{Z}$. By Kronecker products (with permutation matrices) it is well known that the jth eigenvalue of $\boldsymbol{C}_n^{(s)}$ is $\sum_{h=0}^{n-1} c(h)\rho_j^h$, where $\rho_j = e^{i\omega_j}$ is the jth primitive (complex) nth root of 1 and $\omega_j = \frac{2\pi j}{n}$ is the jth Fourier frequency ($j = 0, 1, \dots, n-1$). Further, the eigenvector corresponding to the jth eigenvalue is $(1, \rho_j, \dots, \rho_j^{n-1})^T$; it has norm $\sqrt{n}$. After normalizing with $\frac{1}{\sqrt{n}}$, we get a complete orthonormal set of eigenvectors (of complex coordinates).

When $\boldsymbol{C}(h)$s are $d \times d$ matrices, by inflation techniques and applying Kronecker products, we use blocks instead of entries and the eigenvectors also follow a block structure. By Friedman (1961) and Tee (2007), the eigenvalues and eigenvectors of a general symmetric block circulant matrix are characterized. We apply this result in our situation, when $n = 2k+1$ is odd (for even n similar results hold). Therefore, the spectrum of $\mathfrak{C}_n^{(s)}$ is the union of spectra of the matrices

$$\boldsymbol{M}_j = \boldsymbol{C}(0) + \sum_{h=1}^{k} [\boldsymbol{C}(h)\rho_j^h + \boldsymbol{C}^T(h)\rho_j^{-h}] = \boldsymbol{C}(0) + \sum_{h=1}^{k} [\boldsymbol{C}(h)e^{i\omega_j h} + \boldsymbol{C}^T(h)e^{-i\omega_j h}] \tag{14}$$

for $j = 0, 2, \dots, n-1$, whereas the eigenvectors are obtained by compounding the eigenvectors of these $d \times d$ matrices. So we need the spectral decomposition of the matrices

$$\boldsymbol{M}_0 = \boldsymbol{C}(0) + \sum_{h=1}^{k} [\boldsymbol{C}(h) + \boldsymbol{C}^T(h)]$$

and

$$\boldsymbol{M}_j = \boldsymbol{C}(0) + \sum_{h=1}^{k} [(\boldsymbol{C}(h) + \boldsymbol{C}^T(h))\cos(\omega_j h) + i(\boldsymbol{C}(h) - \boldsymbol{C}^T(h))\sin(\omega_j h)]$$

for $j = 0, 2, \dots, n-1$. Since $\boldsymbol{C}(h) + \boldsymbol{C}^T(h)$ is symmetric and $\boldsymbol{C}(h) - \boldsymbol{C}^T(h)$ is anti-symmetric with 0 diagonal, $\boldsymbol{M}_j$ is self-adjoint for each j and has real eigenvalues with corresponding orthonormal set of eigenvectors of possibly complex coordinates. Indeed, $\boldsymbol{M}_j$ may have complex entries if $j \neq 0$; actually, $\sum_{h=1}^{k} (\boldsymbol{C}(h) + \boldsymbol{C}^T(h))\cos(\omega_j h)$ is the real and $\sum_{h=1}^{k} (\boldsymbol{C}(h) - \boldsymbol{C}^T(h))\sin(\omega_j h)$ is the imaginary part of $\boldsymbol{M}_j$.

It is easy to see that $\boldsymbol{M}_{n-j} = \overline{\boldsymbol{M}_j}$ (entrywise conjugate), therefore, it has the same eigenvalues as $\boldsymbol{M}_j$, but the eigenvectors are the (componentwise) complex conjugates of the eigenvectors of $\boldsymbol{M}_j$. We also need the following form of this matrix:

$$\begin{aligned}\boldsymbol{M}_{n-j} &= \boldsymbol{C}(0) + \sum_{h=1}^{k}[(\boldsymbol{C}(h) + \boldsymbol{C}^T(h))\cos(\omega_j h) - i(\boldsymbol{C}(h) - \boldsymbol{C}^T(h))\sin(\omega_j h)] \\ &= \boldsymbol{C}(0) + \sum_{h=1}^{k}[\boldsymbol{C}(h)e^{-i\omega_j h} + \boldsymbol{C}^T(h)e^{i\omega_j h}], \quad j = 1, \dots, n-1.\end{aligned} \tag{15}$$

