T-1651

Discrete Time- and Band-Limited Functions and Toeplitz Matrices

By

Kempton L. Huehn

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

T-1651

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirement for the degree of Doctor of Philosophy (Mathematics).

Signed: Kempton L. Hicks
Student

Golden, Colorado
Date: October 18, 1974

Approved: Raymond K Mueller
Thesis Advisor

Robert A. Wolf
Head of Department

Golden, Colorado
Date: October 18, 1974

ABSTRACT

A finite set of observations of some Geophysical phenomenon is treated as a single observation (realization) of a finite stochastic process. Each observation lies in an N-dimensional complex vector space. Discrete analogues of time- and band-limited functions, as observations, are defined as well as the discrete Fourier transform.

It is shown that the discrete Fourier transform of a discrete, wide-sense stationary stochastic process with zero mean and band-limited sample paths is an uncorrelated process with zero mean and conversely.

A finite Karhunen-Loeve expansion is defined for a finite stochastic process which is determined by the covariance matrix of the process.

The covariance matrix of a finite wide-sense stationary process is Toeplitz. New matrices, called generalized Toeplitz matrices, are defined and it is shown that an invertible, generalized Toeplitz matrix has a generalized Toeplitz inverse. This property of generalized Toeplitz matrices, along with persymmetric and centrosymmetric matrices, are used to clarify the Trench-Zohar algorithm for inverting Toeplitz matrices. An algorithm is developed which solves a Toeplitz system of algebraic equations and which generalizes existing algorithms.

A search for explicit expressions for the eigenvalues and eigenvectors of Toeplitz matrices is conducted and relationships between a Toeplitz matrix of order n and certain circulant matrices of order 2n and 2n-1 are established. Circulant matrices are special Toeplitz matrices whose eigenvalues and eigenvectors are known. The eigenvalues and eigenvectors are needed in the finite Karhunen-Loeve expansion.

Table of Contents

Acknowledgment

I wish to acknowledge
Dr. R. K. Mueller for his helpful advice in formulating this thesis,
the late Dr. Joeseph R. Lee for his help in providing a teaching assistantship,
the members of the Mathematics and Geophysics Departments at the Colorado School of Mines,
Peggy Randrup for typing this thesis, and
my wife and family.

Chapter 0. Introduction

The revolution in numerical analysis and problem solving resulting from the increased use of modern electronic computers is well known, especially to the analysts working in the sciences, business and economics. This revolution has had an especially profound effect upon those sciences, such as Geophysics, which involve the collection and analysis of large amounts of data.

Because of its speed and capacity, electronic computers can now efficiently solve problems involving large amounts of data and (or) large number of operations which previously were considered unsolvable or too time consuming. Not only can a computer solve many problems more efficiently but it can solve older problems by newer methods and can even suggest new problems.

However, the characteristics of electronic computers require changes in the methods used to solve problems. For example, the computer works with a finite binary representation of a number and this restriction requires the truncation (called "rounding") of input data and internal calculations. The errors which result from this truncating process are called round-off errors. Also, the computer can only perform a finite, albeit large, number of rather simple operations which requires the reduction of complex problems into simpler problems. Often a complicated problem cannot be analyzed in its original form and therefore must be replaced by a simpler problem which is solvable and which approximates the original problem. This simpler problem is often called a model of the original problem. Frequently the errors which result

from reducing a complicated problem, or replacing one problem by another, are not well understood or well known.

In spite of these difficulties, there is a great demand for computer oriented numerical techniques (algorithms, languages, etc.) which can be applied directly toward the solution of specialized problems. Because of the restrictions noted above, theoretical results cannot always be programmed directly on a computer and consequently, entirely distinct numerical algorithms must be developed. That is, numerical techniques must be developed which agree with the theoretical and which take advantage of the speed and capacity of the computer. Often these considerations require the numerical techniques to be modified in ways which are not obvious.

This thesis derives and develops numerical analogues to the time- and band-limited functions described in Slepian and Pollak. These numerical functions are called discrete time- and band-limited functions.

An application of these functions can be seen by considering three problems from Geophysics: polar wandering, earth tide, and free oscillations of the earth. Barber has a very good discussion of these problems and the application of Fourier analysis to them.

Polar wandering is the apparent motion of the north rotational pole relative to some fixed point on the celestial sphere. At the present time Geophysicists are engaged in discovering the causes for this motion. By analyzing the observations of this phenomenon in order to learn the properties which characterize the motion, an understanding of these properties could then be used to predict further motion or supply criteria

by which alternative theories (models) could be evaluated. Polar wandering has been observed for less than 100 years and only at discrete times (a discrete time-limited function). The observational data indicates that the apparent motion of the pole has two recurrent or periodic components and possibly a secular trend. The periodic components have periods of 12 and 14 months. By using the discrete Fourier transform we can estimate the presence, or absence, of certain periodic components in that motion. These periodic components are band-limited functions. Polar wandering is an example of the use of complex-valued (i.e. two dimensional) time series and is related to the variation of latitude problem (Smylie, Clarke and Ulrych).

The earth tide is the reaction of water and rock to the gravitational attraction of the moon and sun (and possibly other celestial bodies). As the apparent motion of the moon and sun are "almost" periodic we would expect that tidal observations would also exhibit this feature, which it does. However tidal observations also have aperiodic components superimposed on the periodic ones. Depending upon the interest or point of view of the observer, these undesirable components would be considered "noise." Most observations have certain undesirable components which often exhibit a randomness due to unknown causes. We will call phenomena possessing this property, random phenomena. In fact, any phenomena whose outcome is not deterministic will be called random.

The tides were probably present long before observations were taken and probably will continue long after we cease taking observations.

This means that our set of observations is just a finite segment of a much larger (possibly infinite) set of observations. The question then arises: how do we analyze this finite segment knowing that our data (information) is incomplete? Further, is it possible to detect the presence of periodic components (band-limited function) from only a finite set of observations (time-limited function)?

Free oscillations of the earth are the vibration modes of an elastic earth when excited by large energy sources such as large earthquakes. While speculation concerning these oscillations arose many years ago, it is only recently that accurate instruments have been constructed which can detect these oscillations whose periods are from 1/2 to 60 minutes long. The availability of high speed computers has certainly aided in the analysis of these oscillations.

Fourier and spectral analysis applied to the seismographic observations of these oscillations has enabled Geophysicists to achieve a high degree of agreement between the theoretical prediction and experimental verification of certain elastic properties of the earth. In fact, the observational data has helped to distinguish between several alternative theories (models of the earth) concerning the composition of the earth's crust. Unfortunately for Geophysicists, these oscillations have only been observed in very large, and hence rare, earthquakes.

The term "Fourier analysis" generally refers to analysis using Fourier integrals or Fourier series.

Let $L^2(R)$ denote the space of complex-valued, Lebesque square-integrable functions defined on the set of real numbers R. Plancherel's Theorem states that the Fourier integral of a function f, in $L^2(R)$, converges in the mean to a function g, in $L^2(R)$. Wiener (1958) shows that the Fourier integral is a linear operator, called the Fourier transform, from $L^2(R)$ into $L^2(R)$ which is invertible; that is, there exists an inverse Fourier transform. One of the important applications of the theory of Fourier transforms is contained in Parseval's Theorem (Parseval's equality).

From this theory Slepian and Pollak defined the subspaces of time- and band-limited functions. A time-limited function in $L^2(R)$ is a function which vanishes outside the interval $(-T, T)$, for some $T > 0$. A band-limited function, in $L^2(R)$, is a function whose Fourier transform vanishes outside the interval $(-\Omega, \Omega)$, for some real $\Omega > 0$. Landau and Pollak showed that the only function which is both time-limited and band-limited is the zero function.

Slepian and Pollak proved that for any $T > 0$ and $\Omega > 0$, there exists a countably infinite set of functions $\{f_0(t), f_1(t), \ldots\}$ and a countably infinite set of positive real numbers $\lambda_0 > \lambda_1 > \ldots$ which satisfy:

1. $\{f_k(t)\}$ are band-limited, orthonormal on R and complete in the space of band-limited functions B,
2. $\{f_k(t)\}$ are orthogonal and complete in $L^2(-T, T)$, and
3. $\{\lambda_k\}$ and $\{f_k(t)\}$ are the eigenvalues and eigenvectors of a certain integral eigenvalue problem, such that $\{f_k(t)\}$ are the prolate spheroidal wave functions.

In this thesis we have found numerical analogues for certain of these results. Unfortunately, we were not able to find numerical analogues for all of them. Specifically, we develop discrete analogues of the time-limited and band-limited functions as two finite dimensional complex vector spaces. Further, we define the discrete Fourier transform as a linear operator from the space of time-limited functions into the space of band-limited functions and we show that this operator is one-to-one and onto. However, because of the isomorphism between two N-dimensional vector spaces, care must be taken in the use of these functions since no non-zero time-limited function can be simultaneously band-limited and conversely. In fact, the lack of specification of the domain and range spaces for the discrete Fourier transform is a frequent source of confusion in the literature on this subject. Whereas in Slepian and Pollak the specification (namely, T and Ω) of the spaces of time-limited and band-limited functions are chosen independently, in our theory the two spaces have the same finite dimension.

Let $L^2(0, 2\pi)$ denote the space of complex-valued, Lebesque square-integrable functions defined on the interval $(0, 2\pi)$. Although it is not usually done, we can define a linear operator (the finite Fourier transform) on $L^2(0, 2\pi)$ whose image is a function (called a sequence) in l^2, the space of all complex-valued sequences $\langle a_n \rangle$ such that $\Sigma |a_n|^2$ converges. Given a function f, defined on $(0, 2\pi)$, the complex Fourier series representation for f(t) is defined as

$$\sum_{n=-\infty}^{\infty} c_n \, e^{int} \qquad (1)$$

where $i^2 = -1$ and for all integers n,

$$c_n = \frac{1}{2\pi}\int_0^{2\pi} f(t)\, e^{-int}\, dt. \tag{2}$$

Since the Fourier series, Equation (1), is a periodic function with period 2π, we can extend f(t) to the set R by requiring f(t) to be periodic with period 2π; that is, we require that, for all t,

$$f(t + 2\pi) = f(t). \tag{3}$$

The Riemann-Lebesque Theorem states that the sequence $\langle c_n \rangle$, in Equation (2), converges to zero as n approaches either plus or minus infinity. The Riesz-Fischer Theorem states that for any trigonometric series, Equation (1), for which $\Sigma |c_n|^2$ converges, there is a function in $L^2(0, 2\pi)$ for which Equation (1) is its Fourier series representation (Carslaw; Apostol). Furthermore, there is a Parseval's Theorem which is analogous to the theorem of the same name in the theory of Fourier transforms.

If we select a finite sequence of complex numbers, $\langle a_0, a_1, \ldots, a_{N-1} \rangle$, a discrete time-limited function, then the finite series

$$\frac{1}{N} \sum_{n=0}^{N-1} a_n\, e^{-int} \tag{4}$$

is a discrete band-limited function when $t = 2\pi k/N$ for $k = 0, 1, \ldots, N-1$. Equation (4) is a finite analogue of a Fourier series since it is a particular Riemann sum of the integral in Equation (2).

Consequently, we define the space of discrete time-limited functions, TL(N), as the space of all complex-valued N-tuples (finite sequences having N terms) and space of discrete band-limited functions, BL(N), as the space of all complex-valued sequences $\langle c_k \rangle$ where

$$c_k = N^{-1} \sum_{n=0}^{N-1} a_n W^{-nk}, \tag{5}$$

$\langle a_0, a_1, \ldots, a_{N-1} \rangle \in TL(N)$, and $W = \exp(i2\pi/N)$. Since $W^k = W^{k+N}$ then $c_k = c_{k+N}$, for $k = 0, 1, \ldots$.

We remark at this point that in most instances where we have specified complex numbers or complex-valued functions, we could have specified real numbers or real-valued functions with little or no changes. This illustrates how easily most results generalize from the real to complex number fields. In addition, we do achieve some economy of form and presentation. For example, the (complex) discrete Fourier transform, Equation (5), is easier to write down than the corresponding real expressions because of the necessity of treating the cases of N odd and N even separately.

Another approach to the development of the discrete theory of Fourier transforms is taken from Dr. Frank A. Hadsell's course notes in Linear Systems given at the Geophysics Department at the Colorado School of Mines (see also Papoulis). This approach uses the theory of generalized functions (distributions) and is also motivated by the desire to find numerical analogues of the results in the theory of Fourier transform which can be applied to electronic computers. We have modified Dr. Hadsell's development slightly.

Given a complex-valued function f(t) which is time-limited to the interval [0, T], we multiply f(t) by a sampling function

$$s_N(t; \tau) = \sum_{m=0}^{N-1} \delta(t - m\tau) \tag{6}$$

where $\delta(t)$ is the Dirac delta (generalized) function, $T = N\tau$ and

$$f_s(t) = f(t)\, s_N(t; \tau) = \sum_{m=0}^{N-1} f(m\tau)\, \delta(t - m\tau). \tag{7}$$

This approach uses the generalized function $f_s(t)$ in place of the discrete time-limited function $\langle f_m = f(m\tau)\rangle$ which we use in our theory. The definition of the Fourier transform for generalized functions results in the Fourier transform of $f_s(t)$ being given by

$$F_s(\omega) = \sum_{m=0}^{N-1} f(m\tau)\, e^{-im\tau\omega}. \tag{8}$$

$F_s(\omega)$ is periodic in ω with period $2\pi/\tau = 2\pi N/T$ (F_s is a band-limited function). Since $F_s(-\omega) = F_s((2\pi/\tau) - \omega)$, the Fourier transform for $|\omega| > \pi/\tau$ is determined ("aliased") by the Fourier transform for $|\omega| \leq \pi/\tau$ (Nyquist frequency).

Since $F_s(\omega)$ is a continuous function, F_s can be sampled at an arbitrary number of values of ω in the interval $[0, 2\pi N/T]$ resulting in a discrete band-limited function.

In contrast to this situation, our discrete band-limited function $\langle c_k\rangle$, given in Equation (5), has only N values (N is determined by the discrete time-limited function) and $\langle c_k\rangle$ is related to $F_s(\omega)$ by

$$c_k = N^{-1} F_s(2\pi k/T) \tag{9}$$

for $k = 0, 1, \ldots, N-1$.

Why does our definition of the discrete Fourier transform require N values whereas the second definition allows an arbitrary number of values of F_s? (this question was one of the motivations for this thesis).

We believe the answer to this question lies in the observation that the second approach assumes the existence of the function f(t) whereas our approach only assumes the existence of the N complex numbers $\langle f_m \rangle$. We observe that the finite sampling process is not invertible. That is, there are many functions whose equally-spaced values result in a given set of N complex numbers $\langle a_n \rangle$. Consequently, our discrete Fourier transform of $\langle a_n \rangle$ represents the discrete Fourier transform of the set of functions whose equally-spaced values are given by $\langle a_n \rangle$. In Chapter Two we derive some of the properties of the discrete Fourier transform of the N-tuple $\langle a_n \rangle$ (without assuming the existence of the function f(t)).

The numerical calculation of the discrete Fourier transform is, at the present time, most efficiently computed using the "fast Fourier transform" algorithm, rediscovered in 1965 by J.W. Cooley and J.W. Tukey. This basic algorithm, leading to numerous variations and modifications, has brought about a substantial saving in computation time (for large N) and has sparked the search for new ways to apply the discrete Fourier transform to other problems. (see Chapter Four for a further discussion).

