CHAPTER 1

INTRODUCTION

Periodically correlated (PC) random processes are random processes in which there exists a periodic rhythm in the structure that is generally more complicated than periodicity in the mean function. We will begin with an illustration of some meteorological data.

The top trace of Figure 1.1 shows a 40 day record of hourly solar radiation levels taken at meteorological station DELTA on Ellsmere Island, N.W.T., Canada.

A daily (24 hour period) rhythm may be observed in this data in two ways: in the periodic average (or mean) and in the variation about the periodic mean. Since solar radiation can be expected to have a 24 hour period, let us compute the average of the 40 measurements for each of the 24 hours. Precisely, if the time series is denoted by $X_t, t = 1, 2, ..., NT$, where $NT = 960$, then the *sample periodic mean* (with period $T = 24$) is computed by

$$\hat{m}_N(t) = \frac{1}{N} \sum_{p=0}^{N-1} X_{t+pT}, \quad t = 1, 2, ..., T, \tag{1.1}$$

Periodically Correlated Random Sequences:Spectral Theory and Practice. By H.L. Hurd and A.G. Miamee

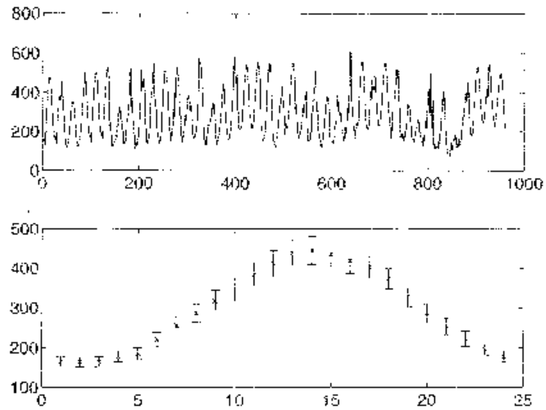

Figure 1.1 (Top) Solar radiation from station DELTA of the Taconite Inlet Project [211]. (Bottom) $\hat{m}_N(t)$ with 95% confidence intervals determined by the Student's t. $T = 24, N = 40$.

and plotted in the bottom trace of Figure 1.1. For t not in the base interval, $\hat{m}_N(t)$ is defined periodically. It is visually clear that the sample periodic mean is not constant (but properly periodic) and a simple hypothesis test for difference in mean, say, between hour 1 and hour 13, indicates a difference with much significance.

We postpone the details of testing for a proper fluctuation in the mean (i.e., for rejection of the hypothesis that the true mean $m(t)$ is constant) to Chapter 9.

The top trace of Figure 1.2 is the deviation $Y_t = X_t - \hat{m}_N(t)$ of X_t from the sample periodic mean $\hat{m}_N(t)$. The bottom trace presents the sample periodic variance,

$$S_N^2(t) = \frac{1}{N-1} \sum_{p=0}^{N-1} Y_{t+pT}^2, \quad t = 1, 2, ..., T, \tag{1.2}$$

and it too appears to have a significant (with the details again postponed) variation through the period. So it is not just the mean that appears to have a periodic rhythm, the variance does too, suggesting that the entire probability law may have a periodic rhythm. We will state this more precisely following some discussion of notation.

First, a stochastic (or random) process $X(t, \omega)$ is taken to be a function $X : \mathbb{I} \times \Omega \to \mathbb{C}$, where $\mathbb{C}$ is the set of complex numbers, $\mathbb{I}$ is called the *index set*, and Ω is a space, on which a sigma-algebra $\mathcal{F}$ of subsets and a probability measure P are defined. An $\mathcal{F}$-measurable function is called a *random variable*, and for a stochastic process, the function $X(t, \cdot)$ is assumed to be a random variable for each $t \in \mathbb{I}$. Although the focus of this book is random sequences

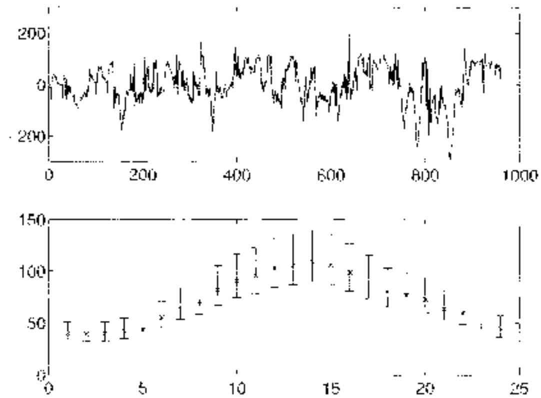

Figure 1.2 (Top) Deviation around sample periodic mean. (Bottom) $S_P(t)$ with 95% confidence limits determined by the chi-squared distribution with $N - 1 = 39$ degrees of freedom.

($\mathbb{I} = \mathbb{Z}$) having a periodic rhythm, extensions of the ideas to fields ($\mathbb{I} = \mathbb{Z}^2$), to processes ($\mathbb{I} = \mathbb{R}$), to multivariate sequences, and to *almost periodic* sequences are briefly described in the supplements to this chapter.

We will most often denote the element of the random sequence by X_t so that the dependence on ω is suppressed and the index is the subscript symbol t, conveying time. The essential structure needed to characterize a stochastic process is its probability law, meaning the collection of finite dimensional distributions, defined as the probabilities

$$P_{t_1,t_2,\dots,t_n}(A_1, A_2, \dots, A_n) = P[X_{t_1} \in A_1, X_{t_2} \in A_2, \dots, X_{t_n} \in A_n] \qquad (1.3)$$

for arbitrary n, collection of times $t_1, t_2, \dots, t_n$ in $\mathbb{Z}$, and Borel sets $A_1, A_2, \dots, A_n$ of $\mathbb{C}$.