Summarizing, for odd $n = 2k+1$, the nd eigenvalues of $\mathfrak{C}_n^{(s)}$ are obtained as the union of the eigenvalues of $\boldsymbol{M}_0$ and those of $\boldsymbol{M}_j$ ($j = 1, \dots, k$) duplicated. Note that for even n similar arguments hold with the difference that there the spectrum of $\mathfrak{C}_n^{(s)}$ is the union of the eigenvalues of $\boldsymbol{M}_0$ and $\boldsymbol{M}_{n-1}$, whereas the eigenvalues of $\boldsymbol{M}_1, \dots, \boldsymbol{M}_{\frac{n}{2}-1}$ are duplicated.

The eigenvectors of $\mathfrak{C}_n^{(s)}$ are obtainable by compounding the d orthonormal eigenvectors of the $d \times d$ self-adjoint matrices $\boldsymbol{M}_0, \boldsymbol{M}_1, \dots, \boldsymbol{M}_{n-1}$ as follows. For $j = 1, \dots, k$: if $\mathbf{v}$ is an eigenvector of $\boldsymbol{M}_j$ with eigenvalue λ, then in Tee(2007) it is proved that the compound vector

$$\mathbf{w} = (\mathbf{v}^T, \rho_j \mathbf{v}^T, \rho_j^2 \mathbf{v}^T, \dots, \rho_j^{n-1}\mathbf{v}^T)^T \in \mathbb{C}^{nd} \tag{16}$$

is an eigenvector of $\mathfrak{C}_n^{(s)}$ with the same eigenvalue λ. Further, if

$$\mathbf{z} = (\mathbf{t}^T, \rho_\ell \mathbf{t}^T, \rho_\ell^2 \mathbf{t}^T, \dots, \rho_\ell^{n-1}\mathbf{t}^T)^T \in \mathbb{C}^{nd}$$

is another eigenvector of $\mathfrak{C}_n^{(s)}$ compounded from an eigenvector $\mathbf{t}$ of another $\boldsymbol{M}_\ell$ ($\ell \neq j$), then $\mathbf{w}$ and $\mathbf{z}$ are orthogonal, irrespective whether $\boldsymbol{M}_\ell$ has the same eigenvalue λ as $\boldsymbol{M}_j$ or not. Similar construction holds starting with the eigenvectors of $\boldsymbol{M}_0$.

Here for each $j = 0, 1, \dots, n-1$, there are d pairwise orthogonal eigenvectors (potential $\mathbf{v}$s) of $\boldsymbol{M}_j$, and the so obtained $\mathbf{w}$s are also pairwise orthogonal. Assume that the eigenvectors of $\boldsymbol{M}_j$ are enumerated in non-increasing order of its eigenvalues, and the inflated $\mathbf{w}$s also follow this ordering, for $j = 0, 1, \dots, n-1$.

As we saw, if $\mathbf{v}$ is an eigenvector of $\boldsymbol{M}_j$ with real eigenvalue λ, then $\overline{\mathbf{v}}$ is the corresponding eigenvector of $\boldsymbol{M}_{n-j}$ with the same eigenvalue λ;

further, the compounded $\mathbf{w}$ and $\overline{\mathbf{w}} \in \mathbb{C}^{nd}$ are orthogonal eigenvectors of $\mathfrak{C}_n^{(s)}$ corresponding to the eigenvalue λ with multiplicity (at least) two; $\mathbf{w}$ and $\overline{\mathbf{w}}$ have the same norm. From them, corresponding to this double eigenvalue λ, the new orthogonal pair of eigenvectors

$$\frac{\mathbf{w}+\overline{\mathbf{w}}}{2} \quad \text{and} \quad i\frac{\mathbf{w}-\overline{\mathbf{w}}}{2} \tag{17}$$

is constructed, but they, in this order, occupy the original positions of $\mathbf{w}$ and $\overline{\mathbf{w}}$. Note that it is necessary to have an orthogonal system of eigenvectors with real coordinates whenever the underlying time series is real, and so, $\mathfrak{C}_n^{(s)}$ is a real symmetric matrix. We do not go in details, neither discuss defective cases.