After we have applied the discrete Fourier transform to a N-tuple $\langle a_n \rangle$, we have the problem of properly interpreting the results. We comment further on this point in Chapter Four.

In Chapter Two, the spaces TL(N), discrete time-limited functions, and BL(N), discrete band-limited functions, are shown to be two N-dimensional complex vector spaces. The discrete Fourier transform is shown to be an isomorphism (a one-to-one mapping) of TL(N) onto BL(N). When two finite dimensional vector spaces are isomorphic to each other, they "tend to look alike algebraically". It is therefore possible to substitute a discrete band-limited function for a discrete time-limited function and vice versa.

What are the ramifications of replacing one discrete function by the other? Given a result about one space, can we obtain a similar result for the other space? (these two questions are also part of the motivation for this thesis).

The first question is particularly pertinent when we attempt to define the discrete analogue of the convolution operation. For example we could define the convolution of two discrete time-limited functions, N-tuples $\langle a_n \rangle$ and $\langle b_n \rangle$, as the (2N-1)-tuple $\langle d_m \rangle$, where, for $m = 0, 1, \ldots, 2N-2$,

$$d_m = \Sigma a_n \, b_{m-n} \tag{10}$$

is defined and summed over all n such that $0 \leq n \leq N-1$ and $0 \leq m-n \leq N-1$. Unfortunately the discrete Fourier transform of $\langle d_m \rangle$ is not related to the transforms of $\langle a_n \rangle$ and $\langle b_n \rangle$ in any nice way. Often, for large N, the numbers d_m become small at each end ($m = 0$ and $m = 2N-2$) because the number of terms in the sum, in Equation (10), is small.

On the other hand, we could define the convolution of two discrete band-limited functions, $\langle a_n \rangle$ and $\langle b_n \rangle$, as the (2N-1)-tuple $\langle d_m \rangle$, in

Equation (10), where now the summation is over all n between 0 and N-1, since $\langle a_n \rangle$ and $\langle b_n \rangle$ can be extended periodically. This is an example where discrete time-limited functions are replaced by discrete band-limited functions. However, this replacement changes the nature of $\langle d_m \rangle$. Discrete band-limited functions $\langle a_n \rangle$ and $\langle b_n \rangle$ lead to discrete band-limited function $\langle d_m \rangle$ which may not become small near the ends; in fact, $\langle d_m \rangle$ is periodic with period N as we prove in Chapter Two. Also, the convolution operation is commutative when band-limited functions are used but noncommutative when time-limited functions are used. Consequently the discrete convolution theory is considerably less satisfactory than the corresponding theory in Fourier transforms and it has been omitted for that reason. A similar difficulty arises in the use of discrete sample paths (realizations) in estimating the covariance function of stochastic processes (Box and Jenkins).

A finite set of observations of any geophysical phenomenon can be viewed as a finite time series. That is, the observations can be considered as a finite segment of a realization of some stochastic process. The question arises as to how we are to consider our finite segment relative to those observations which are not known to us.

The concept of a stationary stochastic process is introduced when the statistical properties of the phenomenon are assumed to be time invariant. The statistical parameters of a stationary time series can be estimated from a segment of the time series.

We could assume that the phenomenon is periodic, in which case we would also assume that the stochastic process is periodic. There may,

in fact, be several different models proposed which approximate the phenomenon under investigation. The observations could then be used to accept, reject, or modify the various models proposed.

One of the reasons why Fourier analysis has been used as an important technique in analyzing various phenomena is that it presents essentially two different views ("time" and "frequency") of the phenomenon. For this reason, Fourier analysis is even applied to phenomenon for which its "frequencies" have no obvious physical interpretation.

In section 2.2 we look at an example, due to Hannan, of a discrete stochastic process which is circular. The N independent random variables of this process forms a basis for a N-dimensional complex vector space. Another basis is obtained which consists of N uncorrelated random variables which are related to the original random variables by the discrete Fourier transform. These two sets of random variables are the statistical counterparts of the discrete time- and band-limited functions.

In Theorems 2.9 and 2.10 we prove that for every discrete periodic stochastic process which has mean value zero and is stationary in the wide sense, there exists another discrete stochastic process which has mean value zero and uncorrelated random variables, and conversely. These two stochastic processes are related by the discrete Fourier transform. The discrete Fourier transform can then be viewed as a linear operator on this vector space mapping one basis (stochastic process) into the other basis.

Since the original stochastic process is stationary in the wide sense, the covariance matrix for this process is a Hermitian, Toeplitz

matrix (Toeplitz matrices are defined in Chapter Three). Theorem 2.9 is then equivalent to the result that under a unitary transformation, the covariance matrix of the original process can be transformed into a diagonal matrix (the diagonal matrix being the covariance matrix for the process having uncorrelated random variables). This is related, in a sense, to the fact that every Hermitian matrix can be diagonalized and consequently we should be able to extend Theorems 2.9 and 2.10 by removing the periodic condition. In Chapter Two we generalize Theorems 2.9 and 2.10 with the result that the two processes are related by a generalized discrete Fourier transform which we call the discrete Karhunen-Loeve expansion.

The expansion is called the discrete Karhunen-Loeve expansion because it is a discrete analogue of the Karhunen-Loeve expansion for continuous random variables. This expansion is determined by the eigenvectors of the covariance matrix of the original process; the eigenvectors form the uncorrelated random variables. The Karhunen-Loeve expansion has certain optimal properties which can be applied to problems in signal detection, factor analysis, and pattern recognition (Watanabe: Davenport and Root; Selin). Unfortunately, we were unable to formulate analogous optimal properties for the discrete Karhunen-Loeve expansion. We will comment further on this point in Chapter Four.

In Chapter Three we define Toeplitz matrices, develop their properties, derive numerical algorithms for (1) inverting Toeplitz matrices, (2) solving a Toeplitz system of algebraic equations, and (3) discuss algorithms for finding eigenvalues and eigenvectors of Toeplitz matrices.

Algorithms for handling Toeplitz matrices are of great interest at the present time due to the role these matrices play in the theory and application of discrete stationary stochastic processes. The applications include estimating certain digital filters (Levinson; Wiener, 1949), fitting autoregressive processes (Box and Jenkins; Jenkins and Watts), and finding the discrete Karhunen-Loeve expansion. Grenander and Szego give other applications.

Levinson derived the first algorithm for solving a Toeplitz system of algebraic equations which made efficient use of the special structure of Toeplitz matrices. There is a slightly different version of the algorithm in Wiggins and Robinson but both algorithms require the Toeplitz matrix to be real and symmetric. We have derived an algorithm in section 3.3 which removes this restriction.

The Trench-Zohar algorithm for inverting a Toeplitz matrix is given in section 3.2. W.F. Trench originally derived this algorithm for Hermitian, Toeplitz matrices. Zohar simplified the derivation and extended the algorithm to a slightly larger class of matrices. However both authors were apparently unaware that, in general, the inverse of invertible Toeplitz matrices is not a Toeplitz matrix.

In section 3.1 we clarify the Trench-Zohar algorithm by finding a larger class of matrices, called generalized Toeplitz matrices, which includes Toeplitz matrices and their inverses (when they exist). We also specify several classes of matrices which are related to Toeplitz matrices.

In sections 3.1 and 3.4 we derive several results relating to the solution of the eigenvalue problem for Toeplitz matrices. Because the discrete Karhunen-Loeve expansion, in Chapter Two, requires the eigenvalues and eigenvectors of Toeplitz matrices, we would like to be able to derive explicit expressions for these eigenvalues and eigenvectors.

While we were not able to find these explicit expressions, we mention that the Jacobi and Householder Methods are numerical algorithms for finding the numerical eigenvalues and eigenvectors. Furthermore, we derive and discuss the eigenvalue problem for circulant matrices which are the (Toeplitz) covariance matrices for periodic processes. Gray and Pearl have derived similar, but different, results on circulant matrices which are "asymptotically Toeplitz" as the order of the matrices increase without bound.

Finally, in Chapter Four we conclude with a discussion of the discrete (numerical) analogues we have derived, their limitations and possible further work.

Chapter 1. Definitions and Examples

1.1 Definition (field). A set of elements F is said to be a field if for every $a,b,c \in F$,

A. (Addition) there is a unique sum $a + b \in F$, and

1. $a + b = b + a$ (commutative)
2. $a + (b + c) = (a + b) + c$ (associative)
3. there is a $0 \in f$ such that $a + 0 = a$
4. for every $a \in F$, there is a $b \in F$ such that $a + b = 0$, where b is usually written as $-a$,

B. (Multiplication) there is a unique product $a\,b \in F$, and

1. $a\,b = b\,a$ (commutative)
2. $a(b\,c) = (a\,b)c$ (associative)
3. there is a $1 \in F$ such that $a1 = 1a = a$
4. for every $a \neq 0$, there is a b such that $ab = 1$ where b is usually written as $1/a$ or a^{-1},

C. 1. $a(b + c) = a\,b + a\,c$ (distributive)

2. $a - b = a + (-b)$ (subtraction).

The set of all rational numbers, real numbers and complex numbers are examples of fields. Since we will be using them later, let us denote the set of all real numbers by R and the set of all complex numbers by C. When we can use either we will denote this by F.

The above definition contains a few of the "laws of Arithmetic" and these can be used to prove the following results: (i) $0 + a = a$, (ii) $-a = (-1)a$, (iii) if $a\,b = 0$ then either $a = 0$ or $b = 0$.

Definition (linear vector space). A set V is said to be a (linear) vector space over the field F (called scalars) if for every $\underline{u}, \underline{v}, \underline{w} \in V$ and every $a,b \in F$,

A. (Addition) there is a unique sum $\underline{u} + \underline{v} \in V$, and

1. $\underline{u} + \underline{v} = \underline{v} + \underline{u}$ (commutative)
2. $\underline{u} + (\underline{v} + \underline{w}) = (\underline{u} + \underline{v}) + \underline{w}$ (associative)
3. there is a $\underline{0} \in V$ such that $\underline{u} + \underline{0} = \underline{u}$
4. for every $\underline{u} \in V$ there is a $\underline{v} \in V$ such that $\underline{u} + \underline{v} = \underline{0}$, where this $\underline{v}$ is usually written as $-\underline{u}$,

B. (Scalar Multiplication) there is a unique product $a\,\underline{u} \in V$, and

1. $a(b\,\underline{u}) = (a\,b)\underline{u} = a\,b\,\underline{u} = \underline{u}\,a\,b$
2. $a(\underline{u} + \underline{v}) = a\,\underline{u} + a\,\underline{v}$ (distributive)
3. $(a + b)\underline{u} = a\,\underline{u} + b\,\underline{u}$ (distributive)
4. $1\,\underline{u} = \underline{u}$,

C. (Subtraction) $\underline{u} - \underline{v} = \underline{u} + (-\underline{v})$.

When the field F is R then V is called a real vector space and when F is C then V is called a complex vector space.

From the above definitions we can obtain the following results: (i) $a\,\underline{0} = \underline{0}$, (ii) $0\,\underline{u} = \underline{0}$, (iii) $-\underline{u} = (-1)\underline{u}$ and (iv) if $a\,\underline{u} = \underline{0}$ then either $a = 0$ or $\underline{u} = \underline{0}$.

The similarity of the two definitions is striking and in fact the field F is a vector space over itself.

We note however that vectors do not possess any multiplicative inverses; that is, there is no division by vectors.

We also note in the above definition that multiplication by scalars illustrates how we can obtain many vectors from a single non-zero vector by choosing different scalars.

This property suggests that we might try to represent every vector in V, if this is possible, as a product of a scalar and some "basic" vector, or as a combination of such products. In fact, this approach is the most common approach to the study of vector spaces. Namely, we seek a subset B of vectors in the vector space V in which each vector in V can be represented as a (unique) sum of scalar products of vectors from B.

Definition (linear combination). A vector $\underline{x} \in V$ is said to be a linear combination of the vectors $\underline{u}, \underline{v} \ldots, \underline{w}$, in V, if for a corresponding set of scalars $\{a, b, \ldots, c\}$, $\underline{x} = a\,\underline{u} + b\,\underline{v} + \ldots + c\,\underline{w}$.

Definition (linear dependence, linear independence). A finite set of vectors $\{\underline{u}, \underline{v}, \ldots, \underline{w}\}$ is said to be linearly dependent if there is a corresponding set of scalars $\{a, b, \ldots, c\}$, not all of which are zero, such that $a\,\underline{u} + b\,\underline{v} + \ldots + c\,\underline{w} = \underline{0}$.
If not, then the set of vectors is said to be linearly independent.

An equivalent definition of linear independence of a finite set of vectors, $\{\underline{u}, \underline{v}, \ldots, \underline{w}\}$, is $a\,\underline{u} + b\,\underline{v} + \ldots + c\,\underline{w} = \underline{0}$ if and only if $a = b = \ldots = c = 0$.

An infinite set of vectors is linearly dependent if any finite subset of these vectors is linearly dependent. An infinite set of vectors is linearly independent if every finite subset of these vectors is linearly independent.

The following is a fundamental result of the above definitions.

Lemma 1.1 The finite set of non-zero vectors $\{\underline{u}, \underline{v} \ldots, \underline{w}\}$ is linearly dependent if and only if for some vector $\underline{x}$ in this set, $\underline{x}$ is a linear combination of the other vectors in the set.

Definition (basis; coordinate system). A basis B for a vector space V is a linearly independent subset of vectors in V such that every vector in V can be expressed as a linear combination of vectors in B.

With the aid of Zorn's Lemma it can be proved that every vector space has a basis.

While, in general there may be many bases for a given finite dimensional vector space, it can be proved that the number of vectors in each basis is the same as in any other basis (Halmos).

Definition (finite or infinite dimensional). A vector space V is said to be finite dimensional if V has a basis consisting of a finite number of vectors. If not, then V is infinite dimensional. A finite dimensional vector space V having a basis containing n vectors is said to be n-dimensional.

Lemma 1.2 For a n-dimensional vector space V,

1. every subset of n+1 vectors from V is linearly dependent,

2. a subset X of V having n vectors is a basis for V if and only if X is linearly independent,
3. A subset X of n vectors from V is a basis if and only if every vector in V can be expressed as a linear combination of vectors in X.

Definition (subspace). A non-empty subset M of a vector space V is said to be a subspace of V if for every pair of vectors $\underline{u}$ and $\underline{v}$ belonging to M, a $\underline{u}$ + b $\underline{v}$ belongs to M for all scalars a and b belonging to F.

Therefore a subspace M is itself a vector space over the field of scalars.

Lemma 1.3 The intersection of any collection of subspaces is a subspace.

Definition (span). If X is a subset of a vector space V then the span of X is the intersection of all subspaces containing X. (X need not be a subspace itself.)

Consequently the span of X is a subspace containing X and is contained in every subspace which contains X.

Lemma 1.4 If X is any subset of a vector space V then S is the span of X if and only if S is the set of all linear combinations of vectors from X.

Therefore we can say that a basis B for a vector space V is a linearly independent subset of vectors from V whose span is V. Further, the dimension of every subspace of a n-dimensional vector space V is less than or equal to n.