Definition 1.1 (Strict Stationarity) *A stochastic process $X_t(\omega)$ is called* (strictly) stationary *if its probability law is invariant with respect to time shifts, or more precisely, if for arbitrary n, collection of times $t_1, t_2, \dots, t_n$ in $\mathbb{Z}$, and Borel sets $A_1, A_2, \dots, A_n$ of $\mathbb{C}$ we have*

$$P_{t_1+1,t_2+1,\dots,t_n+1}(A_1, A_2, \dots, A_n) = P_{t_1,t_2,\dots,t_n}(A_1, A_2, \dots, A_n). \qquad (1.4)$$

Now we can formalize the structure suggested by Figures 1.1 and 1.2.

Definition 1.2 (Periodically Stationarity) *A stochastic sequence $X_t(\omega)$ is called* (strictly) periodically stationary *with period T if, for every n, any collection of times $t_1, t_2, \dots, t_n$ in $\mathbb{Z}$, and Borel sets $A_1, A_2, \dots, A_n$ of $\mathbb{C}$,*

$$P_{t_1+T,t_2+T,\dots,t_n+T}(A_1, A_2, \dots, A_n) = P_{t_1,t_2,\dots,t_n}(A_1, A_2, \dots, A_n), \qquad (1.5)$$

and there are no smaller values of $T > 0$ for which (1.5) holds.

Synonyms for *periodically stationary* include *periodically nonstationary*, *cyclostationary* (think of cyclically stationary), *processes with periodic structure*, and a few others. For a little more on this nomenclature, see the historical notes (Section 1.2) at the end of this chapter.

If (1.5) holds for $T = 1$, then the process (or sequence) is stationary and it is clear that if X_t is periodically stationary with period T, then it is also for period $kT, k \in \mathbb{Z}$. And so we say that a sequence is properly periodically stationary if the least T for which (1.5) holds exceeds 1. Most often we will be considering second order random sequences, so that

$$E\{|X_t|^2\} = \int_\Omega |X_t(\omega)|^2 P(d\omega) < \infty, \quad \text{for all } t \in \mathbb{Z}.$$

We will sometimes just write that $X_t \in L^2$. The mean exists for second order sequences

$$m(t) := \int_\Omega X_t(\omega) P(d\omega), \quad \text{for all } t \in \mathbb{Z}$$

and we define the *covariance* of the pair (X_s, X_t) to be

$$R(s,t) := \text{Cov}\,(X_s, X_t) = E\{[X_s - m_s]\overline{[X_t - m_t]}\}.$$

If there is no ambiguity, we will write $m(t)$ and $R(s,t)$ for the mean and covariance of X_t. Sometimes, in order to conserve space, we will write variables as subscripts rather than in parentheses, such as m_t for $m(t)$ and $R_{s,t}$ for $R(s,t)$.

Since, for a zero mean sequence X_t, the covariance

$$\text{Cov}\,(X_s, X_t) = E\{X_s \overline{X_t}\}$$

is clearly the L^2 inner product, our conclusions about zero mean second order random sequences can be interpreted for sequences of vectors in a Hilbert space. For some topics (e.g., those involving shift operators) it will be more natural to think of X_t in this manner.

The notion of stationarity for second order sequences is expressed in terms of the first two moments.

Definition 1.3 (Weak Stationarity) *A second order random process $X_t \in L^2(\Omega, \mathcal{F}, P)$ with $t \in \mathbb{Z}$ is called (weakly) stationary if for every $s, t \in \mathbb{Z}$*

$$m(t) = m \quad \textit{and} \quad R(s,t) \equiv R(s-t).$$

If X_t is of second order, periodic stationarity induces a rhythmic structure in the mean and covariance.

Definition 1.4 (Periodically Correlated) *A second order process $X_t \in L^2(\Omega, \mathcal{F}, P)$ is called periodically correlated with period T (PC-T) if for every $s, t \in \mathbb{Z}$*

$$m(t) = m(t+T) \tag{1.6}$$

and

$$R(s,t) = R(s+T, t+T) \tag{1.7}$$

and there are no smaller values of $T > 0$ for which (1.6) and (1.7) hold.

It is clear that if the period is T , then (1.6) and (1.7) also hold when T is replaced by kT, for any integer k. If X_t is PC-1 then it is stationary (weakly) because then $R(s,t)$ is a function only of $s-t$. Clearly a stationary sequence is PC with every period.

We will write an indexed collection $\{X_t^j, j = 1, 2, \ldots, q\}$ of random sequences as the vector sequence $\mathbf{X}_t = [X_t^1, X_t^2, \ldots, X_t^q]'$.

Definition 1.5 (Multivariate Stationarity) *A second order q-variate random sequence $\mathbf{X}_t$ with $t \in \mathbb{Z}$ is called (weakly) stationary if*

$$E\{X_t^j\} \equiv m^j \tag{1.8}$$

and

$$R^{jk}(s,t) = \text{Cov}\,(X_s^j, X_t^k) = R^{jk}(s-t) \tag{1.9}$$

for all $s, t \in \mathbb{Z}$ and $j, k \in \{1, 2, \ldots, q\}$. If this is the case, we denote

$$\mathbf{m} = [m^1, m^2, \ldots, m^q]' \quad \textit{and} \quad \mathbf{R}(\tau) = [R^{jk}(\tau)]_{j,k=1}^q$$

.