After normalization, denote by $\mathbf{u}_1, \ldots, \mathbf{u}_{nd}$ the so obtained orthonormal set of eigenvectors (of real coordinates) of $\mathfrak{C}_n^{(s)}$ (in the above ordering) and by $\boldsymbol{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_{nd})$ the $nd \times nd$ orthogonal matrix containing them columnwise; further, let

$$\mathfrak{C}_n^{(s)} = \boldsymbol{U}\boldsymbol{\Lambda}^{(s)}\boldsymbol{U}^T \tag{18}$$

be the corresponding spectral decomposition. After this preparation, we are able to prove the following theorem.

Theorem 5. *Let $\{\mathbf{X}_t\}$ be d-dimensional weakly stationary time series of real components. Denoting by $\boldsymbol{C}(h) = [c_{ij}(h)]$ the $d \times d$ autocovariance matrices ($\boldsymbol{C}(-h) = \boldsymbol{C}^T(h)$, $h \in \mathbb{Z}$) in the time domain, assume that their entries are absolutely summable, i.e., $\sum_{h=0}^{\infty} |c_{pq}(h)| < \infty$ for $p, q = 1, \ldots, d$. Then, the self-adjoint, positive semidefinite spectral density matrix $\boldsymbol{f}(\omega)$ exists in the frequency domain, and it is defined by*

$$\boldsymbol{f}(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \boldsymbol{C}(h) e^{-ih\omega}, \quad \omega \in [0, 2\pi].$$

For odd $n = 2k+1$, consider $\mathbf{X}_1, \ldots \mathbf{X}_n$ with the block Toeplitz matrix $\mathfrak{C}_n$; further, the Fourier frequencies $\omega_j = \frac{2\pi j}{n}$ for $j = 0, \ldots, n-1$. Let $\boldsymbol{D}_n$ be the $dn \times dn$ diagonal matrix that contains the spectra of the matrices $\boldsymbol{f}(0), \boldsymbol{f}(\omega_1), \boldsymbol{f}(\omega_2), \ldots, \boldsymbol{f}(\omega_k), \boldsymbol{f}(\omega_k), \ldots, \boldsymbol{f}(\omega_2), \boldsymbol{f}(\omega_1)$ in its main diagonal, i.e.,

$$\boldsymbol{D}_n = \operatorname{diag}(\operatorname{spec} \boldsymbol{f}(0), \operatorname{spec} \boldsymbol{f}(\omega_1), \ldots, \operatorname{spec} \boldsymbol{f}(\omega_k), \operatorname{spec} \boldsymbol{f}(\omega_k), \ldots, \operatorname{spec} \boldsymbol{f}(\omega_1)).$$

Here spec *contains the eigenvalues of the affected matrix in non-increasing order if not otherwise stated. (The duplication is due to the fact that* $\boldsymbol{f}(\omega_j) = \boldsymbol{f}(\omega_{n-j})$*,* $j = 1, \dots, k$*, for real time series). Then, with the spectral decomposition (18),*

$$\boldsymbol{U}^T \mathfrak{C}_n \boldsymbol{U} - 2\pi \boldsymbol{D}_n \to \boldsymbol{O}, \quad n \to \infty,$$

i.e., the entries of the matrix $\boldsymbol{U}^T \mathfrak{C}_n \boldsymbol{U} - 2\pi \boldsymbol{D}_n$ *tend to 0 uniformly as* $n \to \infty$*.*