The concept of a span of a set of vectors suggests that we can approach the study of vector spaces by considering the span of certain linearly independent sets of vectors.

1.2 Examples of Vector Spaces

This list is by no means complete or exhaustive.

A. Sequence Spaces

1. Let $\{a_1, a_2, \ldots, a_n\}$ be any finite set of non-zero scalars from the field F, F(n) be the set of all "n-tuples" from F,

$\underline{e}_1 = (a_1, 0, 0, \ldots, 0)$,

$\underline{e}_2 = (0, a_2, 0, \ldots, 0)$,

. . . . . .

$\underline{e}_n = (0, 0, 0, \ldots, a_n)$, and

$B = \{\underline{e}_1, \underline{e}_2, \underline{e}_3, \ldots, \underline{e}_n\}$.

If $\underline{x} = (b_1, b_2, \ldots, b_n)$ and $\underline{y} = (c_1, c_2, \ldots, c_n)$ are n-tuples of scalars from F then we define

$\underline{x} + \underline{y} = (b_1 + c_1, b_2 + c_2, \ldots, b_n + c_n)$, and

$a\,\underline{x} = (a\,b_1, a\,b_2, \ldots, a\,b_n)$. F(n), the span of B, is a n-dimensional vector space over the field F.

Definition (onto). A function g defined on a vector space V with range (image) in a vector space W is said to be onto W if for every vector $w \in W$ there exists a vector $v \in V$ such that $g(v) = w$.

Definition (one-to-one). A function g defined on a vector space V with range in a vector space W is said to be one-to-one if for all

$v_1, v_2 \in V$, $v_1 \neq v_2$ implies that $g(v_1) \neq g(v_2)$. A logically equivalent statement is that $g(v_1) = g(v_2)$ implies $v_1 = v_2$, for all v_1, v_2 in V.

Definition (isomorphism). Two vector spaces V and W are said to be isomorphic if there is a function g from V onto W which is one-to-one and $g(a \underline{u} + b \underline{v}) = a\, g(\underline{u}) + b\, g(\underline{v})$.

Lemma 1.5 Every n-dimensional vector space V is isomorphic to F(n).

This lemma suggests, among other things, that the concept of a vector space is too general in that it cannot distinguish between "different" n-dimensional vector spaces over the same field. Later in this chapter we will define some properties which will help distinguish between these different vector spaces.

We remark at this point that when F = R (real numbers) then R(1) is isomorphic to R and R(3) is isomorphic to the set of (equivalence classes of) directed line segments in three dimensional space where vector addition is given by the "parallelogram law."

Definition (sequence). A (infinite) sequence is a function defined on the set of positive integers into the field F. These sequences will be designated by $\langle a_n \rangle$.

2. Let $\{a_1, a_2, \ldots\}$ be any countable set of non-zero complex numbers, V be the set of all complex sequences,

$\underline{e}^1 = \langle a_1, 0, \ldots \rangle$,

$\underline{e}^2 = \langle 0, a_2, \ldots \rangle$, and so on, with

$B = \{\underline{e}^1, \underline{e}^2, \ldots\}$.

If $\underline{x} = \langle b_i \rangle$ and $\underline{y} = \langle c_i \rangle$ are sequences of complex numbers then we define

$\underline{x} + \underline{y} = \langle b_1 + c_1, b_2 + c_2, \ldots \rangle = \langle (b_i + c_i) \rangle$, and

$a\,\underline{x} = \langle a\,b_1, a\,b_2, \ldots \rangle = \langle (a\,b_i) \rangle$.

V, the span of B, is a complex vector space over C which is infinite dimensional.

If we make a slight change in notation in C(n), example 1 above, then we see that C(n) is (isomorphic to) a finite dimensional subspace of V, namely, the subspace of V consisting of all vectors $\underline{x} = \langle b_i \rangle$ such that $b_i = 0$ for integers i greater than the (given) positive integer n. We will have more to say about C(n) in Chapter Two.

3. Let l_0 be the subset of V, in example 2, consisting of all complex sequences which possess only a finite number of non-zero values; that is, $\underline{x} = \langle b_i \rangle \in l_0$ if and only if $b_i = 0$ for all integers i greater than some positive integer. Clearly C(n) is a subspace of l_0 and l_0 is infinite dimensional.

Definition (convergence). A (real) complex sequence $\langle a_i \rangle$ is said to converge if there is a (real) complex scalar a, called the limit of the sequence, such that for each positive real number ε there is a positive integer N such that $|a_n - a| < \varepsilon$ whenever n is greater than N.

Definition (absolute convergence). A (real) complex sequence $\langle a_n \rangle$ is said to be absolutely convergent if the infinite series $\Sigma|a_n|$ converges. (Here the vertical lines denote absolute value.)

Lemma 1.6 An absolutely convergent sequence converges to zero.

Definition (Cauchy sequence). A (real) complex sequence $\langle a_n \rangle$ is said to be Cauchy if for each positive real number ε there is a positive integer N such that $|a_n - a_m| < \varepsilon$ whenever n and m are greater than N.

Lemma 1.7 A (real) complex sequence converges if and only if it is Cauchy.

Definition (bounded sequence). A (real) complex sequence is said to be bounded if there is a positive real number K such that $|a_n| \leq K$ for all integers n.

4. Let c_0 be the subset of V, the set of all complex sequences, consisting of those sequences which converge to the limit zero.

5. Let c be the subset of V consisting of those sequences which are Cauchy.

6. Let l_1 be the subset of V consisting of all absolutely convergent sequences.

7. Let b be the subset of V consisting of all bounded sequences.

Lemma 1.8 $C(n) \subset l_0 \subset l_1 \subset c_0 \subset c \subset b$. That is, each of these sets is a subspace of the succeeding vector space.

8. Let l^p be the set of all sequences $\langle a_n \rangle$ such that the infinite series $\Sigma|a_n|^p$, $1 \leq p < \infty$, converges.

Since a sequence is a function defined on the set of positive integers it is not suprising that other collections of functions also form vector spaces.

B. Function Spaces

Consider the set of all real-valued functions defined on a given (fixed) bounded interval. Let $\underline{x} = f(t)$ and $\underline{y} = g(t)$ for each t in the real interval I. Define

$$\underline{x} + \underline{y} = f(t) + g(t) \quad , t \in I$$

$$a\,\underline{x} = a\,f(t) \quad , t \in I, a \in R.$$

Most of the vector spaces derivable from these definitions will be infinite dimensional.

In some instances the requirement that the given interval be bounded can be removed but invariably the interval must be given in advance (i.e. fixed).

9. Let P be the set of all polynomials with (real) complex coefficients defined on (R) C where addition and scalar multiplication are the usual addition of polynomials and multiplication by a (real) complex number. If the degrees of the polynomials are required to be less than the positive integer n then we obtain a n-dimensional subspace of P, designated as P_n. The zero vector in P is the zero polynomial.

10. Let $L^p(I)$, $1 \leq p < \infty$ be the set of all functions f defined on a given (fixed) interval I such that $|f|^p$ is Lebesque integrable on I.

11. Let Per(T) be the set of all continuous periodic functions having a given (fixed) period $T > 0$; that is, if $f \in Per(T)$ then $f(t + T) = f(t)$, for all t.

The set of all mxn matrices whose elements are (real) complex numbers is a vector space.

1.3 Definition (normed linear space). A vector space V is said to be a normed linear space if there is a function (called a norm) defined on V into R, designated by $||\ ||$, satisfying

1. $||\underline{x}|| \geq 0$; $||\underline{x}|| = 0$ if and only if $\underline{x} = \underline{0}$
2. $||a\underline{x}|| = |a|\ \ ||\underline{x}||$
3. $||\underline{x} + \underline{y}|| \leq ||\underline{x}|| + ||\underline{y}||$, for all $\underline{x}$ and $\underline{y}$ in V.

What has been defined up to this point establishes some of the algebraic "structure" of vector spaces. With the introduction of a norm we have added a topological "structure" characterizing a class of topological (vector) spaces. We note in passing that a normed linear space is a metric space, an important example of a topological space.

We will now give definitions for sequences of vectors in a normed linear space analogous to the definitions given above for sequences of (real) complex numbers.

Definition (Cauchy sequence). A sequence $\langle \underline{x}_n \rangle$ of vectors in a normed linear space V is said to be a Cauchy sequence if for each $\varepsilon > 0$ there is an integer N such that $||\underline{x}_n - \underline{x}_m|| < \varepsilon$ whenever n and m are greater than N.

Definition (limit). A sequence $\langle \underline{x}_n \rangle$ of vectors in a normed linear space is said to have the limit $\underline{x}$ if for each $\varepsilon > 0$ there is an integer N such that $||\underline{x}_n - \underline{x}|| < \varepsilon$ whenever $n > N$.

Definition (complete). A normed linear space V is said to be complete if every Cauchy sequence of vectors in V has a limit in V.

We will be primarily interested in the following norms in the remainder of this thesis:

A. Sequence Spaces. For $\underline{x} = \langle b_i \rangle$

$$||\underline{x}|| = \left(\sum_{i=1}^{\infty} |b_i|^2 \right)^{1/2}$$

B. Functions Spaces. For $\underline{x} = f(t)$, $t \in I$,

$$||\underline{x}|| = \left(\int_I |f|^2 \right)^{1/2}$$

Complete normed linear spaces are called Banach Spaces.

1.4 Definition (inner product space). A vector space is said to be an inner product space if there is a function defined on V x V into R, designated by (,), satisfying for all $\underline{x}, \underline{y}, \underline{z} \in V$,

1. $(\underline{x},\underline{x}) \geq 0$; $(\underline{x},\underline{x}) = 0$ iff $\underline{x} = \underline{0}$ (non-negative definite)
2. $(\underline{x},\underline{y}) = (\underline{y},\underline{x})$ (symmetric)
3. $(a\underline{x} + b\underline{y}, \underline{z}) = a(\underline{x},\underline{z}) + b(\underline{y},\underline{z})$ (linearity)

If the field of scalars is C instead of R then (2) becomes $(\underline{x},\underline{y}) = (\underline{y},\underline{x})^*$, where the asterisk means complex conjugate. Given an inner product we can always define a norm, induced by the inner product, by $||\underline{x}||^2 = (\underline{x},\underline{x})$, so that every inner product space is a normed linear space.

Definition (orthogonal; orthonormal). Two vectors $\underline{x}$ and $\underline{y}$, in an inner product space V, are said to be orthogonal if $(\underline{x},\underline{y}) = 0$. They are said to be orthonormal if they are orthogonal and $||\underline{x}|| = 1 = ||\underline{y}||$.

1.5 Definition (linear operator; linear functional). A function A from a vector space V into a vector space W, designated $A:V\to W$, is said to be a linear operator if $A(a\underline{x} + b\underline{y}) = aA\underline{x} + bA\underline{y}$ for all $\underline{x},\underline{y} \in V$ and for all $a,b \in F$. A linear functional is a linear operator whose range is the field of scalars. We note that $A\underline{0} = \underline{0}$. We will be chiefly interested in linear operators from a vector space V into itself, or an isomorphic image of V.

Two special linear operators are (1) the zero operator which maps every $\underline{x} \in V$ into $\{\underline{0}\}$, and (2) the identity operator which maps every $\underline{x} \in V$ onto itself. It is relatively easy to show that the set of all linear operators defined on a vector space V over a field F, into V, is itself a vector space over the same field. Indeed a great deal more can be said about this vector space since we can define a product of two such linear operators.

Definition (product of linear operators). If A and B are two linear operators defined on V, into V, then C = BA is said to be the product of A followed by B where C is a linear operator from V into V provided $C\underline{x} = B(A\underline{x})$ for all $\underline{x} \in V$.

We remark that the product of two linear operators is, in general, non-commutative and may possess divisors of zero. That is, (1) in general, $AB \neq BA$, and (2) AB may be the zero operator without either A or B being the zero operator. These two properties distinguish products (multiplication) of linear operators from products previously defined for a vector space.

Considering the analogy of the multiplication defined previously, we seek linear operators A and B such that AB and BA are the identity operator. Operator B would then be called the multiplicative inverse operator for A. Unfortunately, requiring the operator A to be non-zero is not the correct condition to impose on A in order that A possess an inverse operator.

Definition (one-to-one; onto; invertible). A linear operator A from a vector space V into V is said to be one-to-one if $\underline{x}_1 \neq \underline{x}_2$ implies that $A\underline{x}_1 \neq A\underline{x}_2$. A is said to be onto if for every vector $\underline{y} \in V$ there is a vector $\underline{x} \in V$ such that $\underline{y} = A\underline{x}$. A is said to be invertible if there is an operator B, also from V into V, such that $AB = I = BA$.

Lemma 1.9 If A is a linear operator defined on a finite dimensional vector space V, into V, then A is one-to-one if and only if A is onto.

Lemma 1.10 If A is a linear operator defined on a finite dimensional vector space V, into V, then A is invertible if and only if A is one-to-one.

Definition (matrix of A). If A is a linear operator on a n-dimensional vector space V, into V, having a basis $B = \{\underline{x}_1, \underline{x}_2, \ldots \underline{x}_n\}$ then the set $\{a_{ij}\}$ of n^2 scalars is said to be the matrix of A associated with basis B if

$$A\underline{x}_j = \sum_{i=1}^{n} a_{ij}\underline{x}_i \ .$$

1.6 We now define the concepts of a stochastic process and a time series. These concepts involve collections of random variables defined on a probability space which is associated with a phenomenon described by probabilistic laws. These random variables arise when we attempt to construct quantitative (mathematical) models to explain the behavior of random phenomena (physical, economical, social, etc.) whose laws are either not well understood or for which a deterministic (non-probabilistic) description is inadequate or unknown.

Relative to a given random phenomenon, we make the following definitions. (Hoog and Craig; Parzen)

Definition (sample space). The set S of all possible outcomes associated with a given random phenomenon is called the sample space for that phenomenon.

Definition (sigma field). A collection $\mathcal{E}$ of subsets of a set S is called a sigma field if

1. S belongs to $\mathcal{E}$,
2. the set complement of every set in $\mathcal{E}$ also belongs to $\mathcal{E}$, and
3. for every sequence of sets $\langle S_n \rangle$, S_n belonging to $\mathcal{E}$, the union $\bigcup_{n=1}^{\infty} S_n$, belongs to $\mathcal{E}$.

Definition (events). The sigma field $\mathcal{E}$ of the sample space S which includes (are generated by) certain subsets (events) of S will be called the collection of events.

The same sample space may lead to many different collections of events depending upon which events are of interest.

The sigma field of all subsets of R, the set of real numbers, generated by all open intervals is called the family of Borel sets for R. The sigma field of C, the set of complex numbers, generated by all sets of the form $I_1 \times I_2$ (Cartesian product of I_1 and I_2) where I_1 and I_2 are intervals of real numbers, is called the family of Borel sets for C.

The collection of events $\mathcal{E}$ is defined specifically so that we can assign a probability measure on $\mathcal{E}$. That is, we can assign a probability to each event in $\mathcal{E}$.