Multivariate (or vector) sequences obtained from the *blocking* of univariate (or scalar) sequences will be indexed by n and thus denoted as $\mathbf{X}_n$. That is, the univariate sequence X_t is related by T-blocking to the T-variate sequence $\mathbf{X}_n$ by

$$[\mathbf{X}_n]^j = X_{j+nT}, \; n \in \mathbb{Z}, \; j = 0, 1, \ldots, T-1. \tag{1.10}$$

The following proposition is a simple matter of following the indices.

Proposition 1.1 (Gladyshev) *A second order random sequence $\{X_t : t \in \mathbb{Z}\}$ is PC with period T if and only if the T is the smallest integer for which the T-variate blocked sequence $\mathbf{X}_n$ (1.10) is stationary.*

Proof. Considering the covariance $\text{Cov}([\mathbf{X}_n]^j, [\mathbf{X}_m]^k) = \text{Cov}(X_{j+nT}, X_{k+mT})$, then stationarity of $\mathbf{X}_n$ implies

$$\text{Cov}\,([\mathbf{X}_n]^j, [\mathbf{X}_m]^k) = R^{jk}(n-m) = \text{Cov}\,(X_{j+nT}, X_{k+mT}),$$

which implies (1.7) holds for X_t, and conversely. The same argument applies to the mean. ∎

Periodically correlated sequences are generally *nonstationary* but yet they are nonstationary in a very simple way that, when the period T is known, makes them equivalent to vector valued stationary processes.

The term *periodically correlated* was introduced by E. G. Gladyshev [77], but the same property was introduced by W. R. Bennett [12] who called them *cyclostationary*.

Since PC sequences are so closely related to stationary vector sequences, which are rather well understood, then one can legitimately ask: why go to the effort to study the structure of these processes? There are several answers. First, the value of T, required to transform a PC sequence to a vector stationary sequence, sometimes is not known prior to the analysis of an observed time series. Thus studying the time and spectral structure of the process using its natural time organization can provide clues to help us develop tests for PC structure and estimators for the period T. Second, the issues concerning innovation rank are more easily understood for PC sequences than for multivariate sequences because the natural time order eliminates some ambiguity. Third, the methods developed here for sequences naturally carry over to continuous time and to the almost periodic case; and in those cases it is not generally possible to block the process into a stationary sequence of finite dimensional vectors.

We will often assume that $E\{X_t\} \equiv 0$ as it is the covariance (or quadratic) structure that is of most interest. However, we shall carefully discuss the issue of the additive periodic terms of a PC sequence, and how they can be conceptually viewed, and how they can be treated in the analysis of time series.

There are several ways in which two sequences can be considered equal. For example, two random processes X_t and Y_t can be called equal if for each $\omega \in \Omega$ their respective sample paths $X_t(\omega)$ and $Y_t(\omega)$ are the same. However, throughout this book, unless otherwise specified, we take two processes X_t and Y_t to be equal if

$$E \mid X_t - Y_t \mid^2 = 0, \quad \text{for every } t \in \mathbb{I}.$$

1.1 SUMMARY

This summary provides a little more detail about the contents with enough precision to make our direction clear, but not with the same care we will give subsequently. And it also provides further discussion of notation.

Chapter 1: Introduction. Gives an introductory empirical example to motivate the definitions, and then this summary followed by a historical development of the study of these processes. In this we do not attempt a complete bibliography but concentrate on the beginnings of the topic and give additional references that contain more complete bibliographies.

Chapter 2: Examples, Models, and Simulations. Presents simple models for constructing PC sequences, usually by combining randomness (usually through stationary sequences) with periodicity. Some important examples are sums and products of periodic sequences and stationary sequences, time scale modulation of stationary sequences, pulse amplitude modulation, periodic autoregressions, periodic moving averages, and periodically perturbed dynamical systems.

For most of these examples, results of simulations are presented to show the extent to which some sort of periodic rhythm is visually perceptible in the time series. These also illustrate that the usual periodogram typically does not reveal the presence of the periodic structure in PC sequences, and the periodogram of the squares sometimes can reveal the periodic structure, but not always.

Chapter 3: Review of Hilbert Spaces. Presents the basic facts about Hilbert space that will be needed. After definitions of vector space, inner product, and Hilbert space, general properties of (linear) operators are discussed. Of particular interest are projection operators, which have an important use in prediction, and unitary operators, which have a fundamental role in stationary and PC sequences. Finally, we review the spectral theory for unitary operators, including spectral measures, integrals, and the representation

$$U = \int_0^{2\pi} e^{i\lambda} E(d\lambda). \tag{1.11}$$

This spectral representation plays a critical role in the spectral theory for stationary and PC sequences.

Chapter 4: Stationary Random Sequences. Emphasizes the role of the unitary operator and its spectral representation as we believe this helps to give a clear view of PC sequences. The core result is that if X_t^j, $j = 1, 2, \ldots, q$ are *jointly* (weakly) stationary and $\mathcal{H} = \overline{sp}\{X_t^j : j = 1, 2, ..., q,\ t \in \mathbb{Z}\}$, the stationary covariance structure allows one to prove quite easily that there exists a *unitary* operator $U : \mathcal{H} \to \mathcal{H}$ for which

$$X_{t+1}^j = UX_t^j \tag{1.12}$$

for every $j = 1, 2, \ldots, q$ and $t \in \mathbb{Z}$. Iterating (1.12) gives $X_t^j = U^t X_0^j$ for all t, and by applying the spectral representation (1.11) we obtain the spectral representation of the sequence

$$X_t^j = \int_0^{2\pi} e^{i\lambda t} \xi^j(d\lambda) \tag{1.13}$$

where ξ_j is orthogonally scattered.