Proof. We saw that $\boldsymbol{U}^T \mathfrak{C}_n^{(s)} \boldsymbol{U} = \boldsymbol{\Lambda}^{(s)}$. Recall that the eigenvalues in the diagonal of $\boldsymbol{\Lambda}^{(s)}$ comprise the union of spectra of the matrices $\boldsymbol{M}_0$ and those of $\boldsymbol{M}_1, \dots, \boldsymbol{M}_{n-1}$, which are the same as the eigenvalues of $\boldsymbol{M}_0$ and those of $\boldsymbol{M}_{n-1}, \dots, \boldsymbol{M}_{n-k}$ of (15), duplicated. But these matrices are finite sub-sums (for $|h| \le k$) of the infinite summations

$$2\pi \boldsymbol{f}(\omega_j) = \sum_{h=-\infty}^{\infty} \boldsymbol{C}(h) e^{-ih\omega} = \boldsymbol{C}(0) + \sum_{h=1}^{\infty} [\boldsymbol{C}(h) e^{-i\omega_j h} + \boldsymbol{C}^T(h) e^{i\omega_j h}],$$

so (by the continuity of the spectra), the pairwise distances between the eigenvalues of $\boldsymbol{M}_j$ and the corresponding eigenvalues of $2\pi \boldsymbol{f}(\omega_j)$ (both in non-increasing order) tend to 0 as $n \to \infty$, for $j = 0, 1, \dots, k$. Here we used the absolute summability of the entries of $\boldsymbol{C}(h)$s, which fact implies that the diagonal entries of the diagonal matrix $\boldsymbol{\Lambda}^{(s)} - 2\pi \boldsymbol{D}_n$ are bounded in absolute value by

$$\max_{p,q \in \{1,\dots,d\}} \sum_{|h|>k} |c_{pq}(h)| \to 0, \quad n = 2k+1 \to \infty.$$

So the matrix $\boldsymbol{\Lambda}^{(s)} - 2\pi \boldsymbol{D}_n$ tends to the zero matrix entrywise as $n \to \infty$. Therefore, it remains to show that the entries of $\mathbf{U}^T \mathfrak{C}_n \mathbf{U} - \mathbf{U}^T \mathfrak{C}_n^{(s)} \mathbf{U}$ tend to 0 uniformly as $n \to \infty$.

Before doing this, some facts should be clarified.

- The pth row sum of $\boldsymbol{M}_j$ is bounded by

$$\sum_{q=1}^{d} |c_{pq}(0)| + \sum_{q=1}^{d} \sum_{h=1}^{k} |c_{pq}(h)| + \sum_{q=1}^{d} \sum_{h=1}^{k} |c_{qp}(h)| \le d c_{pp}(0) + 2dL,$$

  for $p \in \{1, \dots, d\}$ with $L = \max_{p,q \in \{1,\dots,d\}} \sum_{h=1}^{\infty} |c_{pq}(h)| > 0$, independently of n, because of the absolute summability of the entries of $\boldsymbol{C}(h)$.

This is true for any $j \in \{0, 1, \ldots, n-1\}$. For simplicity, consider (any) one of the $\boldsymbol{M}_j$s, and denote it by $\boldsymbol{M} = [m_{pq}]_{p,q=1}^d$. Then

$$\|\boldsymbol{M}\|_\infty = \max_{p \in \{1,\ldots,d\}} \sum_{q=1}^d |m_{pq}| \le d \max_{p \in \{1,\ldots,d\}} c_{pp}(0) + 2dL = K.$$

As the spectral radius of $\boldsymbol{M}$ is at most $\|\boldsymbol{M}\|_\infty$, any eigenvalue λ of $\boldsymbol{M}$ is bounded in absolute value by K (independenty of n).

- Let $\mathbf{v}$ be an eigenvector of $\boldsymbol{M}_j$ with eigenvalue λ, we can assume that $\|\mathbf{v}\| = \sqrt{\mathbf{v}^*\mathbf{v}} = 1$. Then the vector $\mathbf{w}$ in Equation (16) is an eigenvector of $\mathfrak{C}_n^{(s)}$. Since

  $$\mathbf{w}^*\mathbf{w} = \mathbf{v}^*\mathbf{v}(1 + \rho_j\rho_j^{-1} + \rho_j^2\rho_j^{-2} + \cdots + \rho_j^{n-1}\rho_j^{-(n-1)}) = n,$$

  the (complex) vector $\frac{1}{\sqrt{n}}\mathbf{w}$ will have unit norm. Further, by transformation (17), the coordinates of any (real) unit-norm eigenvector $\mathbf{u}$ are bounded by $\sqrt{\frac{2}{n}}$ in absolute value.