Definition (probability function). The probability function associated with an event E, in $\mathcal{E}$, is a real-valued function defined on $\mathcal{E}$, denoted by P[ ], satisfying the properties:

1. $0 \leq P[E] \leq 1$, for every E in $\mathcal{E}$,
2. P[S] = 1, S is the "certain event", and
3. if for every sequence of events $\langle E_n \rangle$ where the E_n's are mutually exclusive events (i.e. disjoint sets), then

$$P[\bigcup_{n=1}^{\infty} E_n] = \sum_{n=1}^{\infty} P[E_n]$$

Definition (discrete sample space). A sample space S is said to be discrete if S has a countable (denumerable) number of elements.

Frequently we are interested in functions, called random variables, defined on the sample space S which enable us to quantify the events associated with the given random phenomenon. This is particularly helpful when we are interested in constructing quantitative models which are to be used in the prediction of the behavior of this random phenomenon.

In addition, the realizations of these random variables are used to estimate the statistical parameters which determine the nature of the underlying sample space. Most notable of these statistical parameters are the mean and variance.

Definition (random variable; realization). A random variable X is a complex-valued function defined on the sample space S for which the collection of events $\mathcal{E}$ has a probability function defined on it and for which the set {s: X(s) in B} is an event, for every Borel set B in C. The realization of X is the complex number X(s), for each $s \in S$.

Furthermore, we can define a probability function , denoted by Pr[ ], associated with the random variable X, which is induced by the probability function P[ ], defined on S by Pr[B] = P[s: $s \in S$ and $X(s) \in B$] for all Borel sets B in C.

Definition ($X \leq x$). The expression $X \leq x$, for each random variable X and complex number x, means that the real part of the realization of X is less than or equal to the real part of x and the imaginary part of the realization of X is less than or equal to the imaginary part of x. In abbreviated form we write $X \leq x$ if and only if $Re(X) \leq Re(x)$ and $Im(X) \leq Im(x)$.

Definition (distribution function). For each random variable X with its associated probability function Pr[ ], the distribution function for X is a real-valued function F defined on C by $F(x) = Pr[X \leq x]$. This distribution function, which always exists, can be viewed as a

joint distribution function of two real variables, namely, the real and imaginary part of the complex variable x. As such, this distribution function is a monotone non-decreasing function of each variable defined on R^2 into the closed interval [0, 1] and it has a right hand limit at every real value of each real variable.

We will be primarily interested in random variables (said to be of the finite type) whose range is a countable set. Consequently the following definitions will reflect this interest by involving summations rather than the integrations required in the more general case.

Definition (probability mass function). For every random variable X and its associated probability function Pr[ ], the probability mass function p(x) is a real-valued function defined on C by $p(x) = Pr[X = x]$ (when it exists).

At each x in C, this probability mass function can be viewed as a joint probability mass function of the real and imaginary parts of x.

This probability mass function satisfies (1) $0 \leq p(x) \leq 1$, for each x, and (2) $\Sigma\, p(x) = 1$ where the summation is over all x in C for which $p(x) > 0$.

Definition (stochastic process). A stochastic process is a collection of random variables denoted by $\{X_t: t \in T\}$ where T is some index set. The stochastic process is said to be discrete if T is a countable set and continuous otherwise.

Definition (time series). A time series associated with the stochastic process $\{X_t: t \in T\}$ is the set $\{X_t(s): t \in T\}$ for some $s \in S$. A time series is a realization of the stochastic process for each $s \in S$.

Definition (joint distribution function; joint probability mass function). The joint distribution function for random variables X and Y is given by $F(x,y) = Pr[X \leq x, Y \leq y]$ for every x,y in C. The joint probability mass function for (discrete) X and Y is given by $p(x,y) = Pr[X = x, Y = y]$ for every x, y in C (when it exists).

In an analogous manner we could give definitions for the joint distribution and probability mass functions for any finite collection of random variables. In addition, marginal and conditional distribution functions could be defined but we will not need them.

Definition (expectation). The expectation of a function G of a random variable X (of the finite type) is given by $E[G(X)] = \Sigma G(x)p(x)$, if it converges absolutely, where p(x) is the probability mass function for X and the summation is over the (countable) set of x's for which $p(x) > 0$.

Definition (mean; variance). The mean and variance of a random variable X are given by

$$Mn[X] = E[X],$$

$$Var[X] = E[|X - E(X)|^2] = E[|X|^2] - |E[X]|^2$$

respectively, when they exist.

Var[X] is a non-negative real number, when it exists, and its square root is the standard deviation of X.

Definition (characteristic function). The characteristic function associated with the complex-valued random variable X is a function of the real variables u and v given by $E[\exp(iuX_1 + ivX_2)]$, where X_1 is the real part of X, X_2 is the imaginary part of X and $i^2 = -1$.

Definition (independent random variables). The two random variables X and Y are said to be independent if $E[G(X)H(Y)] = E[G(X)]E[H(Y)]$ for all functions G and H for which these expectations exist.

There are other equivalent definitions of independence.

Definition (covariance). The covariance of two random variables X and Y, both having zero means, is given by $Cov[X,Y] = E[XY^*]$ when it exists. That is, $Cov[X,Y] = \Sigma\Sigma xy^*p(x,y)$.

Lemma 1.11 If X and Y are two independent random variables with zero means then $Cov[X,Y] = 0$.

Definition (stationary). The stochastic process $\{X_t: t \in T\}$ is said to be stationary if for every finite collection of random variables $\{X_1, X_2, \ldots, X_m\}$, the joint distribution function satisfies $F_{1, 2, \ldots, m}(x_1, x_2, \ldots, x_m) = F_{1+h, \ldots, m+h}(x_1, \ldots, x_m)$ for every h in which the joint distribution function on the right hand side exists.

Lemma 1.12 If the stochastic process $\{X_t: t \in T\}$ is stationary then

1. $Mn[X_t]$ is a constant

2. $\mathrm{Cov}[X_t, X_s]$ is a function only of $t - s$.

Definition (stationary in the wide-sense). A stochastic process $\{X_t: t \in T\}$ is said to be stationary in the wide-sense if $\mathrm{Mn}[X_t]$ is a constant and $\mathrm{Cov}[X_t, X_s]$ is a function only of $t - s$, for each $t,s \in T$.

Definition (correlation). The correlation of two random variables X and Y, both having zero means, is given by

$$\mathrm{Cor}[X, Y] = \frac{\mathrm{Cov}[X,Y]}{[\mathrm{Var}[X]\ \mathrm{Var}[Y]]^{1/2}} .$$

Definition (autocorrelation). The autocorrelation of the stochastic process $\{X_t: t \in T\}$ which is stationary in the wide-sense is given by $A_u = \mathrm{Cor}[X_t, X_{t+u}]$ whenever it exists.

We will assume throughout the rest of this thesis that our stochastic processes are stationary in the wide-sense and have zero mean value. The definition of stationary in the wide-sense is sometimes called covariance stationary or second-order stationary.

Chapter 2. Time- and Band-Limited Functions

2.1 Suppose we wish to describe (i.e. to model) the behavior of some random phenomenon which has an identifiable outcome (configuration, observable state); specifically, an outcome with some observable state given by a single complex variable at each instant of time. Let the totality of all such outcomes form the sample space S associated with the random phenomenon. Let the observable state be given by the complex-valued random variable $X_t(s)$ at time t and $s \in S$.

If we observe this phenomenon for N successive instants of time then we have a discrete (finite) family of random variables $\{X_t: t = 0, 1, 2, \ldots, N-1\}$ which forms a stochastic process and whose values form a time series.

We will not be concerned with outcomes which have occured before t = 0 but we will assume that the phenomenon continues to be observable for all succeeding times following t = 0 subject to the restriction that if $z_0, z_1, z_2, \ldots$ is a time series of the process (a sequence of observed values) then $\Sigma|z_n|^2$, over all n, $n \geq 0$, is convergent.

These observed values could, for example, be taken at equal intervals of time. The requirement that $\Sigma|z_n|^2$ be convergent implies that $|z_n|$ approach zero as n grows without bound (but not conversely).

The question arises as to how we should view our finite sample of observations in relation to the space of all possible sequences of observations (all possible time series of the discrete family of random variables).

Mathematically speaking we are concerned with the problem of embedding C(N), the space of all N-tuples of complex numbers, into l^2, the space of sequences, $\langle z_n \rangle$, of complex numbers such that $\Sigma|z_n|^2$ is convergent.

In one sense, we can view our N-tuple of observed values as an approximation to some assumed sequence of values in l^2. In most cases, however, the choice of the positive integer N is determined by factors unrelated to the phenomenon under observation and, in addition, the degree of approximation is usually not known.

One obvious way to embed C(N) into l^2 is to map C(N) isomorphically onto a N-dimensional subspace of l^2. This is done in the next section and this subspace is designated as TL(N).

However, we will see that there are advantages in considering other spaces in addition to TL(N). In particular we will be interested in another N-dimensional space, designated by BL(N), whose properties are quite distinct from TL(N) and yet, TL(N) and BL(N) are related by what we will call the discrete Fourier transform. This discrete Fourier transform is really analogous to the Fourier series associated with a complex-valued function f(t) defined on the interval [a, b] for which the Lebesque integral $\int_a^b |f(t)|^2 dt$ is finite.

Therefore, given N complex numbers we will assume that this N-tuple forms a finite segment of a time series associated with some stochastic process and that this time series is an element of l^2.

If we assume that realizations of X_t, for t = N, N+1, ..., are zero then the N-tuple will be considered an element of TL(N). On the other hand, if we assume that realizations of X_t are repetitive (i.e. the time series is a periodic sequence with period N) then the N-tuple will be considered an element of BL(N).

These two assumptions are not, of course, the only alternatives. However, without the presence of other information to the contrary, these two alternatives are reasonable and they do allow the use of certain analytic tools (e.g. discrete Fourier transform) which are not available under other assumptions.

2.2 Let us begin by specifying several of the inner product spaces previously defined with which we will be concerned.

Let l^2 be the set of all (infinite) sequences of complex numbers such that if $z = \langle z_n \rangle \in l^2$ then

$$||z||^2 = \sum_{n=0}^{\infty} |z_n|^2 \quad \text{is convergent.} \tag{1}$$

We will call $||z||^2$ the total energy of the sequence z. z_n will be called the nth component of z. The inner product of $u = \langle u_n \rangle$ and $v = \langle v_n \rangle$ in l^2 is defined by

$$(u, v) = \sum_{n=0}^{\infty} u_n v_n^* , \tag{2}$$

where (Cauchy-Schwartz inequality)

$$\left| \sum_{n=0} u_n v_n^* \right|^2 \leq \sum_{n=0}^{\infty} |u_n|^2 \sum_{m=0}^{\infty} |v_m|^2 . \tag{3}$$

l^2 is a complete inner product (Euclidean) space (every Cauchy sequence of elements in l^2 converges to an element of l^2) (Shilov).

Let TL(N) be the subspace of l^2 consisting of those sequences $a = \langle a_n \rangle$ for which $a_n = 0$ for all integers n greater than N-1. TL(N) is a N-dimensional Euclidean space of N-tuples of complex numbers which is isomorphic to C(N). We will frequently identify the elements $a = \langle a_n \rangle \in$ TL(N) by $a = \langle a_0, a_1, \ldots, a_{N-1} \rangle$. We will indicate the inner product of $a = \langle a_0, a_1, \ldots, a_{N-1} \rangle$ and $b = \langle b_0, b_1, \ldots, b_{N-1} \rangle$ in TL(N) by

$$(a, b)_N = \sum_{n=0}^{N-1} a_n b_n^* . \tag{4}$$

TL(N) will be called the space of time-limited functions.

Consequently, we can define a linear transformation (operator) T from l^2 into TL(N) by

$$T\langle u_n \rangle = \langle a_n \rangle \tag{5}$$

where
$$a_n = \begin{cases} u_n, & n = 0, 1, 2, \ldots, N-1 \\ 0, & n = N, N+1, \ldots . \end{cases}$$

We will refer to the sequence $\langle a_n \rangle$ as the time-limited version of the sequence $\langle u_n \rangle$. We extend the definition of $(\ ,\ )_N$ to l^2 by

$$(u, v)_N = \sum_{n=0}^{N-1} u_n v_n^* \tag{6}$$

and

$$||u||_N^2 = (u, u)_N \ . \tag{7}$$

This last real number is the energy of the sequence $u = \langle u_n \rangle$ contained in the first N components of u.

Definition (discrete Fourier transform). If $a = \langle a_n \rangle \in TL(N)$ then $c = \langle c_k \rangle$ will be called the discrete Fourier transform of a, denoted by $F\langle a_n \rangle = \langle c_k \rangle$, provided that for $k = 0, 1, \ldots$ and $i^2 = -1$,

$$c_k = N^{-1} \sum_{n=0}^{N-1} a_n \exp(-i2\pi kn/N). \tag{8}$$

(Definitions differ as to the presence or absence of 1/N in front of the summation sign.) We observe that the numbers c_k are, in general, complex even when the a_n's are real. We will also observe shortly that $\langle c_k \rangle$ is not an element of l^2 and, hence, not in TL(N).

We will develop a few of the properties of this discrete Fourier transform after first listing several results from complex variables in the following lemma.

Lemma 2.1 If α, β, β_n are complex numbers and * denotes complex conjugation, then

1. $\alpha^{**} = \alpha$
2. $(\alpha\beta)^* = \alpha^* \beta^*$
3. $|\alpha|^2 = \alpha\alpha^*$

4. $(\Sigma\alpha_n)^* = \Sigma\alpha_n{}^*$

5. $W^* = W^{-1}$

6. $W^N = W^0 = 1$

7. $W^{k+N} = W^k$, for all k, where we define

$$W = \exp(i2\pi/N). \tag{9}$$

Definition (periodic sequence). A sequence $u = \langle u_n \rangle$ is said to be periodic if there is a positive integer N such that $a_n = a_{n+N}$, for $n = 0, 1, 2, \ldots$. N will be called a period for such a sequence.

Lemma 2.2 For integers $k, N > 0$ and for $\Delta(k; N) = \sum_{n=0}^{N-1} W^{kn}$ we have that

1. for fixed N, $\Delta(k; N)$ is the kth component of a periodic sequence with period N,

2. $\Delta(k; N) = \begin{cases} N, \text{ for } k = 0, N \\ 0, \text{ for } k = 1, 2, 3, \ldots, N-1, \text{ and} \end{cases}$

3. the results are also true with W replaced by W^{-1}.

Proof: From Euler's formula: $\exp(ix) = \cos x + i \sin x$, we obtain

i) $\exp(i2\pi k) = 1$

ii) $\exp(\pm i\pi k) = \cos\pi k = (-1)^k$

iii) $W^{k+N} = W^k$, which proves (1).

For (2), we have

Case 1. $k = 0, N$: $\Delta(0; N) = \Delta(N; N) = \sum_{n=0}^{N-1} 1 = N.$

Case 2. $k = 1, 2, \ldots, N-1$: $W^k \neq 1$, and

$$\Delta(k; N)[1 - W^k] = \sum_{n=0}^{N-1} [W^{kn} - W^{kn+k}] = W^0 - W^{kN} = 0.$$

If we replace W by W^{-1} then, by Lemma 2.1, we obtain

$$\Delta(-k; N) = \sum_{n=0}^{N-1} W^{-kn} = \sum_{n=0}^{N-1} W^{-kn}W^{kN} = \sum_{m=0}^{N-1} W^{km} = \Delta(k; N).$$

Lemma 2.3 If $\langle a_n \rangle \in TL(N)$ and $\langle c_k \rangle$ is the discrete Fourier transform of $\langle a_n \rangle$, then $\langle c_k \rangle$ is a periodic sequence with period N.