We then discuss the main topics connected with prediction, regularity and singularity, the Wold decomposition, innovations, the predictor expressed by innovations, the connection between spectral theory and prediction, and finally, finite past prediction. We also discuss the issue of rank in connection with innovations and spectral theory.

Chapter 5: Harmonizable Sequences. Presents the main facts about *harmonizable* random sequences with emphasis on what is important to PC sequences. As a generalization of the spectral representation for stationary sequences (and also for continuous time), M. Loève [138], who also wrote about (strongly) harmonizable processes in the first edition of Probability Theory [139], defined a sequence to be *harmonizable* if it has a spectral representation

$$X_t = \int_0^{2\pi} e^{i\lambda t} \xi(d\lambda), \tag{1.14}$$

where $\xi(\cdot)$ is an $L^2(\Omega, \mathcal{F}, P)$ valued measure but no longer has orthogonally scattered (or uncorrelated) increments as it does in the stationary case. In order to convey the precise meaning of (1.14), we discuss vector valued measures and integration with respect to such measures. Then we discuss *weakly* and *strongly* harmonizable sequences, their connection to projections of stationary sequences, and spectral representation

$$R(s, t) = \int_0^{2\pi} \int_0^{2\pi} e^{i\lambda_1 s - i\lambda_2 t} F(d\lambda_1, d\lambda_2) \tag{1.15}$$

of the covariance, where the sense of integration depends on whether X_t is weakly or strongly harmonizable.

Finally, we show how time invariant linear filtering affects the spectral representation of a harmonizable sequence (and of its covariance).

Chapter 6: Fourier Theory of the Covariance. This is a topic introduced and mainly completed by Gladyshev [77]. The bijection between PC-T sequences and T-variate stationary vector sequences makes it no surprise that the Fourier theory for the covariance for PC sequences is very much related to the Fourier theory for the covariance of stationary vector sequences.

The PC structure in the covariance (1.7) implies easily that

$$R(t + \tau, t) = \sum_{k=0}^{T-1} B_k(\tau) e^{i2\pi kt/T}, \tag{1.16}$$

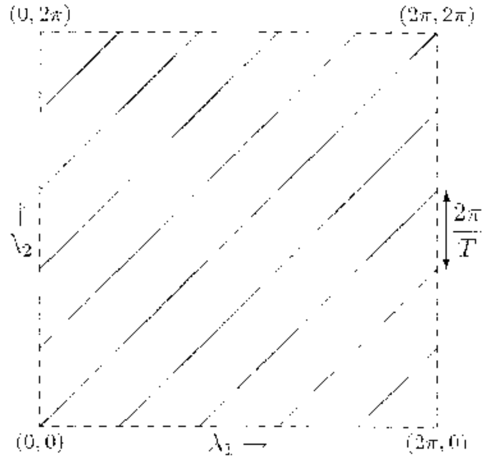

Figure 1.3 Support S_T of the spectral measure F for a periodically correlated sequence.

where $B_k(\tau) = T^{-1}\sum_{t=0}^{T-1} e^{-i2\pi kt/T} R(t+\tau, t)$. Using the connection to stationary vector sequences, Gladyshev argued that the coefficient functions $\{B_k(\tau) : k = 0, 1, \dots, T-1\}$ are Fourier transforms

$$B_k(\tau) = \int_0^{2\pi} e^{i\lambda\tau} dF_k(\lambda). \tag{1.17}$$

We show it by use of a characterization of Fourier transforms based on a theorem of Riesz.

The plausibility that $R(s,t)$ given by (1.16) can be put into the form (1.15), which would make the covariance strongly harmonizable, turns out to be a fact, so every PC sequence is strongly harmonizable. The defining rhythm (1.7) associated with a PC sequence constrains the support set of the spectral measure F appearing in (1.15) to the $2T-1$ diagonal lines

$$S_T = \{(\lambda_1, \lambda_2) \in [0, 2\pi)^2 : \lambda_2 = \lambda_1 - 2\pi k/T, k = -(T-1), \dots, T-1\}, \tag{1.18}$$

as illustrated in Figure 1.3. The support lines of F may be identified with the sequence $\{F_k(\cdot) : k = 0, \dots, T-1\}$ of complex measures whose Fourier transforms are $B_k(\tau)$.

We discuss the Lebesgue decomposition of F and the issue of point masses in the random spectral measure $\xi(\cdot)$, some of which are produced by the mean $m(t)$. The effects of time invariant and periodic filtering, sampling, and random time shifting of PC sequences are examined. We also give the mapping between the spectral measure F and the matrix valued spectral measure $\mathbf{F}$ of the (blocked) vector stationary sequence $\mathbf{X}_n$.

Chapter 7: Representations of PC Sequences. Addresses various representations of PC sequences, with an emphasis on the connection to the unitary operator of a PC sequence. The basic covariance structure (1.7) implies that on the Hilbert space $\mathcal{H} = \overline{sp}\{X_t : t \in \mathbb{Z}\}$ there exists a unitary operator $U : \mathcal{H} \mapsto \mathcal{H}$ for which

$$X_{t+T} = UX_t \tag{1.19}$$

for every $t \in \mathbb{Z}$. Thus U is a shift operator for X *but only for shifts of length* T. The most basic consequence of (1.19) is that we can find (derived from U) another unitary operator $V : \mathcal{H} \mapsto \mathcal{H}$ and a periodic function P_t taking values in $\mathcal{H}$ for which