Now we are ready to show that

$$|\mathbf{u}_i^T \mathfrak{C}_n^{(s)} \mathbf{u}_j - \mathbf{u}_i^T \mathfrak{C}_n \mathbf{u}_j| \to 0, \quad n \to \infty$$

uniformly in $i, j \in \{1, \ldots, nd\}$. Recall that in the $nd \times nd$ matrices $\mathfrak{C}_n^{(s)}$ and $\mathfrak{C}_n$ the (m, ℓ) blocks are the same if $|m - \ell| \le k$. Denote by $\mathbf{u}_{i,m}$ and $\mathbf{u}_{j,\ell}$ the mth and ℓth blocks of the unit-norm eigenvectors $\mathbf{u}_i$ and $\mathbf{u}_j$, respectively. Recall (see their description preceding the theorem) that they both were compounded from n vectors of length d, and their coordinates are bounded by $\sqrt{\frac{2}{n}}$ in absolute value. Then

$$
\begin{aligned}
|\mathbf{u}_i^T(\mathfrak{C}_n^{(s)} - \mathfrak{C}_n)\mathbf{u}_j| &= 2\left|\sum_{m=1}^{k}\sum_{\ell=1}^{m} \mathbf{u}_{i,\ell}^T(\boldsymbol{C}(m) - \boldsymbol{C}(n-m))\mathbf{u}_{j,n-m+\ell}\right| \\
&\le 2\sqrt{\frac{2}{n}}\left|\sum_{m=1}^{k} \mathbf{1}_d^T(\boldsymbol{C}(m) - \boldsymbol{C}(n-m))\sum_{\ell=1}^{m}\mathbf{u}_{j,n-m+\ell}\right| \\
&\le 2\sqrt{\frac{2}{n}}\sqrt{\frac{2}{n}}\left|\sum_{m=1}^{k} m\mathbf{1}_d^T(\boldsymbol{C}(m) - \boldsymbol{C}(n-m))\mathbf{1}_d\right| \\
&\le \frac{4}{n}\left(\sum_{m=1}^{k} m\sum_{p=1}^{d}\sum_{q=1}^{d}|c_{pq}(m)| + \sum_{m=1}^{k} m\sum_{p=1}^{d}\sum_{q=1}^{d}|c_{pq}(n-m)|\right) \\
&\le 4d^2\left(\max_{p,q\in\{1,\dots,d\}}\sum_{m=1}^{k}\frac{m}{n}|c_{pq}(m)| + \max_{p,q\in\{1,\dots,d\}}\sum_{m=1}^{k}\frac{m}{n}|c_{pq}(n-m)|\right) \\
&\le 4d^2\left(\max_{p,q\in\{1,\dots,d\}}\sum_{m=1}^{k}\frac{m}{n}|c_{pq}(m)| + \max_{p,q\in\{1,\dots,d\}}\sum_{m=n-k}^{n-1}\frac{k}{n}|c_{pq}(m)|\right),
\end{aligned}
$$

where $\mathbf{1}_d \in \mathbb{R}^d$ is the vector of all 1 coordinates and so, the quadratic form $\mathbf{1}_d^T(\boldsymbol{C}(m) - \boldsymbol{C}(n-m))\mathbf{1}_d$ is the sum of the entries of $\boldsymbol{C}(m) - \boldsymbol{C}(n-m)$. In the last line, the second term converges to 0, since it is bounded by $\sum_{m=k}^{\infty}|c_{pq}(m)|$, and together with n, k tends to ∞ too; further, it holds uniformly for all $p, q \in \{1, \dots, d\}$. The first term for every p, q pair also tends to 0 as $n \to \infty$ by the discrete version of the dominated convergence theorem (for series), see the forthcoming Lemma 1. Indeed, the summand is dominated by $|c_{pq}(m)|$ and $\sum_{m=1}^{\infty}|c_{pq}(m)| < \infty$; further, $\frac{m}{n}|c_{pq}(m)| \to 0$ as $n \to \infty$, for any fixed m. Consequently, $\sum_{m=1}^{\infty}\frac{m}{n}|c_{pq}(m)|$ tends to 0, and so does $\sum_{m=1}^{k}\frac{m}{n}|c_{pq}(m)|$ as $n \to \infty$. It holds uniformly for all p, q, and also for all i, j, so the proof is complete. □