Proof: The results follows from the definition, equation (8), and Lemma 2.1, since $W^{-n(k+N)} = W^{-nk}W^{-nN} = W^{-nk}$.

Theorem 2.4 If $\langle a_n \rangle \in TL(N)$ and $F\langle a_n \rangle = \langle c_k \rangle$ then

$$a_n = \sum_{k=0}^{N-1} c_k W^{kn}, \text{ for } n = 0, 1, 2, \ldots, N-1, \tag{10}$$

and

$$\sum_{n=0}^{N-1} |a_n|^2 = N \sum_{k=0}^{N-1} |c_k|^2 . \tag{11}$$

This last equation is known as Parseval's Equality. Proof: From the definition of F, Lemma 2.2 and for $n = 0, 1, 2, \ldots, N-1$,

$$\sum_{k=0}^{N-1} c_k W^{kn} = \frac{1}{N} \sum_{k=0}^{N-1} \sum_{m=0}^{N-1} a_m W^{-km}W^{kn} = \frac{1}{N} \sum_{m=0}^{N-1} a_m \sum_{k=0}^{N-1} W^{k(n-m)} = a_n.$$

Furthermore, $\sum_{n=0}^{N-1} |a_n|^2 = \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} c_k W^{nk} \sum_{m=0}^{N-1} c_m^* W^{-nm}$

$$= N \sum_{k=0}^{N-1} c_k c_k^* = N \sum_{k=0}^{N-1} |c_k|^2 .$$

We observe in Lemma 2.3, that the discrete Fourier transform of $\langle a_n \rangle \in TL(N)$ is a periodic sequence and consequently is not an element of l^2. We observe in Theorem 2.4 that the element $\langle a_n \rangle$ can be recovered from the first N components of $\langle c_k \rangle$ and that the total energy of $a = \langle a_n \rangle$ is given by

$$||a||^2 = (a, a)_N = \sum_{n=0}^{N-1} |a_n|^2 = N \sum_{k=0}^{N-1} |c_k|^2 \quad . \tag{12}$$

This suggests that there is N times "as much energy" in $\langle a_n \rangle$ as in one period of the discrete Fourier transform of $\langle a_n \rangle$.

Let BL(N) denote the space of all periodic sequences of complex numbers with period N with the usual definitions of addition and scalar multiplication. We will define the inner product on BL(N) by

$$(c, d)_N = N \sum_{k=0}^{N-1} c_k d_k^* \tag{13}$$

for $c = \langle c_k \rangle$, $d = \langle d_k \rangle \in BL(N)$.

In view of Theorem 2.4, the total energy of c is given by

$$||c||_N^2 = (c, c)_N \tag{14}$$

We will call BL(N) the space of band-limited sequences and it forms a complete complex Euclidean space. It is intuitatively clear that BL(N) is N-dimensional since each sequence is determined by the first N components and the requirement that it is periodic with period N.

We will show that BL(N) is N-dimensional by showing that the discrete Fourier transform F is one-to-one (F is invertible).

Theorem 2.5 The discrete Fourier transform F is a linear transformation from TL(N) into BL(N) which is invertible. Proof: The linearity F follows from $F(\langle a_n\rangle + \langle b_n\rangle) = F\langle a_n + b_n\rangle = F\langle a_n\rangle + F\langle b_n\rangle$. Let $\{e^1, e^2, \ldots, e^N\}$ be a basis for TL(N) where $e^1 = \langle 1,0,0,\ldots,0\rangle$, $\ldots$, $e^N = \langle 0, 0, 0, \ldots, 1\rangle$. The set $\{Fe^1, Fe^2, \ldots, Fe^N\}$ is in BL(N), where the kth component of Fe^j is $N^{-1}W^{-kj}$, and

$$\sum_{j=1}^{N} \alpha_j Fe^j = F\left(\sum_{j=1}^{N} \alpha_j e^j\right) = F\langle\alpha_n\rangle = \langle 0, 0, 0, \ldots, 0\rangle \text{ implies that}$$

$$0 = N^{-1} \sum_{n=1}^{N} \alpha_n W^{-kn} \text{ , for } k = 0, 1, 2, \ldots, N-1.$$

If we multiply the kth equation by W^{km} and sum on k, we obtain

$$0 = N^{-1} \sum_{k=0}^{N-1} W^{mk} \sum_{n=1}^{N} \alpha_n W^{-kn} = N^{-1} \sum_{n=1}^{N} \alpha_n \sum_{k=0}^{N-1} W^{k(m-n)} = \alpha_m \text{ ,}$$

for m = 1, 2, ..., N by Lemma 2.2. Hence $\{Fe^1, Fe^2, \ldots, Fe^N\}$ is a linearly independent set in BL(N). Now let $\langle c_k\rangle$ be any arbitrary sequence in BL(N). If we let

$$a_n = \sum_{k=0}^{N-1} c_k W^{kn}, \text{ for } n = 0, 1,2, \ldots, N-1, \text{ then}$$

$\mathbf{a} = \langle a_0, a_1, \ldots, a_{N-1}\rangle \in TL(N)$ and $F\langle a_n\rangle = \langle c_k\rangle$. Hence, F is onto and therefore F is invertible. Consequently BL(N) is N-dimensional.

Since TL(N) and BL(N) are both N-dimensional and therefore isomorphic to each other, we can establish another isomorphism between TL(N) and BL(N) by stating that $\mathbf{a} = \langle a_n\rangle \in TL(N)$ is isomorphic to $\mathbf{d} = \langle d_k\rangle \in BL(N)$ iff $d_k = a_k$ for k = 0, 1, 2, ..., N-1 . (15)

By altering the definition, we could define the discrete Fourier transform as a linear transformation from TL(N) onto TL(N), or from C(N) onto C(N). This situation tends to make TL(N) and BL(N) "look alike" yet, as we will see, it is necessary to distinguish between them.

Consequently we can define a linear transformation B from l^2 onto BL(N) by letting $B = F \circ T$. That is, B is the composition of T, defined in Equation (5), followed by F. Since F is invertible, $T = F^{-1} \circ B$ where the inverse of F is given implicitly by Equation (10).

Lemma 2.6 If $\langle a_n \rangle \in TL(N)$ and $F\langle a_n \rangle = \langle c_k \rangle$ then

1. $F\langle a_n^* \rangle = \langle d_k \rangle$, and
2. $a_n = a_n^*$ if and only if $c_k = d_k$ where

$$d_k = c^*_{mN-k} , \tag{16}$$

for some positive integer m such that mN-k is positive.

Proof: From Lemma 2.1 and the definition of F,

$$d_k = N^{-1} \sum_{n=0}^{N-1} a_n^* W^{-kn} = N^{-1} \left(\sum_{n=0}^{N-1} a_n W^{n(k-mN)} \right)^* = c^*_{mN-k} ,$$

where m is chosen so that mN-k is positive. Because of the periodicity it is sufficient to let m = 1.

Let us look at an example of how the discrete Fourier transform plays a part in the linear and statistical analysis of a finite stochastic process.

Suppose we have the set $B = \{X_0, \ldots, X_{N-1}\}$ of N linearly independent random variables and let V(N) be the N-dimensional complex vector space formed by the span of B.

Let us assume that these N random variables form a discrete (finite) stochastic process which is stationary in the wide sense and the covariance matrix for these random variables is given by

$$\mathrm{Cov}[X_t, X_s] = E[X_t X_s^*] = E[X_{t-s} X_0^*] \tag{17}$$

where $E[X_t] = 0$, for $t, s = 0, \ldots, N-1$.

If we define a linear transformation U on V(N) such that

$$UX_t = X_{t+1}, \quad t = 0, 1, \ldots, N-2$$

and

$$UX_{N-1} = X_0$$

(the process B is called a circular process). B forms a basis for V(N).

Define an inner product on V(N) by

$$(X, Y) = \mathrm{Cov}[X, Y] = \sum_t \sum_s \alpha_t \beta_s^* E[X_t X_s^*] \tag{18}$$

where $X = \Sigma \alpha_t X_t$, $Y = \Sigma \beta_s X_s$.

Here we have made use of the fact that the covariance matrix $E[X_t X_s^*]$ is Hermitian non-negative definite (Jenkins and Watts). These terms are defined in Chapter Three.

The matrix representation of U with respect to the basis B is

$$\begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

This matrix is a special case of a circulant matrix (Aitken) and a Toeplitz matrix, defined in Chapter Three.

The determinant of U is given by $\det(U) = (-1)^{N-1}$ since N-1 interchanges of the rows of U results in the identity matrix of order N.

The characteristic equation for U is

$$\det(U - \lambda I) = \lambda^N - 1 = 0 \tag{19}$$

whose eigenvalues are, for $j = 0, 1, \dots, N-1$,

$$\lambda_j = W^j = \exp(i2\pi j/N), \tag{20}$$

where W^{-j} could have been chosen instead. The eigenvector Y_j associated with the eigenvalue λ_j is a non-zero solution to $UY_j = \lambda_j Y_j$ given by

$$Y_j = \sum_{k=0}^{N-1} W^{jk} X_k \quad . \tag{21}$$

If we multiply Y_j by W^{-jm}, sum on j, and apply Lemma 2.2, we obtain for $m = 0, 1, \dots, N-1$,

$$X_m = N^{-1} \sum_{j=0}^{N-1} W^{-jm} Y_j \quad . \tag{22}$$

Since the eigenvalues are distinct, the set $\{Y_0, Y_1, \dots, Y_{N-1}\}$ forms an orthogonal basis for V(N).

The transformation U can be decomposed (Halmos) into

$$U = \sum_{k=0}^{N-1} W^k U_k \tag{23}$$

where the U_k's satisfy

$$\begin{aligned} &1.\ U_k^2 = U_k,\ k = 0, 1, \ldots, N-1, \\ &2.\ U_k U_j = 0,\ k \neq j, \text{ and} \\ &3.\ \sum_{k=0}^{N-1} U_k = I \text{ (identity transformation).} \end{aligned} \tag{24}$$

U_k is the (unitary) projection of V(N) into the span of Y_k and

$$U_k X_m = N^{-1} W^{-km} Y_k = n^{-1} \sum_{j=0}^{N-1} W^{k(j-m)} X_j \ . \tag{25}$$

Therefore the matrix associated with U_k with respect to the basis B has the element $N^{-1}W^{k(j-m)}$ in the (m+1) row and the (j+1) column.

In addition we have

$$X_m = U^m X_0 = \sum_{k=0}^{N-1} W^{-km} U_k X_0 = N^{-1} \sum_{k=0}^{N-1} W^{-km} Y_k, \tag{26}$$

$$Y_k = N U_k X_0 = \sum_{j=0}^{N-1} W^{kj} X_j, \tag{27}$$

$$(Y_j, Y_k) = N^2 (U_j X_0, U_k X_0) = 0,\ j \neq k, \tag{28}$$

and

$$(Y_k, Y_k) = \sum_{t=0}^{N-1} \sum_{s=0}^{N-1} W^{k(t-s)} E[X_t X_s^*]. \tag{29}$$

This example was taken from Hannan in a slightly modified form.

2.3 We would like to investigate several bases in the spaces TL(N) and BL(N). Since TL(N) and BL(N) are isomorphic to each other, the discrete Fourier transform could have been defined on either TL(N) or BL(N) with its range (or image) in either space. In fact the lack of specification of the domain and range of the discrete Fourier transform is often a source of confusion. Recall that we have defined F from TL(N) onto BL(N).

Definition (orthogonal; orthonormal). A set of two or more sequences in l^2 is said to be orthogonal if for every pair $a,b \in l^2$, $a = \langle a_n \rangle$, $b = \langle b_n \rangle$, $a \neq b$, then

$$(a, b) = \sum_{n=0}^{\infty} a_n b_n^* = 0. \tag{30}$$

The set is said to be orthonormal if it is orthogonal and $||a||^2 = (a, a) = 1$ for every a in the set.

Suppose that we have a countable, complete orthonormal set of sequences in l^2 given by $\{e_n = \langle e_{ni} \rangle\}$. (It is well known that there is at least one such set in l^2.)

For $x \in l^2$, let

$$x = \sum_{n=0}^{\infty} \alpha_n e_n. \tag{31}$$

That is, if $x = \langle x_j \rangle \in l^2$ then, for $j = 0, 1, 2, \ldots,$

$$x_j = \sum_{n=0}^{\infty} \alpha_n e_{nj}, \tag{32}$$

where, for n = 0, 1, 2, ...,

$$\alpha_n = (x, e_n) = \sum_{j=0}^{\infty} x_j e_{nj}^*, \tag{33}$$

since

$$(e_n, e_m) = \sum_{j=0}^{\infty} e_{nj} e_{mj}^* = \delta_{mn}, \tag{34}$$

and

$$\delta_{mn} = 0, \text{ for } m \neq n, \text{ and } \delta_{mm} = 1. \tag{35}$$

We have already exhibited one complete orthonormal set of sequences in l^2 which we can write as

$$\begin{aligned} e_1 &= \langle 1, 0, 0, \ldots \rangle \\ e_2 &= \langle 0, 1, 0, \ldots \rangle \\ &\vdots \end{aligned} \tag{36}$$

Theorem 2.7 The discrete Fourier transform of an orthogonal set of sequences in TL(N) is an orthogonal set of sequences in BL(N).

Proof: If $\{e_n\}$ is any orthogonal set of sequences in TL(N) then

$$\begin{aligned} (Fe_n, Fe_m)_N &= (F\langle e_{nj}\rangle, F\langle e_{mj}\rangle)_N \\ &= N^{-1} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} e_{nk} W^{-jk} \sum_{p=0}^{N-1} e_{mp}^* W^{jp} \\ &= N^{-1} \sum_{k=0}^{N-1} \sum_{p=0}^{N-1} e_{nk} e_{mp}^* \sum_{j=0}^{N-1} W^{j(p-k)} \\ &= \sum_{k=0}^{N-1} e_{nk} e_{mk}^* = (e_n, e_m)_N. \end{aligned}$$

So $\{Fe_n\}$ is an orthogonal set of sequences in BL(N).

Let X denote a discrete stochastic process $\{X_0, X_1, \ldots\}$ which, for convenience, has mean value zero. We will say that such a process is periodic with period N (a positive integer) if, for m = 0, 1, 2, ...,

$$\Pr[X_{m+N} - X_m = 0] = 1. \tag{37}$$

We observe that a periodic (or cyclic) process satisfies a type of stationarity condition. We also observe that every realization of the process is an element of BL(N), with probability one.

For m, n = 0, 1, 2, ..., let

$$R_{mn} = \mathrm{Cov}[X_m, X_n] = E[X_m X_n^*] \tag{38}$$

be the covariance function of X_m and X_n. If, in addition, the X-process is stationary in the wide sense then $R_{mn} = R_{m-n}$.