$$X_t = V^t P_t, \text{ for every } t \in \mathbb{Z}. \tag{1.20}$$

Using the spectral theorem for the unitary V leads to a spectral representation $X_t = \int_0^{2\pi} e^{i\lambda t}\xi_1(d\lambda, t)$, where $\xi_1(\cdot, t)$ is orthogonally scattered for all t whereas harmonizability implies that PC sequences *also* have a spectral decomposition (1.14) with respect to a *time invariant* random spectral measure ξ that is not orthogonally scattered. With the aid of (1.20) we explicitly construct the time independent random measure ξ. By expanding P_t in a Fourier series we obtain the Gladyshev representation $X_t = \sum_{k=0}^{T-1} Z_t^k e^{i2\pi kt/T}$ as a Fourier series having jointly stationary coefficients $\{Z_t^k : k = 0, 1, \ldots, T-1\}$. We show (see [160]) how to explicitly construct a dilated sequence Y_t such that X_t can be recovered by projection, $X_t = PY_t$.

Chapter 8: Prediction of PC Sequences. Treats the prediction problem for PC sequences, again with the help of the unitary operator U. We discuss regularity and singularity, the Wold decomposition and innovations, where we find that, at any t, the dimension d_t of the innovation space is either 0 or 1, and $d_t = d_{t+T}$. It follows that a regular PC-T sequence has an infinite moving average representation with respect to the orthonormal sequence $\{\xi_t : t \in D^+\}$,

$$X_t = \sum_{j \geq 0: t-j \in D^+} a_t^j \xi_{t-j},$$

where the ℓ^2 sequence of coefficients $A_t = \{a_t^j : j \geq 0\}$ is periodic $A_t = A_{t+T}$ and $D^+ = \{t : d_t > 0\}$ is the set of times where nontrivial innovation occurs. The number $r = \sum_{j=1}^{T} d_{t+j}$ is a constant (independent of t) and is defined to be the *rank* of a PC-T sequence. A PC-T sequence is of *full rank* whenever $r = T$. For a simply constructed PC sequence of less than full rank, let $\{\xi_t, t \in \mathbb{Z}\}$ be an orthonormal sequence; the sequence $\{\ldots, \xi_{-1}, \xi_{-1}, \xi_0, \xi_0, \xi_1, \xi_1, \ldots\}$ is PC-2 but of rank 1. We discuss the prediction problem for infinite and finite sets of predictors and give some illustrative results for periodic autoregressions of order 1, which, although simple, may also be of less than full rank. We then discuss prediction based on a finite past, periodic partial autocorrelations, and the periodic Durbin Levinson algorithm. We also give the innovation algorithm for nonnegative definite (and hence possibly of deficient rank) covariances, along with a Cholesky decomposition for NND matrices.

Chapter 9: Estimation of Mean and Covariance. Addresses the problems of estimation of the time-varying mean $m_t = E\{X_t\} = m_{t+T}$ and covariance $R_{t+\tau,t} = E\{[X_{t+\tau} - m_{t+\tau}][X_t - m_t]\} = R_{t+T+\tau,t+T}$, and their Fourier coefficients $\widetilde{m}_k = T^{-1}\sum_{t=0}^{T-1} m_t e^{-i2\pi kt/T}$ and $B_k(\tau) = T^{-1}\sum_{t=0}^{T-1} R_{t+\tau,t}e^{-i2\pi kt/T}$. Here, X_t is taken to be a real valued PC-T sequence. The corresponding estimators, which may be motivated by the lifting to the stationary vector sequence $\mathbf{X}_n$, are given by the following: for $\widehat{m}_{t,N}$ see (1.1),

$$\widehat{\widetilde{m}}_{k,N} = \frac{1}{NT}\sum_{j=0}^{NT-1} X_j e^{-i2\pi kj/T} = \frac{1}{T}\sum_{t=0}^{T-1} \widehat{m}_{t,N} e^{-i2\pi kt/T},$$

$$\widehat{R}_N(t+\tau,t) = \frac{1}{N}\sum_{k=0}^{N-1}[X_{t+kT+\tau} - \widehat{m}_{t+\tau,N}][X_{t+kT} - \widehat{m}_{t,N}],$$

$$\widehat{B}_{k,NT}(\tau) = \frac{1}{NT}\sum_{t\in I_{NT,\tau}} [X_{t+\tau} - \widehat{m}_{t+\tau,N}][X_t - \widehat{m}_{t,N}]e^{-i2\pi kt/T}.$$

For $\widehat{m}_{t,N}$ and $\widehat{\widetilde{m}}_{k,N}$, we give conditions for mean square consistency in spectral terms, express the limits spectrally, and discuss the connection to the mean ergodic theorem. We show how to use the random time shift to give almost sure consistency in terms of B_0 and F_0 from known stationary results. The random time shift is used again to obtain asymptotic normality via mixing for linear PC sequences. The practical estimation programs `permest.m` and `permcoeff.m` are presented and demonstrated. To test for a proper periodic mean (null is $m(t) = m$), the former produces confidence intervals based on Student's t and an ANOVA test; the latter uses a variance contrast method applied to the periodogram to produce p-values for $\widehat{m}_k = 0$.

For $\widehat{R}_N(t+\tau,t)$ we give conditions for consistency in probability for a linear PC sequence using the lifted correlations and the fact that a linear PC sequence, when lifted, is a linear T-variate stationary sequence so known results may be applied. For X_t with bounded fourth moments, various conditions on the second moments of $Z^\dagger_{t,\tau} = [X_{t+\tau} - m_{t+\tau}][X_t - m_t] - R(t+\tau,t)$ ensure mean square consistency; and other conditions give almost sure consistency. Asymptotic normality is obtained using either of two approaches: (1) a condition on the covariance of $Z^\dagger_{t+jT,\tau}$ along with ϕ-mixing, and (2) normality of X_t and a summability condition on the covariance, namely, $\sum_{t=0}^{T-1}\sum_{\tau=-\infty}^{\infty}|R(t+\tau,t)|^2 < \infty$. Similar results are obtained for consistency of $\widehat{B}_{k,NT}(\tau)$.