Lemma 1 (Dominated convergence theorem for sums, discrete version). *Consider* $\sum_{m=1}^{\infty} f_n(m)$ *and Assume that* $|f_n(m)| \le g(m)$ *with* $\sum_{m=1}^{\infty} g(m) < \infty$. *If* $\lim_{n\to\infty} f_n(m) = f(m)$ *exists* $\forall m \in \mathbb{N}$, *then*

$$\lim_{n\to\infty}\sum_{m=1}^{\infty} f_n(m) = \sum_{m=1}^{\infty} f(m).$$

Remark 2. *To perform Principal Component Analysis (PCA), we use the eigenvectors in the columns of $\boldsymbol{W}$ (of complex coordinates) in the ordering described preceding the theorem; assume that they are already normalized to have a complete orthonormal system in $\mathbb{C}^{nd}$. Let $\mathbf{Z} = (\mathbf{Z}_1^T, \dots, \mathbf{Z}_n^T)^T$ denote the random vector obtained by*

$$\mathbf{Z} = \boldsymbol{W}^* \mathbf{X}.$$

Its (complex) components are also uncorrelated and $\mathbb{E}\mathbf{Z}\mathbf{Z}^* \sim 2\pi \boldsymbol{D}_n$ again. Instead, we consider the blocks $\mathbf{Z}_j$s of it, and perform a 'partial principal component transformation' (in d-dimension) of them. Let $\mathbf{w}_{1j}, \dots, \mathbf{w}_{dj}$ be the columns of $\mathbf{W}$ corresponding to the coordinates of $\mathbf{Z}_j$. In view of (16), $\mathbf{Z}_j$ can be written as

$$\mathbf{Z}_j = \frac{1}{\sqrt{n}} (\boldsymbol{V}_j^* \otimes \mathbf{r}^*) \mathbf{X},$$

where $\mathbf{r}^* = (1, \rho_j^{-1}, \rho_j^{-2}, \dots, \rho_j^{-(n-1)})$ and $\boldsymbol{V}_j$ is the $d \times d$ unitary matrix in the spectral decomposition $\boldsymbol{M}_j = \boldsymbol{V}_j \boldsymbol{\Lambda}_j \boldsymbol{V}_j^*$. Because of $\mathbb{E}\mathbf{Z}_j \mathbf{Z}_j^* = \boldsymbol{\Lambda}_j$ (apparently from the proof of Theorem 5), we have that

$$\mathbb{E}(\boldsymbol{V}_j \mathbf{Z}_j)(\boldsymbol{V}_j \mathbf{Z}_j)^* = \boldsymbol{V}_j \boldsymbol{\Lambda}_j \boldsymbol{V}_j^* = \boldsymbol{M}_j.$$

At the same time,

$$\boldsymbol{V}_j \mathbf{Z}_j = \frac{1}{\sqrt{n}} \boldsymbol{V}_j (\boldsymbol{V}_j^* \otimes \mathbf{r}^*) \mathbf{X} = \frac{1}{\sqrt{n}} (\boldsymbol{I}_d \otimes \mathbf{r}^*) \mathbf{X} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \mathbf{X}_t e^{-it\omega_j}, \quad j = 1, \dots, n.$$

This is the formal DFT of $\mathbf{X}_1, \dots, \mathbf{X}_n$, see (8).