Lemma 2.8 Let X be a discrete stochastic process with mean value zero. If X is periodic with period N then its covariance function R_{mn}, is periodic with period N in each variable. Proof: if X is periodic with period N and has mean value zero then

$$\begin{aligned} R_{m+N,n} &= \mathrm{Cov}[X_{m+N}, X_n] = R_{mn} \\ &= \mathrm{Cov}[X_m, X_{n+N}]. \end{aligned}$$

We now require the X-process to be stationary in the wide sense and, for k = 0, 1, ..., N-1, let

$$Y_k = \sum_{n=0}^{N-1} X_n W^{nk} \tag{39}$$

and

$$S_k = \sum_{m=0}^{N-1} R_m W^{mk} \tag{40}$$

where $W = \exp(i2\pi/N)$ and R_m is the covariance function of X_{m+n} and X_n. Equation (39) defines another discrete stochastic process Y: $\{Y_0, Y_1, \ldots, Y_{N-1}\}$. For every realization of X which is in BL(N), the realization of Y will be in TL(N). In addition, we have, for $n = 0, 1, 2, \ldots,$

$$X_n = N^{-1} \sum_{k=0}^{N-1} Y_k W^{-nk} . \tag{41}$$

Theorem 2.9 If X is a discrete stochastic process which has mean value zero, is stationary in the wide sense, and is periodic with period N, then the random variables defined in Equation (39) have mean value zero and are uncorrelated. Proof: from Equation (39), we have

$$E[Y_k] = \sum_{n=0}^{N-1} E[X_n] W^{nk} = 0.$$

For $k, j = 0, 1, \ldots, N-1$, and from Equations (38), (39), (40), and Lemma 2.2,

$$\begin{aligned} \mathrm{Cov}[Y_k, Y_j] &= E[Y_k Y_j^*] \\ &= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} E[X_n X_m^*] W^{nk-mj} \\ &= N^{-1} \sum_{p=0}^{N-1} S_p \sum_{n=0}^{N-1} W^{n(k-p)} \sum_{m=0}^{N-1} W^{m(p-j)} \\ &= N S_k \delta_{kj}. \end{aligned}$$

(δ_{kj}: Kronecker delta) since, for $m = 0, 1, \ldots, N-1$,

$$R_m = \mathrm{Cov}[X_{m+n}, X_n] = E[X_m X_0^*] = N^{-1} \sum_{k=0}^{N-1} S_k W^{-mk} . \tag{42}$$

The converse of Theorem 2.9 is also true.

Theorem 2.10 Let Y: $\{Y_0, \ldots, Y_{N-1}\}$ be a finite stochastic process which has mean value zero. If Y_k and Y_j, for all $k \neq j$, are uncorrelated then the discrete stochastic process X defined by Equation (41) has mean value zero, is stationary in the wide sense and is periodic with period N.

Proof: The mean value zero and periodicity follow directly from Equation (41) and the properties of $W = \exp(i2\pi/N)$. In addition, we have

$$\begin{aligned} \mathrm{Cov}[X_m, X_n] &= E[X_m X_n^*] \\ &= N^{-2} \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} E[Y_k Y_j^*] W^{nj-mk} \\ &= N^{-2} \sum_{k=0}^{N-1} E[Y_k Y_k^*] W^{k(n-m)} \\ &= R_{m-n}. \end{aligned} \tag{43}$$

Therefore, X is stationary in the wide sense.

We next show that an expansion similar to Equation (41) can be found for discrete non-periodic stochastic processes.

Suppose we have a finite stochastic process X: $\{X_0, X_1, \ldots, X_{N-1}\}$ which, again for convenience, has mean value zero. We seek another finite stochastic process Y: $\{Y_0, Y_1, \ldots, Y_{N-1}\}$ and a finite set of N-tuples $\{f_n = \langle f_{nk} \rangle\}$ such that

$$X_m = \sum_{k=0}^{N-1} f_{km} Y_k \tag{44}$$

satisfying

$$\sum_{m=0}^{N-1} f_{km} f_{jm}^* = \delta_{kj}, \tag{45}$$

$$E[Y_k Y_j^*] = |\sigma_k|^2 \delta_{kj}. \tag{46}$$

If the Y-process and the N-tuples $\{f_n\}$ can be found then, for $m, n = 0, 1, \ldots, N-1$, it is necessary that

$$\begin{aligned} R_{mn} &= E[X_m X_n^*] \\ &= \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} E[Y_k Y_j^*] f_{km} f_{jn}^* \\ &= \sum_{k=0}^{N-1} |\sigma_k|^2 f_{km} f_{kn}^* , \end{aligned} \tag{47}$$

and for $j = 0, 1, \ldots, N-1$,

$$\begin{aligned} \sum_{n=0}^{N-1} R_{mn} f_{jn} &= \sum_{k=0}^{N-1} |\sigma_k|^2 f_{km} \sum_{n=0}^{N-1} f_{kn}^* f_{jn} \\ &= |\sigma_j|^2 f_{jm} . \end{aligned} \tag{48}$$

Equation (48), written as a matrix equation of the form

$$Rf = \lambda f, \tag{49}$$

is the eigenvalue problem for the NxN matrix $R = R_{mn}$ given in Equation (47), with eigenvalue λ and eigenvector f. We also have

$$Y_k = \sum_{m=0}^{N-1} f_{km}^* X_m. \tag{50}$$

The covariance matrix R, whose elements are R_{mn}, satisfy, for m, n = 0, 1, ..., N-1,

1. $R_{mn} = R^*_{nm}$ (Hermitian)
2. R is non-negative definite
3. the eigenvalues of R are real and non-negative
4. $|R_{mn}|^2 \leq R_{nn}R_{mm}$

If, in addition, the X-process is stationary in the wide sense then $R_{mn} = R_{m-n}$ and R is called a Toeplitz matrix. Properties of Toeplitz matrices are the subject of Chapter 3.

We also observe that if we extend the X-process so that it is periodic with period N, then Equations (41) and (44) are equivalent with $f_{km} = W^{-km}$.

We now see that given the finite stochastic process X, the existence of the finite Y-process and the N-tuples $\{f_n\}$ depends upon the existence of the eigenvalues and eigenvectors of the covariance matrix R. The N-tuples $\{f_n\}$ are then orthogonal (Equation (45)) and the Y-process has uncorrelated random variables (Equation (46)).

The existence of the finite Y-process and the N-tuples $\{f_n\}$ depends upon the existence of the eigenvalues and eigenvectors of the covariance matrix R and not upon the particular X-process. Equation (44) defines a finite Karhunen-Loeve expansion for the X-process and it forms a canonical (orthogonal) expansion for the N-dimensional vector space determined by the random variables $X_0, X_1, \ldots, X_{N-1}$.

When the covariance matrix R is unknown then R can be estimated from realizations of the stochastic process. The eigenvalues and eigenvectors are then determined from the estimated covariance matrix. Numerically, the determination of the eigenvalues and eigenvectors becomes more difficult as the order of R increases. For this reason we attempted to find explicit expressions for the eigenvalues and eigenvectors of Toeplitz matrices. Our search was, however, only partially successful. The results of this search are given in Chapter Three, especially in sections 3.1 and 3.4.

Toeplitz matrices arise in a number of different circumstances in addition to the Karhunen-Loeve expansion mentioned above. For example, Toeplitz matrices arise in the determination of the Wiener filter (Levinson; Wiener). Toeplitz matrices also appear in the discussion of autoregressive time series (Grenander and Szego; Box and Jenkins; Widom).

Because of the importance and present interest in Toeplitz matrices, we feel that a discussion and development of the properties of these matrices should extend beyond the eigenvalue problem needed for the Karhunen-Loeve expansion above. Consequently, in Chapter Three we define several classes of matrices which are related to Toeplitz matrices. Also, since part of the motivation of this thesis is numerical, we have included several numerical algorithms dealing with Toeplitz matrices.

The application of the results of this chapter have been discussed in the Introduction. The shortcomings of these results are discussed in Chapter Four along with possible future work.

Chapter 3. Toeplitz Matrices

3.1 In this chapter we will investigate the properties of Toeplitz matrices. We will derive numerical algorithms for (1) inverting a Toeplitz matrix and (2) solving a Toeplitz system of linear algebraic equations. We will discuss algorithms for (3) finding the eigenvalues and eigenvectors of Toeplitz matrices.

Toeplitz matrices arise primarily as the covariance matrix of a discrete stationary time series (Box and Jenkins; Grenander and Szego).

Levinson derived the first algorithm for solving a Toeplitz system of linear algebraic equations when he constructed an iterative procedure for finding the coefficients of the discrete Wiener filter, the linear filter which minimizes the prediction error of the future values of a time series in the least square sense (Levinson; Wiener). This algorithm, and a modified algorithm which appears in Wiggins and Robinson, require the Toeplitz matrix to be real and symmetric. We derive an algorithm in section 3.3 which avoids these restrictions and uses the Trench-Zohar algorithm discussed next.

In 1964, William F. Trench derived an algorithm for inverting a Toeplitz matrix. Shalov Zohar extended this algorithm slightly and gave a simpler derivation. However, both authors were apparently unaware that the inverse of an invertible Toeplitz matrix is not, in general, a Toeplitz matrix.

We have derived a new class of matrices, called generalized Toeplitz matrices, which are generalizations of Toeplitz matrices and which have

the property that the inverse of an invertible generalized Toeplitz matrix is a generalized Toplitz matrix. With this clarification, we have given Zohar's derivation of the Trench-Zohar algorithm in section 3.2 with only slight modifications.

The Karhunen-Loeve expansion, discussed in section 2.3, requires the determination of the eigenvalues and eigenvectors of Hermitian, positive definite matrices. The Jacobi and Householder methods are two examples of well known numerical algorithms for finding the eigenvalues and eigenvectors of Hermitian matrices. For theoretical reasons, it would be desirable to have explicit expressions for the eigenvalues and eigenvectors of Toeplitz matrices. In section 3.4, we search for these explicit expressions.

Since the eigenvalues and eigenvectors of circulant matrices (special Toeplitz matrices) are known, we have shown in section 3.4 that any Toeplitz matrix can be embedded into circulant matrices of order 2n or 2n-1. Several recent articles by R.M. Grey and Judea Pearl have investigated similar, but different, embeddings of Toeplitz matrices into circulant matrices. They use the circulant matrices to determine the distribution of the eigenvalues of the Toeplitz matrices.

A circulant matrix is the covariance matrix for a periodic time series and its eigenvalues and eigenvectors can be calculated easily using the "fast Fourier transform" algorithm.

Definition (main diagonal; secondary diagonal). The main diagonal of any matrix consists of the elements in the (1, 1), (2, 2), ... positions. The secondary diagonal of a nxn matrix consists of the elements in the (1, n), (2, n-1), ..., (n, 1) positions.

In this chapter we will have occasion to use the following special matrices:

1. 0, zero matrix, is a matrix which has zeros in every position,
2. I, identity matrix, is a nxn matrix which has ones on the main diagonal and zeros elsewhere, and
3. J, a permutation matrix, is a nxn matrix which has ones on the secondary diagonal and zeros elsewhere.

Unless otherwise stated, the matrices in this chapter have complex elements.

In addition, throughout this chapter we will follow the convention that

1. A, B, ..., Z will denote nxn matrices (n is a positive integer; we will use a subscript to denote the order of a matrix when that order is not n),
2. a, b, c, ..., h will denote nx1 matrices (vectors), unless otherwise stated,
3. $(A)_{ij}$ is the element in the ith row and jth column of A.

Definition (Toeplitz matrix). A Toeplitz matrix T of order n is a nxn matrix such that for i, j = 1, 2, ..., n,

$$(T)_{ij} = t_{i-j} \quad . \tag{1}$$

That is, Toeplitz matrices have the same element along each diagonal "parallel" to the main diagonal. Consequently, every nxn Toeplitz matrix is determined once the 2n-1 elements of the first row and first column are specified.

Apparently these matrices were first studied as a special class of matrices by O. Toeplitz about 1910 when he investigated the distribution of the eigenvalues of these matrices as the order n grew without bound (Grenander and Szego). However, we have not found any reference in which the eigenvalues are explicitly determined for general Toeplitz matrices.

Definition (Hermitian). A nxn matrix A is said to be Hermitian if

$$A = A^{*\prime}, \tag{2}$$

where ' denotes matrix transposition and * denotes complex conjugation. That is, A is Hermitian if and only if, for i, j = 1, 2, ..., n,

$$(A)_{ij} = (A)^*_{ji} \quad . \tag{3}$$

We will use the term "symmetric" several times in this chapter and we will give it a broader meaning than is usually the case.

The usual definition of a symmetric matrix is given in equation (2), or (3), when the matrix A has real elements. To avoid confusion with other "symmetries", we will say that A is real-symmetric when Equation (2), or (3) is satisfied and A has real elements.

We will say that the nxn matrix A is symmetric about the main diagonal when A = A' (the elements of A may be complex).

Consequently, a Toeplitz matrix T of order n is Hermitian if, for $k = 0, 1, 2, \ldots, n-1$,

$$t_{-k} = t_k^* \ , \tag{4}$$

where t_k is an element of the first column of T and t_{-k} is an element of the first row.

Definition (positive definite). A nxn Hermitian matrix A is said to be positive definite if for all nx1 non-zero matrices x, the quadratic form x'Ax* is positive. That is, A is positive definite if and only if

$$\sum_{i=1}^{n} \sum_{j=1}^{n} (A)_{ij} x_i x_j^* \geq 0 \ , \tag{5}$$

where x_i are the elements of $x \neq 0$.

Definition (eigenvalue; eigenvector). If for a nxn matrix A and a scalar λ, there is a non-zero vector (nx1 matrix) x such that

$$Ax = \lambda x \tag{6}$$

then λ is an eigenvalue of A and x is an eigenvector of A associated with λ.

The eigenvalue problem for a matrix A is the problem of finding the eigenvalues and eigenvectors of A.

The eigenvalues of a nxn matrix A consists of the solutions (roots) of the characteristic polynomial: $\det(A - \lambda I) = 0$, where det(A) means the determinant of A.

Lemma 3.1 (Vinograde) Every nxn matrix A has (at most) n eigenvalues. The eigenvalues of Hermitian matrices are real numbers.

Lemma 3.2 (Vinograde) If A is a nxn Hermitian matrix then A is positive definite if and only if all eigenvalues of A are postive.

Definition (principal minor). If A_k is a kxk submatrix obtained from the nxn matrix A by deleting n-k rows and the same n-k numbered columns, $0 \leq k < n$, then $\det(A_k)$ is called a principal minor of A of order k. The descending principal minors are those numbers, $\det(A_k)$, $k = 1, 2, \ldots, n$, for which

$$(A_k)_{ij} = (A)_{ij} \tag{7}$$

for $i, j = 1, 2, \ldots, k$.

Lemma 3.3 (Vinograde) If A is a nxn real-symmetric matrix then A is positive definite if and only if all principal minors of A are positive.

Lemma 3.4 If A, I, and J are nxn matrices then, for $i, j = 1, 2, \ldots, n$,

1. $(JA)_{ij} = (A)_{n+1-i,j}$ (8)

(the rows of JA are the rows of A in reverse order)

2. $(AJ)_{ij} = (A)_{i,\, n+1-j}$ (9)

(the columns of AJ are the columns of A in reverse order)

3. $J^2 = I$ and $J' = J$.

Proof: the results follow directly from the definition of J.

Definition (persymmetric; PS). A nxn matrix A is said to be persymmetric, abbreviated to PS, if $(A)_{ij}$ is a function of i+j only, for $i, j = 1, 2, \ldots, n$ (Aitken).