The practical estimation programs persigest.m and Bcoeff.m are presented and demonstrated. Program peracf.m computes $\widehat{R}_{N_{t,\tau}}(t+\tau,t)$ and $\widehat{\rho}_{N_{t,\tau}}(t+\tau,t)$. Assuming normal X_t, confidence limits for the latter are computed (and plotted) by use of the Fisher transformation. Also computed are tests for (1) equality of correlations $\rho(t+\tau,t) = \rho(\tau)$, where $\rho(\tau)$ is some unknown constant; and (2) for $\rho(t+\tau,t) \equiv 0$ for some specific τ ($\equiv$ means for all t in a period). Program Bcoeff.m computes (and plots) $\widehat{B}_{k,NT}(\tau)$ for $k = 0,1,\ldots,\lfloor (T-1)/2 \rfloor$ (real X_t) via the sample Fourier transform applied to $Y_{t,\tau} = [X_{t+\tau} - \widehat{m}_{t+\tau,N}][X_t - \widehat{m}_{t,N}]$. This permits the computing of p-values for the test $\widehat{B}_{k,NT}(\tau) = 0$, based on the variance contrast method of Section 9.2.2. Also, program persigest.m computes $\widehat{\sigma}_N(t)$ along with confidence intervals based on normal (χ^2 distribution with $N-1$ degrees of freedom) and the Bartlett test for heterogeneous variances.

Chapter 10: Spectral Estimation. Addresses the problems of estimation of the possibly complex density functions $f_k(\lambda)$ when the $F_k(\cdot)$ in (1.17) are absolutely continuous with respect to Lebesgue measure. The principal idea for the estimation of $f_k(\lambda)$ is based on smoothing the two-dimensional periodogram

$$I(N,\lambda_1,\lambda_2) = \frac{1}{2\pi N}\tilde{X}_N(\lambda_1)\overline{\tilde{X}_N(\lambda_2)} \tag{1.21}$$

along lines of support of F in $[0,2\pi)\times[0,2\pi)$, where

$$\tilde{X}_N(\lambda) = \sum_{t=0}^{N-1}[X_t - m_t]e^{-i\lambda t}$$

is the sample Fourier transform of $[X_t - m_t], t = 0,1,\ldots,N-1$. Note the usual estimators for the spectral density in the stationary case are formed by smoothing $I(N,\lambda_1,\lambda_2)$ along the main diagonal $\lambda_1 = \lambda_2$. We begin by showing that $\tilde{f}_{k,N}(\lambda) = I_N(\lambda,\lambda-2\pi k/T)$ is the Fourier transform of $\widehat{B}_{k,N}(\tau)$ and if $\sum_{\tau=-\infty}^{\infty}\sum_{t=0}^{T-1}|R(t+\tau,t)| < \infty$, then $\tilde{f}_{k,N}(\lambda)$ is an asymptotically unbiased estimator for $f_k(\lambda)$. By assuming X_t is Gaussian, we obtain $\lim_{N\to\infty}\mathrm{Var}\,[\tilde{f}_{k,N}(\lambda)] = f_0(\lambda)f_0(\lambda-2\pi k/T)$ if $\lambda \neq \pi n/T$ and the limit is $f_0(\lambda)f_0(\lambda-2\pi k/T) + |f_{n-k}(\pi n/T)|^2$ if $\lambda = \pi n/T$, thus showing, as in the stationary case, that the estimator is not consistent. However, as in the stationary case, consistency can be achieved by smoothing $\tilde{f}_{k,N}(\lambda)$ by the Fourier transform $W(\lambda)$ of a summable weight sequence $w(j)$, $\widehat{f}_{k,N}(\lambda) = \frac{1}{\mu_N}\int_0^{2\pi}W((\sigma-\lambda)/\mu_N)\tilde{f}_{k,N}(\sigma)d\sigma$, and where μ_N is a positive sequence with $\mu_N \to 0$ and $N\mu_N \to \infty$ as $N\to\infty$. We give conditions under which estimators formed in this manner are consistent and asymptotically normal. If X_t is a Gaussian PC-T sequence for which $\sum_{\tau=-\infty}^{\infty}\left[\sum_{t=0}^{T-1}|R(t+\tau,t)|^2\right]^{1/2} < \infty$, then there exists a $K > 0$ for which $N\mu_N \text{Cov}[\hat{g}_j(\lambda_1), \hat{g}_k(\lambda_2)] \leq K$ for any $j, k \in [0, 1, ..., T-1]$ and $\lambda_1, \lambda_2 \in [0, 2\pi)$. If X_t is periodically stationary with fourth moments and uniformly ϕ-mixing with $\sum_{n=-\infty}^{\infty} (\phi_n)^{1/2} < \infty$ and $k(j)$ is any sequence with $\sum_{j=-\infty}^{\infty} k(j)|j|^{1/2} < \infty$, then

$$\lim_{N\to\infty} \text{Cov}[\hat{f}_j(\lambda_1), \hat{f}_k(\lambda_2)] = 0$$

if $\mu_N \to 0$, $N\mu_N^3 \to \infty$ as $N \to \infty$. In addition, we discuss the empirical spectral analysis of harmonizable sequences, which leads naturally to the notion of spectral coherence, which may be defined theoretically as