It is also in accord with the definition of the orthogonal increment process $\{\mathbf{Z}_\omega\}$ of which $\boldsymbol{V}_j \mathbf{Z}_j \sim \mathbf{Z}_{\omega_j}$ is the discrete analogue. Also, $\mathbf{Z}_1, \dots \mathbf{Z}_n$ are asymptotically pairwise orthogonal akin to $\boldsymbol{V}_1 \mathbf{Z}_1, \dots, \boldsymbol{V}_n \mathbf{Z}_n$. Further,

$$\mathbb{E}(\boldsymbol{V}_j \mathbf{Z}_j)(\boldsymbol{V}_j \mathbf{Z}_j)^* \sim 2\pi \boldsymbol{f}(\omega_j),$$

and it is in accord with the fact that

$$\mathbb{E}\mathbf{Z}_j \mathbf{Z}_j^* \sim 2\pi \operatorname{diag} \operatorname{spec} \boldsymbol{f}(\omega_j),$$

for $j = 1, \dots, n$.

To find the best k-rank approximation of the original process, the d-dimensional vectors $\boldsymbol{V}_j \mathbf{Z}_j$s, obtained by DFT, should be projected onto the *subspace spanned by the k leading eigenvectors of $\boldsymbol{V}_j$. Assume that the eigenvalues in $\boldsymbol{\Lambda}_j$ are in non-increasing order. Let us denote the k leading eigenvectors by $\mathbf{v}_{j1}, \dots, \mathbf{v}_{jk}$. Then*

$$(\boldsymbol{V}_j \mathbf{Z}_j)^{(k)} = \mathrm{Proj}_{\mathrm{Span}\{\mathbf{v}_{j1},\dots,\mathbf{v}_{jk}\}} \boldsymbol{V}_j \mathbf{Z}_j = \sum_{\ell=1}^{k} (\mathbf{v}_{j\ell}^* \boldsymbol{V}_j \mathbf{Z}_j) \mathbf{v}_{j\ell} = \sum_{\ell=1}^{k} Z_{j\ell} \mathbf{v}_{j\ell},$$

where $Z_{j\ell}$ denotes the ℓth coordinate of $\mathbf{Z}_j$.

This transformation gives rise to dimension reduction in the frequency domain, then via IDFT (due to $\mathbf{X} = \boldsymbol{W}\mathbf{Z}$), in the time domain too:

$$(\mathbf{X}_t)^{(k)} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (\boldsymbol{V}_j \mathbf{Z}_j)^{(k)} e^{it\omega_j} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left(\sum_{\ell=1}^{k} Z_{j\ell} \mathbf{v}_{j\ell} \right) e^{it\omega_j}, \quad t = 1, \dots, n.$$

Proposition 2. *Theorem 5 implies the following. Assume that for the spectra of the spectral densities $\boldsymbol{f}$ of the d-dimensional weakly stationary process $\{\mathbf{X}_t\}$ of real coordinates the following hold:*

$$m := \inf_{\omega \in [0,2\pi], q \in \{1,\dots,d\}} \lambda_q(\boldsymbol{f}(\omega)) > 0,$$
$$M := \sup_{\omega \in [0,2\pi], q \in \{1,\dots,d\}} \lambda_q(\boldsymbol{f}(\omega)) < \infty.$$

(Note that under the conditions of Theorem 5, $\boldsymbol{f}(\omega) > 0$ and it is continuous almost everywhere on $[0, 2\pi]$, so the above conditions are readily satisfied.)

Then for the eigenvalues $\lambda_1 \le \lambda_2 \le \dots \le \lambda_{nd}$ of the block Toeplitz matrix $\mathfrak{C}_n$ the following holds:

$$2\pi m \le \lambda_1 \le \lambda_{nd} \le 2\pi M.$$

5 Summary

Based on the finite set $\mathbf{X}_1, \dots, \mathbf{X}_T$ of observations, the parameters of the process are estimated as follows:

$$\hat{\boldsymbol{C}}(h) = \begin{cases} \frac{1}{T}\sum_{t=1}^{T-h}(\mathbf{X}_{t+h} - \bar{\mathbf{X}}_T)(\mathbf{X}_t - \bar{\mathbf{X}}_T)^*, & 0 \le h \le T-1 \\ \hat{\boldsymbol{C}}^*(-h), & -T+1 \le h < 0, \end{cases}$$

while $\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}}_T = \frac{1}{T}\sum_{t=1}^{T}\mathbf{X}_t$. In practice, we usually estimate from a single trajectory ($T \to \infty$), so ergodicity is of distinguished importance. It is proved that $\bar{\mathbf{X}}_T$ is ergodic for the mean if and only if $\boldsymbol{F}(\{0\}) = 0$, that is, the spectral measure has no atom at $\{0\}$. It surely holds if $d\boldsymbol{F}$ is absolutely continuous w.r.t. the Lebesgue measure (that is, has a spectral density matrix $\boldsymbol{f}$).