We observe that a matrix is PS if and only if all elements along diagonals "parallel" to the secondary diagonal are equal. Hence a PS matrix is determined once the 2n-1 elements of the first row and last column are specified. For example, J is a PS matrix.

We mentioned in Chapter 1 that the set of all nxn matrices with complex elements form a complex vector space. It can be shown that such a space has dimension n^2. In the next theorem we will state and prove some of the properties of PS matrices which form a 2n-1 dimensional vector space.

Theorem 3.5 Let A and J be nxn matrices.

1. Sums and scalar multiples of PS matrices are PS matrices.
2. If A is a PS matrix then A = A'.
3. A is a PS matrix if and only if JA, or AJ, is a Toeplitz matrix.

Proof:

1. If A and B are nxn PS matrices then for i, j = 1, 2, ..., n, $(A+B)_{ij} = (A)_{ij} + (B)_{ij}$ and $(\alpha A)_{ij} = \alpha(A)_{ij}$ (sums and scalars multiples of functions of i+j are again functions of i+j).
2. If A is a PS matrix and $(A)_{ij} = a_{i+j}$ then $(A)_{ij} = a_{i+j} = a_{j+i} = (A)_{ji}$.
3. If A is a PS matrices and $(A)_{ij} = a_{i+j}$ then by Lemma 3.4, $(JA)_{ij} = (A)_{n+1-i,\, j} = a_{n+1+j-i}$ (a function of i-j, for i, j = 1, 2, ..., n).

We also have $AJ = A'J' = (JA)'$.

If JA, or AJ, is a Toeplitz matrix and

$(AJ)_{ij} = b_{i-j}$ then $A = AJJ$ and

$(A)_{ij} = (AJJ)_{ij} = (AJ)_{i,\, n+1-j} = b_{i+j-n-1}$

(a function of $i+j$ for $i, j = 1, 2, \ldots, n$).

Definition (centrosymmetric; CS). A nxn matrix A is said to be centrosymmetric, abbreviated to CS, if for $i, j = 1, 2, \ldots, n$,

$$(A)_{ij} = (A)_{n+1-i,\, n+1-j}. \tag{10}$$

We observe that a matrix A is a CS matrix if and only if A is symmetric about the "center" of the matrix. For example, I and J are CS matrices.

In the next theorem we will state and prove some of the properties of CS matrices. The set of CS matrices of order n form a vector space whose dimension is $n^2/2$, when n is an even integer, and $(n^2+1)/2$, when n is an odd integer.

Theorem 3.6 Let A and J be nxn matrices.

1. A is a CS matrix if and only if
$$JAJ = A \tag{11}$$
2. Sums, scalar multiples, products and transpositions of CS matrices are CS matrices.
3. If A is a CS matrix and A^{-1} exists then A^{-1} is a CS matrix.

Proof:

1. This follows from Equations (8), (9), and (10).

2. If A and B are CS matrices then

   $J(A+B)J = JAJ + JBJ = A + B$,

   $J(\alpha A)J = \alpha JAJ = \alpha A$,

   $JABJ = JAJJBJ = AB$, and

   $JA'J = J'A'J' = (JAJ)' = A'$.

3. From $I = AA^{-1} = JAJJA^{-1}J = A(JA^{-1}J)$ we obtain $A^{-1} = JA^{-1}J$ since inverses of matrices are unique when they exist.

Next we will define a class of matrices which apparently has not been studied before as a special class of matrices.

Definition (generalized Toeplitz; GT). A nxn matrix T is said to be a generalized Toeplitz matrix, abbreviated to GT, if

$$JTJ = T' \tag{12}$$

As we will show in the next theorem, all Toeplitz matrices are GT matrices. The converse, however, is not true. The following example of a GT matrix is not a Toeplitz matrix:

$$\begin{bmatrix} 1 & 4 & 6 \\ 2 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}.$$

In recent literature dealing with algorithms involving Toeplitz matrices, the definition of a Toeplitz matrix is the same as we have given above. However the property of Toeplitz matrices which is most frequently used is given in Equation (12). It is interesting that up to the present time no one has recognized that Equation (12) represents a generalization of Toeplitz matrices.

In the next theorem we state and prove some of the properties of GT matrices. As a consequence the set of all nxn GT matrices forms a vector space of dimension n(n+1)/2.

We observe that GT matrices are symmetric about the secondary diagonal.

Theorem 3.7 Let I, J, and T be nxn matrices.

1. If T is a Toeplitz matrix then JTJ = T'.
2. If T is a GT matrix and T^{-1} exists then T^{-1} is a GT matrix (inverses of Toeplitz matrices are not necessarily Toeplitz matrices).
3. Sums, scalar multiples and transpositions of GT matrices are GT matrices.
4. Products of GT matrices are GT matrices if and only if they commute.
5. T is a Toeplitz matrix if and only if JT, or TJ, is a PS matrix.

Proof:

1. If $(T)_{ij} = t_{i-j}$, for i, j = 1, 2, ..., n then $(JTJ)_{ij} = (T)_{n+1-i,\, n+1-j} = t_{j-i} = (T')_{ij}$.
2. If T is a GT matrix and T^{-1} exists then $I = TT^{-1} = JTJJT^{-1}J = T'(JT^{-1}J)$ and $JT^{-1}J = (T')^{-1} = (T^{-1})'$.
3. If T and S are nxn GT matrices then $J(T+S)J = JTJ + JSJ = T' + S' = (T + S)'$ and $JT'J = J'T'J' = (JTJ)' = T'' = T$.

4. If T and S are nxn GT matrices then

   JTSJ = JTJJSJ = T'S' = (ST)'. Hence JTSJ = (ST)' = (TS)' if and only if ST = TS.

5. This follows from Theorem 3.5.

Theorem 3.8 A nxn GT matrix T is CS if and only if T = T'.

Proof: This follows directly from Equations (11) and (12).

We will illustrate the four types of matrices defined above by considering the decomposition of a matrix of each type into matrices of lower order. These decompositions are of value, when they exist, in that they reduce the eigenvalue problem of a large matrix to eigenvalue problems of smaller matrices.

If S_{2n} is a 2nx2n CS matrix and

$$J_{2n} = \begin{bmatrix} 0 & J \\ J & 0 \end{bmatrix} \tag{13}$$

then, from Equation (11), we have

$$S_{2n} = \begin{bmatrix} A & B \\ JBJ & JAJ \end{bmatrix}, \tag{14}$$

where A, B, and J are nxn matrices but A and B are not necessarily CS matrices. This means that the CS property is not preserved under this decomposition.

If S_{2n} is a 2nx2n PS matrix then

$$S_{2n} = \begin{bmatrix} A & B \\ B & C \end{bmatrix} \tag{15}$$

where A, B, and C are nxn PS matrices. The PS property is preserved under this decomposition.

If T_{2n} is a 2nx2n GT matrix then from Equation (12), we have

$$T_{2n} = \begin{bmatrix} A & B \\ C & JA'J \end{bmatrix}, \tag{16}$$

where B, C are GT matrices but A is not necessarily a GT matrix.

If T_{2n} is a 2nx2n Toeplitz matrix and for i, j = 1, 2, ..., 2n,

$$(T_{2n})_{ij} = t_{i-j} \tag{17}$$

then let

$$T_{2n} = \begin{bmatrix} T & U \\ V & T \end{bmatrix} \tag{18}$$

where T, U, and V are nxn Toeplitz matrices and for i, j = 1, 2, ..., n,

$$(T)_{ij} = t_{i-j}\ ,\ (V)_{ij} = t_{i-j+n}\ ,\ \text{and}\ (U)_{ij} = t_{i-j-n}\ . \tag{19}$$

Our interest in this decomposition stems from the fact that the Toeplitz property is preserved under this decomposition. We would like to use this decomposition to express the determinant of T_{2n}, $\det(T_{2n})$, in terms of the determinants of matrices of order n.

Lemma 3.9 If A and B are nxn matrices then det(AB) = det(A)det(B).

Lemma 3.10 If I and J are nxn matrices then det (I) = 1 and det (J) = $(-1)^{n(n-1)/2}$.

Proof: clearly det(I) = 1 for all n.

The number of row (or column) interchanges necessary to transform J into I, is given by

$$\sum_{i=1}^{n-1} (n-i) = n(n-1)/2.$$

If S_{2n} is the 2nx2n CS matrix given in Equation (14) then Aitken considers the following matrix product:

$$\begin{bmatrix} I & J \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ JBJ & JAJ \end{bmatrix} \begin{bmatrix} I & -J \\ 0 & I \end{bmatrix} = \begin{bmatrix} A+BJ & 0 \\ JBJ & JAJ-JB \end{bmatrix}. \tag{20}$$

If we apply Lemmas 3.9, 3.10 and Laplace's expansion on the first n columns, we obtain

$$\det \begin{bmatrix} I & J \\ 0 & I \end{bmatrix} = 1 = \det \begin{bmatrix} I & -J \\ 0 & I \end{bmatrix}$$

and consequently the above matrix product gives

$$\det(S_{2n}) = \det(A+BJ)\det(JAJ-JB). \tag{21}$$

Mullen applied this technique to Toeplitz matrices by requiring $T'_{2n} = T_{2n}$ (so that T_{2n} is a CS matrix). If T_{2n} is a 2nx2n Toeplitz matrix then from Equation (18) we obtain

$$T_{2n} = \begin{bmatrix} T & U \\ V & T \end{bmatrix}.$$

If we require $T'_{2n} = T_{2n}$ then we have $T = T' = JTJ$, $U = V' = JVJ$, and

$$\begin{bmatrix} I & J \\ 0 & I \end{bmatrix} \begin{bmatrix} T & V' \\ V & T \end{bmatrix} \begin{bmatrix} I & -J \\ 0 & I \end{bmatrix} = \begin{bmatrix} T+JV & 0 \\ V & T-VJ \end{bmatrix}.$$

By taking determinants of these matrices we obtain

$$\det(T_{2n}) = \det(T+JV)\det(T-VJ). \tag{22}$$

There is a similar decompositon for Toeplitz matrices of order 2n+1.

However, while T, U, and V are Toeplitz matrices, JV and VJ are not Toeplitz (they are PS matrices by Theorem 3.7) and, therefore, this decomposition cannot be continued.

Let us instead consider the following matrix product:

$$\begin{bmatrix} I & I \\ 0 & I \end{bmatrix} \begin{bmatrix} T & U \\ V & T \end{bmatrix} \begin{bmatrix} I & -I \\ 0 & I \end{bmatrix} = \begin{bmatrix} T+V & U-V \\ V & T-V \end{bmatrix}, \tag{23}$$

especially the case where U = V. An interesting special case occurs when T_{2n} is a circulant matrix.

Definition (circulant). A nxn matrix T is said to be circulant if T is a Toeplitz matrix and for k = 1, 2, ..., n-1,

$$t_k = t_{k-n} \tag{24}$$

(a circulant matrix is determined once the n elements of the first row are specified).

Consequently, if T_{2n} is a circulant matrix then it satisfies Equations (17), (18), (19), $t_k = t_{k-2n}$ and U = V. Therefore the determinant of Equation (23) gives

$$\det(T_{2n}) = \det(T+V)\det(T-V). \tag{25}$$

Since T, U = V, T+V, T-V are Toeplitz matrices, we can continue this decomposition. There is, however, no need to continue as the determinant of any circulant matrix is known (Aitken).

When T_{2n+1} is a Toeplitz matrix of order 2n+1, the decomposition of T_{2n+1} is given by

$$T_{2n+1} = \begin{bmatrix} T & Ja & U \\ b'J & t_0 & a' \\ V & b & T \end{bmatrix} \tag{26}$$

where, for i, j = 1, 2, ..., n, $(T)_{ij} = t_{i-j}$, $(U)_{ij} = t_{i-j-n-1}$, $(V)_{ij} = t_{i-j+n+1}$, $a = [t_{-1}\ t_{-2}\ \ldots\ t_{-n}]'$, and $b = [t_1\ t_2\ \ldots\ t_n]'$.

For a matrix product similar to the one given in Equation (23), we obtain

$$\begin{bmatrix} T+V & Ja+b & U-V \\ b'J & t_0 & a'-b'J \\ V & b & T-V \end{bmatrix}. \tag{27}$$

The matrix product given in Equation (23) was suggested by some results due to Levinson.

The matrix products in Equations (20) and (23) are particular examples of similarity (congruent) transformations. It is known (Vinograde) that if S_{2n} or T_{2n} are Hermitian then they are similar to diagonal matrices. Equations (20) and (23) are attempts to find relatively simple unitary matrices which transform S_{2n} or T_{2n} into diagonal matrices under a similarity transformation.

Lastly, if we assume that T^{-1} exists and T_{2n} is given as in Equation (18), then pre-multiplication of T_{2n} by

$$\begin{bmatrix} I & 0 \\ -VT^{-1} & I \end{bmatrix} \text{ and } \begin{bmatrix} I & -UT^{-1} \\ 0 & I \end{bmatrix}$$

gives

$$\begin{aligned} \det(T_{2n}) &= \det(T)\det(T-VT^{-1}U) \\ &= \det(T)\det(T-UT^{-1}V) \end{aligned} \tag{28}$$

and

$$\det(T-VT^{-1}U) = \det(T-UT^{-1}V).$$

We will return to these results in Section 3.4 in connection with the eigenvalue problem for Toeplitz matrices.

3.2 We next derive a recursive algorithm for computing the inverse of a Toeplitz matrix. This derivation is a slight modification of an algorithm given by Zohar which, in turn, was a modification of the original derivation given by Trench.

We will derive an explicit expression for T_{n+1}^{-1} assuming that T_k^{-1} exists for k = 1, 2, ..., n+1. This condition is called "strongly nonsingular" by Zohar and is more general than the condition of positive definiteness of T_{n+1} required by Trench.

Let us assume that the complex numbers $t_0, t_1, t_2, \ldots, t_n, t_{-1}, t_{-2}, \ldots, t_{-n}$ are given and, for i, j = 1, 2, ..., n,

$$(T)_{ij} = t_{i-j} , \tag{29}$$

$$a = [t_1 \; t_2 \; \ldots \; t_n]' \tag{30}$$

$$b = [t_{-1} \; t_{-2} \; \ldots \; t_{-n}]' \tag{31}$$

where ' denotes matrix transposition. Equation (29) specifies that T is a Toeplitz matrix of order n.

If T_{n+1} is a nonsingular Toeplitz matrix of order n+1 then

$$T_{n+1}^{-1} = \frac{\text{adj}(T_{n+1})}{\det(T_{n+1})}$$

where $\text{adj}(T_{n+1})$ and $\det(T_{n+1})$ are the adjoint and determinant of T_{n+1}, respectively. If the complex number in the (1, 1) position of the matrix T_{n+1}^{-1} is denoted by $1/\lambda_n$ then

$$\det(T_{n+1}) = \lambda_n \det(T). \tag{32}$$

We can express T_{n+1} and T_{n+1}^{-1} as

$$T_{n+1} = \begin{bmatrix} t_0 & b' \\ a & T \end{bmatrix} \tag{33}$$

and

$$T_{n+1}^{-1} = \frac{1}{\lambda_n} \begin{bmatrix} 1 & d' \\ c & M \end{bmatrix} \tag{34}$$

where t_0, 1, and λ_n are scalars, a, b, c, d are nx1 matrices, and T, M are nxn matrices.