$$\begin{aligned}\gamma(\lambda_1, \lambda_2) &= \lim_{N\to\infty} \frac{\text{Cov}[\tilde{X}_N(\lambda_1), \tilde{X}_N(\lambda_2)]}{\text{Var}^{1/2}[\tilde{X}_N(\lambda_1)]\text{Var}^{1/2}[\tilde{X}_N(\lambda_2)]} \\ &= \lim_{N\to\infty} \text{Corr}[\tilde{X}_N(\lambda_1), \tilde{X}_N(\lambda_2)],\end{aligned}$$

and whose squared magnitude may be estimated by

$$|\gamma(\lambda_p, \lambda_q, M)|^2 = \frac{|\sum_{m=1}^{M} \tilde{X}_{p-M/2+m} \overline{\tilde{X}}_{q-M/2+m}|^2}{\sum_{m=1}^{M} |\tilde{X}_{p-M/2+m}|^2 \sum_{m=1}^{M} |\tilde{X}_{q-M/2+m}|^2}$$

where $\tilde{X}_p$ is the sample Fourier transform of X_t of length N. For PC-T sequences, setting $\lambda_1 = \lambda$ and $\lambda_2 = \lambda_1 - 2\pi k/T$, we obtain

$$\gamma(\lambda, \lambda - 2\pi k/T) = \frac{f_k(\lambda)}{f_0^{1/2}(\lambda) f_0^{1/2}(\lambda - 2\pi k_0/T)}.$$

Spectral coherence is useful because (1) it may be estimated in a simple manner, (2) its distribution is known under a null hypothesis condition, and (3) in the case of PC sequences, it gives a way to judge the largeness of $\hat{f}_{k,N}(\lambda)$. If a harmonizable sequence has a jump (or atom) in its random spectral measure at λ_a and at λ_b, then the theoretical spectral coherence at (λ_a, λ_b) will be unity. Hence it becomes important to sense the presence of discrete spectral components in a time series and to remove them. Such methods are also discussed in this chapter. Finally, we present programs `fkest.m` and `scoh.m` that implement the estimator $\hat{f}_{k,N}(\lambda)$ and the empirical spectral coherence $|\gamma(\lambda_p, \lambda_q, M)|^2$ given above.

Chapter 11: A Paradigm for Nonparametric Analysis of PC Time Series. Suppose one is given a sample of a time series and asked the question: Does this series exhibit the PC property or not? If so, what can we say about it? This chapter summarizes and organizes the methods discussed in previous chapters into a procedural outline, or paradigm, for answering these questions only within the scope of nonparametric time series analysis. That is, we consider only the tools of mean, correlation, and spectral measurements. Obviously, our ability to answer these questions, especially regarding characterization, will substantially improve by the inclusion of PARMA time series analysis, a topic to be addressed in future writings.

1.2 HISTORICAL NOTES

The notion of PC processes seems to have begun with W. R. Bennett [12] who observed their presence in a communication theoretic context and called them *cyclostationary*. L. I. Gudzenko [84] initiated the subject of nonparametric spectral analysis for PC processes. V. A. Markelov [149] addressed some level crossing problems for Gaussian PC processes. A short time later E. G. Gladyshev [77] published the first analysis of spectral properties and representations based on the connection between PC sequences and stationary vector sequences. He gave necessary and sufficient conditions, in the spirit of A. Khintchine [129], for a doubly indexed sequence $R(s,t)$ to be the correlation of a PC sequence and argued that all PC sequences are strongly harmonizable and showed that their spectral support consists of a family of lines parallel to the main diagonal and having spacing of $2\pi/T$. He also gave two representations for the processes and conditions for the processes to be purely nondeterministic. In 1963 Gladyshev [78] treated continuous time PC processes and introduced the almost periodically correlated processes.

In a series of papers L. J. Herbst [90–96] explored sequences and processes whose variances may be periodic or almost periodic with respect to time; this work was done without the benefit of the PC structure. W. M. Brelsford [25] obtained, for PC sequences, a spectral-like representation of mixed summation and integral form. He also presented methods for estimation of the periodic coefficients in periodic autoregression models.

In various investigations of asymptotic stationarity, J. Kampé de Fériet [126], J. Kampé de Fériet and F. N. Frenkiel [127], and Parzen [178, 180] mentioned processes that are PC in nature but their work concentrated on the estimation of the asymptotic correlation and spectral density functions (i.e., on estimation of $B_0(\tau)$ and $f_0(\lambda)$; see the preceding summary of Chapter 6 for this notation).

To give some of the early connections to applications, Markelov [149] states that the noise output of a parametric amplifier has the PC property if the noise input is stationary; along similar lines, Parzen [179] suggested that a Poisson process with time periodic parameter would be a model for electron emissions from the cathode of a temperature limited diode whose filament was heated by an alternating current. A. S. Monin [165] suggests using PC processes as models for meteorological time series and R. H. Jones and Brelsford [122] do this for a time series of temperatures.

Several additional books touch on various aspects of PC processes. First, A. Papoulis [176] discusses various properties of PC processes in an early edition of his book on probability and stochastic processes; he calls them periodically stationary. L. E. Franks [58] discusses cyclostationary processes in a book on communication theory.