The time series $\{\mathbf{X}_t\}$ is called *ergodic for the covariance* if it is ergodic for the mean, and in addition, for each $j, k = 1, \dots, d$, the empirical covariance $\hat{c}_{jk}(h)$ converges in mean square to the true covariance $c_{jk}(h)$ for any $0 \le h \le h_T$ as $T \to \infty$, where $h_T = o(T)$. A necessary and sufficient condition for this is described in Theorem 4.

By DFT, we can assign d-dimensional *spectral amplitudes* $\hat{\mathbf{Z}}_{\omega_j}$ to the sample:

$$\hat{\mathbf{Z}}_{\omega_j} := T^{-\frac{1}{2}}\sum_{t=0}^{T-1}\mathbf{X}_t e^{-it\omega_j}, \quad j = 0, \dots, T-1,$$

where $\omega_j = j\frac{2\pi}{T}$ is the jth Fourier frequency, $j = 0, 1, \dots, T-1$.

Conversely, $\{\mathbf{X}_t\}$ can be obtained by IDFT from the spectral amplitudes:

$$\mathbf{X}_t = T^{-\frac{1}{2}}\sum_{j=0}^{T-1}\hat{\mathbf{Z}}_{\omega_j} e^{it\omega_j}, \quad t = 0, 1, \dots, T-1.$$

This is a discrete version of the spectral representation of the time series $\{\mathbf{X}_t\}$.

The *periodogram* of the sample is the sequence

$$\{\boldsymbol{I}_T(\omega_0), \boldsymbol{I}_T(\omega_1), \dots, \boldsymbol{I}_T(\omega_{T-1})\}$$

of the $d \times d$ *intensity matrices*

$$\boldsymbol{I}_T(\omega_j) := \hat{\mathbf{Z}}_{\omega_j}\hat{\mathbf{Z}}_{\omega_j}^* = \frac{1}{T}\sum_{t,s=0}^{T-1}\mathbf{X}_t\mathbf{X}_s^* e^{-i(t-s)\omega_j}.$$

Also,

$$\boldsymbol{I}_T(\omega_j) = \sum_{h=-T+1}^{T-1} \hat{\boldsymbol{C}}(h)\, e^{-ih\omega_j}.$$

So $\hat{\boldsymbol{f}}(\omega_j) := \frac{1}{2\pi}\boldsymbol{I}_T(\omega_j)$ is a discrete estimate of the spectral density $\boldsymbol{f}(\omega)$, $\omega \in [0, 2\pi)$. This estimate is asymptotically unbiased.

If $d = 2$, bivariate periodograms can be defined that are building blocks of the estimates for the spectral density matrix up to a factor 2π. If $d > 1$, we can plot the (real) eigenvalues of the spectral density matrix at the Fourier frequencies. In case of real time series, it suffices to stay on the $[0, \pi]$ interval.

The spectrum of the block Toeplitz matrix $\mathfrak{C}_T$ asymptotically comprises the union of the spectra of the matrices $\boldsymbol{f}(\omega_j)$s ($j = 0, 1, \ldots, T-1$) as $T \to \infty$. Theorem 5 is about this issue. By means of the complex eigenvectors we can make PCA and law rank approximation of the process itself. Actually, $\hat{\mathbf{Z}}_{\omega_j}$s give a discrete approximation of the orthogonal increment process, and the smoothed periodogram is proportional to the spectral density, that shows the spectra of the process. In 1D, the periodogram is proportional to the spectral density itself, the areas under which are the variances of the PC-transformed process. The PCs are continuum many, they extend to the whole $[0, 2\pi]$ interval, and the place of the peak of the periodogram shows the position of the PC with the highest variance. From this, we can draw conclusions for the periods of the underlying time series, see the Example in the next lesson.