For the continuous time case, H. L. Hurd [101] showed the nature of the spectral support (an extension of Gladyshev's theorem) for strongly harmonizable PC processes, identified the connection between PC processes and those that can be made stationary by an independent uniformly distributed time shift, and obtained consistency results for estimation of the coefficient functions $B_k(\tau)$ and the densities $f_k(\lambda)$. H. Ogura [174] presented some of the spectral theory based on harmonizable processes. W. A. Gardner, in his dissertation [63] developed various representations of continuous time PC processes and used them in the solution of estimation problems. Much of this appears in the paper by Gardner and Franks [64]. After the initial work by Gladyshev, the topic was seriously examined in the former Soviet Union by Y. A. Dragan and his colleagues; much of their work seems to be summarized in three books [50–52], all in Russian. A. M. Yaglom [227] gives many references and a nice exposition of many of the basic relationships; we recommend this as a reference for those who wish to work in the topic.

Much work on cyclostationary processes followed, mainly motivated by communications problems and lead by Gardner, resulting in two books [66,67]. In these books Gardner principally takes a viewpoint much like Wiener's generalized harmonic analysis, where an observed sequence is considered to be a nonrandom sequence. The usual notion of probability is replaced with a limit of occupation time above a threshold, or fraction of time [67,166]. The approach produced understanding and solutions to many problems and is therefore interesting and useful. A discussion of the two views (random and nonrandom) may be found in [69]. Some later efforts to clarify the Wold isomorphism [67,226] applied to random and nonrandom cyclostationary sequences are given in [116,117].

In 1992 Gardner initiated a large meeting, whose subject was *Cyclostationarity in Communications and Signal Processing* [73], which brought together engineers, statisticians, and mathematicians. Many new problems and collaborations came from this meeting.

Subsequent work took several directions, in addition to continued work on communications problems. In statistics, work ensued on structural theory of PC and almost PC processes [23,44,46,65,97,98,102–104,106,110–113,116–118,143,145,157,160] on spectral and covariance estimation [5–7,30,39–41,43,75,105,108,109,135,212,228], on testing, [7,37,72,107] and on PARMA time series [10,137,140,175,205,212,213,217–219].

Interesting work on PC processes, a little under-represented here, exists in some other fields of study. As pointed out earlier by Monin [165] and Jones and Brelsford [122], there is a natural application in meteorology due to the obvious daily or annual forcing. Applications of both parametric and nonparametric methods are indicated and have been used rather extensively. We find the connection between PC processes and Bloch's theorem, pointed out by K. Kim, G. North, and J. Huang [130, 131] to be of great interest. An extensive body of related work on periodic control [15–17] has a direct relation to PC sequences. For connected work in economics, see the book by P. H. Franses [60] and the references therein.

A recent survey on cyclostationarity by Gardner, A. Napolitano, and L. Paura [74] contains a very complete bibliography.

During this period of development of PC processes, the theory of harmonizable processes also matured. H. Cramér [36] attributes the word *harmonizable* to M. Loève [138], who also wrote about (strongly) harmonizable processes in [139]. Yu A. Rozanov [200] made an important early contribution, before the more recent developments [1, 31, 38, 82, 99, 100, 144, 155, 169, 189].

PROBLEMS AND SUPPLEMENTS

1.1 PC fields indexed on $\mathbb{Z}^2$. A collection of second order random variables $X_{s,t}$ indexed on $\mathbb{Z}^2$ is called a (strongly) PC field with period (S, T) if its mean and covariance functions satisfy

$$m(s,t) = m(s + kS, t + lT), \tag{1.22}$$

$$R(s,t,s',t') = R(s + kS, t + lT, s' + kS, t' + lT) \tag{1.23}$$

for all integers s, t, s', t' and k, l in $\mathbb{Z}$.

The second order random field $X_{s,t}$ is called *weakly PC* with *period* (S, T) if

$$R(s,t,s',t') = R(s + S, t + T, s' + S, t' + T) \tag{1.24}$$

for every s, t, s', t'. Here we require $S \geq 0$ and $T \geq 0$ but do not permit $S = T = 0$ because this would put no constraint on the covariance structure of the field.

A weakly PC random field is essentially a countable collection of PC sequences arranged along parallel lines of slope T/S in $\mathbb{Z}^2$. If X is strongly PC, then it is also weakly PC.

1.2 Multivariate PC sequences. The multivariate sequence $\mathbf{X}_t = [X_t^1, X_t^2, \ldots, X_t^N]'$ is PC with period T if

$$m^j(t) = E\{X_t^j\} = m^j(t + T) \tag{1.25}$$

and

$$R^{jk}(s,t) = E\{[X_s^j - m^j(s)][\overline{X_t^k - m^k(t)}]\} = R^{jk}(s - T, t - T) \tag{1.26}$$

for every $j,k = 1,2,...,N$ and $s,t \in \mathbb{Z}$.

1.3 Almost PC sequences. A complex valued nonrandom sequence f_t is called *almost periodic* in the Bohr sense if for every $\varepsilon > 0$ the set

$$E(\varepsilon) = \{u : \sup_t |f_{t+u} - f_t| < \varepsilon\} \tag{1.27}$$

has bounded gaps, meaning there is a real number A for which every interval of length A intersects $E(\varepsilon)$. A L^2 sequence is *almost PC* if for every τ the sequence $R(t+\tau, t)$ is Bohr AP with respect to t.

1.4 Continuous time processes and fields. For $\mathbb{I} = \mathbb{R}$, the defining equations (1.6) and (1.7) as well as (1.22) and (1.23) read exactly the same. Continuous time APC processes are L^2 processes for which $R(t+\tau, t)$ is Bohr AP with respect to t for each τ. A function $f : \mathbb{R} \mapsto \mathbb{C}$ is Bohr almost periodic if it is continuous and for every $\varepsilon > 0$ the set $E(\varepsilon)$ defined by (1.27) has bounded gaps.