If I, I_{n+1} denote the identity matrices of order n and n+1, respectively, then $T_{n+1}\ T_{n+1}^{-1} = I_{n+1} = T_{n+1}^{-1}\ T_{n+1}$,

$$\lambda_n\ I_{n+1} = \begin{bmatrix} t_0+b'c & t_0d'+ b'M \\ a+Tc & ad'+TM \end{bmatrix},$$

and

$$\lambda_n\ I_{n+1} = \begin{bmatrix} t_0+d'a & b'+d'T \\ t_0c+Ma & cb'+MT \end{bmatrix}.$$

We now equate these two expressions for $\lambda_n\ I_{n+1}$ and obtain the following relations:

$$\lambda_n = t_0+b'c = t_0+d'a, \tag{35}$$

$$0 = a+Tc = t_0c+Ma, \tag{36}$$

$$0 = t_0d'+b'M = b'+d'T, \tag{37}$$

and

$$\lambda_n\ I = ad'+TM = cb'+MT. \tag{38}$$

From these four relations

$$b'c = d'a, \tag{39}$$

$$0 = c+T^{-1}a, \tag{40}$$

$$0 = b'T^{-1}+d', \tag{41}$$

$$\lambda_n T^{-1} = T^{-1}ad'+M = cb'T^{-1}+M. \tag{42}$$

From these four relations we obtain

$$b'c = d'a = -b'T^{-1}a \tag{43}$$

$$c = -T^{-1}a \tag{44}$$

$$d = -T^{-1}'b \tag{45}$$

and

$$M = \lambda_n T^{-1}+cd' = \lambda_n T^{-1}+T^{-1}ab'T^{-1}. \tag{46}$$

These equations are satisfied for matrices of order n = 1, 2, ...

If we substitute these last four relations into Equation (34), we can express the inverse of T_{n+1} as

$$T_{n+1}^{-1} = \frac{1}{\lambda_n}\begin{bmatrix} 1 & d' \\ c & \lambda_n T^{-1}+cd' \end{bmatrix}. \tag{47}$$

Thus we are able to express T_{n+1}^{-1} explicitly in terms of the elements of T^{-1} and T_{n+1}, and ultimately in terms of T_{n+1}.

Except for Equations (29), (30), and (31), we have not made use of the Toeplitz property. With the explicit use of this property we will be able to compute T_{n+1}^{-1} without the explicit computation of the intermediate inverses. This is acheived by developing a recursive algorithm for the computation of λ_k, c_k, d_k, for k = 1, 2, ..., n.

If we require T_{n+1} to be a Toeplitz matrix then T is a Toeplitz matrix and T^{-1} and T_{n+1}^{-1} are generalized Toeplitz matrices by Theorem 3.7. Applying Equation (12) to Equation (47) we also obtain

$$\lambda_n T_{n+1}^{-1} = \begin{bmatrix} \lambda_n T^{-1}+Jdc'J & Jd \\ c'J & 1 \end{bmatrix}. \tag{48}$$

For each i, j = 1, 2, ..., n, we obtain from Equations (47) and (48),

$$\lambda_n(T_{n+1}^{-1})_{ij} = \lambda_n(T^{-1})_{ij} + (Jdc'J)_{ij} \tag{49}$$

and

$$\lambda_n(T_{n+1}^{-1})_{i+1,j+1} = \lambda_n(T^{-1})_{ij} + (cd')_{ij}. \tag{50}$$

If we eliminate the common term on the right hand side of these two equations, we get

$$\lambda_n(T_{n+1}^{-1})_{i+1,j+1} = \lambda_n(T_{n+1}^{-1})_{ij} + (cd'-Jdc'J)_{ij}. \tag{51}$$

This equation illustrates how the entire matrix T_{n+1}^{-1} can be found once the first row and first column are known. The first row and first column are determined by λ_n, c_n, d_n.

We will now show how λ_n, c_n, d_n can be found by recursive generation.

1. From Equation (40) we have, for all n,

$$c = -T^{-1}a = -JT^{-1'}Ja$$

where we have used the generalized Toeplitz property, Equation (12). Consequently, for order n+1, we get

$$c_{n+1} = -\frac{1}{\lambda_n}\begin{bmatrix} \lambda_n T^{-1}+Jdc'J & Jd \\ c'J & 1 \end{bmatrix}\begin{bmatrix} a \\ t_{n+1} \end{bmatrix}.$$

By performing the indicated matrix multiplication we can express this product as

$$c_{n+1} = \begin{bmatrix} c \\ 0 \end{bmatrix} - \frac{c'Ja+t_{n+1}}{\lambda_n}\begin{bmatrix} Jd \\ 1 \end{bmatrix}. \tag{52}$$

2. From Equation (41) we have, for all n, $d = -T^{-1'}b = -JT^{-1}Jb$

where, again, we have used the generalized Toeplitz property.

Consequently, for order n+1, we get

$$d_{n+1} = -\frac{1}{\lambda_n}\begin{bmatrix} \lambda_n T^{-1}+Jdc'J & Jd \\ c'J & 1 \end{bmatrix}' \begin{bmatrix} b \\ t_{-n-1} \end{bmatrix}.$$

By performing the indicated matrix multiplication we can express this product as

$$d_{n+1} = \begin{bmatrix} d \\ 0 \end{bmatrix} - \frac{d'Jb+t_{-n-1}}{\lambda_n}\begin{bmatrix} Jc \\ 1 \end{bmatrix}. \tag{53}$$

3. From Equation (35), we obtain for order n+1,

$$\lambda_{n+1} = t_0+b'_{n+1}c_{n+1} = t_0+d'_{n+1}a_{n+1}. \tag{54}$$

Alternatively, from Equation (50) we obtain

$$\lambda_n(T^{-1}_{n+1})_{n+1,n+1} = \lambda_n(T^{-1})_{nn} + (cd')_{nn} = 1$$

and, replacing n by n+1 in the middle expression,

$$\lambda_{n+1}(T^{-1}_{n+1})_{n+1,n+1} + (c_{n+1}d'_{n+1})_{n+1,n+1} = 1.$$

If we eliminate $(T^{-1}_{n+1})_{n+1,n+1}$ from these two equations, we then have

$$\lambda_{n+1} = \lambda_n[1 - (c_{n+1}d'_{n+1})_{n+1,n+1}] \tag{55}$$

$$= \lambda_n - (c'Ja+t_{n+1})(d'Jb+t_{-n-1})/\lambda_n,$$

where $\lambda_n \neq 0$ for n = 1, 2, ..., which was assumed in the beginning. We can use Equations (52) and (53) to show that Equations (54) and (55) are equivalent.

4. Finally, we need to specify the initial values for the recursive generation of λ_n, c_n, d_n. We have

$$\lambda_1 = \det(T_2)/\det(T_1) = (t_0^2-t_1t_{-1})/t_0 \tag{56}$$

$$c_1 = -t_1/t_0, \tag{57}$$

and

$$d_1 = -t_{-1}/t_0 \ . \tag{58}$$

Trench-Zohar Algorithm:

Given the complex numbers t_k, for $k = 0, 1, \ldots, n, -1, -2, \ldots, -n$, we

1. compute λ_1, c_1, d_1 by Equations (56), (57), and (58),
2. compute λ_k, c_k, d_k, for each $k = 2, 3, \ldots, n+1$ using a_n and b_n (Equations (30) and (31)) by Equations (52), (53), and (54),
3. compute T_{n+1}^{-1} by Equations (47) and (51) where

$$\lambda_n (T_{n+1}^{-1})_{11} = 1$$

$$\lambda_n (T_{n+1}^{-1})_{1,j+1} = (d_n)_j, \quad j = 1, \ldots, n,$$

$$\lambda_n (T_{n+1}^{-1})_{i+1,1} = (c_n)_i, \quad i = 1, \ldots, n,$$

$$\lambda_n (T_{n+1}^{-1})_{i+1,j+1} = \lambda_n (T_{n+1}^{-1})_{i,j} + (c_n d_n' - J d_n c_n' J)_{ij}$$

for $i, j = 1, 2, \ldots, n$. We note that the presence of J does not actually require a matrix multiplication but rather a "reshuffling".

We have programmed and tested this algorithm for relatively simple examples. Akaike has extended this algorithm to block Toeplitz matrices.

When the matrix T_{n+1} is Hermitian then $b^* = a$ and $d^* = c$. If T_{n+1} is real-symmetric and positive definite then $\lambda_n > 0$.

3.3 We will now derive an algorithm for solving a system of linear algebraic equations whose coefficient matrix is a nxn Toeplitz matrix. That is, we would like to solve the matrix equation

$$Tx = y \tag{59}$$

where T is a given Toeplitz matrix of order n, x is a unknown nx1 matrix and y is a given nx1 matrix.

Chronologically, the first algorithm for this system was derived by Levinson in 1947. A modification of Levinson's algorithm appears in Wiggins and Robinson. This latter algorithm also appears in the Geophysics Department (Colorado School of Mines) computer library under the name of TOEPLZ. However, both of these algorithms require the Toeplitz matrix to be real-symmetric.

We now derive an algorithm based upon the derivation given in the previous section which avoids the real-symmetric restriction.

We are going to derive a recursive algorithm for finding the solution x_{n+1} of the system

$$T_{n+1}x_{n+1} = y_{n+1} \tag{60}$$

assuming that we know T_{n+1}, y_{n+1} and the solution x of the system given in Equation (59). Let

$$x_{n+1} = [z' \ \alpha]', \ y_{n+1} = [y' \ \beta]' \tag{61}$$

where α,β are scalars and y, z are nx1 matrices.

Since T_{n+1} is a Toeplitz matrix of order n+1, we can apply the generalized Toeplitz property, Equation (12), to T_{n+1} in Equation (33)

and express Equation (60) as

$$\begin{bmatrix} T & Jb \\ a'J & t_0 \end{bmatrix} \begin{bmatrix} z \\ \alpha \end{bmatrix} = \begin{bmatrix} y \\ \beta \end{bmatrix}. \tag{62}$$

From Equation (59), we have $y = Tx$ and substituting this relation into Equation (62) we obtain

$$\begin{bmatrix} T & Jb \\ a'J & t_0 \end{bmatrix} \begin{bmatrix} z-x \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \beta - a'Jx \end{bmatrix}. \tag{63}$$

If we assume that T_{n+1}^{-1} exists ($\lambda_n \neq 0$) then from Equation (48), we get

$$\lambda_n \begin{bmatrix} z-x \\ \alpha \end{bmatrix} = \begin{bmatrix} \lambda_n T^{-1} + Jdc'J & Jd \\ c'J & 1 \end{bmatrix} \begin{bmatrix} 0 \\ \beta - a'Jx \end{bmatrix} = (\beta - a'Jx) \begin{bmatrix} Jd \\ 1 \end{bmatrix}$$

and, therefore,

$$x_{n+1} = \begin{bmatrix} x \\ 0 \end{bmatrix} - \frac{xJa' - \beta}{\lambda_n} \begin{bmatrix} Jd \\ 1 \end{bmatrix}. \tag{64}$$

We note the similarity of Equations (52) and (64). We also note that although the nx1 matrix c does not explicitly appear in Equation (64), it is needed in the calculation of the nx1 matrix d.

Algorithm:

Given the complex numbers t_k, $k = 0, 1, \ldots, n, -1, -2, \ldots, -n$ and $y_{n+1} = [\beta_1 \ \beta_2 \ \cdots \ \beta_n \ \beta_{n+1}]$, we

1. compute λ_1, c_1, d_1 by Equations (56), (57), and (58),
2. compute $x_1 = \beta_1 / t_0$,
3. compute λ_k, c_k, d_k, for each $k = 2, 3, \ldots, n$ by Equations (52), (53), (54) using a_n and b_n (Equations (30) and (31)),
4. compute x_k, for $k = 2, 3, \ldots, n+1$ by Equation (64).

3.4 In this last section we are interested in the eigenvalue problem for Toeplitz matrices. It was our original intention to derive, if possible, explicit expressions for the eigenvalues and eigenvectors of Toeplitz matrices. We have been only partially successful.

If T is a nxn Toeplitz matrix, λ is a scalar, and x is a non-zero nx1 matrix such that

$$Tx = \lambda x \tag{65}$$

then λ is an eigenvalue of T and x is an eigenvector of T associated with λ. The eigenvalues of T are solutions of the characteristic equation

$$\det(T - \lambda I) = 0. \tag{66}$$

If T is Hermitian then all eigenvalues are real numbers. If T is positive definite and real-symmetric then all eigenvalues are non-negative real numbers.

If T is a Toeplitz matrix then $T - \lambda I$ is also a Toeplitz matrix for every λ. Hence, the problem of finding eigenvalues of Toeplitz matrix is equivalent to the problem of finding the determinant of a Toeplitz matrix.

Most numerical algorithms which compute the eigenvalues of matrices, do not depend upon explicit expressions for eigenvalues.

For example, the Jacobi method is an algorithm which computes the eigenvalues and eigenvectors of real-symmetric matrices using a sequence of orthogonal transformations. The Householder method is similar to the Jacobi method but it uses unitary transformations instead.

There are numerical algorithms which determine the eigenvalues of quite general matrices. Most of these numerical algorithms are iterative and the convergence is only guaranteed when the matrix is positive definite. The above mentioned algorithms are well known and therefore, they are not reproduced in this thesis.

The Karhunen-Loeve expansion derived in section 2.4, Equation (44), requires the determination of the eigenvalues and eigenvectors of Hermitian, Toeplitz matrices, and as indicated above, there are numerical algorithms which will handle these matrices. However, round-off errors increase as the order of the matrices increases. Consequently, explicit expressions for the eigenvalues and eigenvectors of Toeplitz matrices would be desirable.

In fact explicit expressions for the eigenvalues and eigenvectors of circulant matrices do exist. (see sections 2.2 and 3.1)

$C = C_n$ is a nxn circulant matrix whose elements are given by

$$(C)_{ij} = \begin{cases} c_{j-i} & , i \leq j \\ c_{n-i+j} & , i > j. \end{cases} \tag{67}$$

The eigenvalues, λ, and eigenvectors, x, of C are solutions to the equations

$$Cx = \lambda x. \tag{68}$$

That is, the eigenvalues, λ, and eigenvectors, x, are solutions to the n difference equations

$$\sum_{j=1}^{m-1} c_{n-m+j}x_j + \sum_{j=m}^{n} c_{j-m}x_j = \lambda x_m, \tag{69}$$

for m = 1, 2, ..., n, which are linear and have constant coefficients. If we change the indices of summation then we get

$$\sum_{j=0}^{n-m} c_j x_{j+m} + \sum_{j=n-m+1}^{n-1} c_j x_{j-n+m} = \lambda x_m. \tag{70}$$

Since the equations are linear and have constant coefficients we can assume that $x_j = r^j$ for some constant r. Substituting $x_j = r^j$, $r \neq 0$, into Equation (70) and dividing by r^m, we obtain

$$\lambda = \sum_{j=0}^{n-m} c_j r^j + r^{-n} \sum_{j=n-m+1}^{n-1} c_j r^j. \tag{71}$$

If we select $r = W^k = \exp(i2\pi k/n)$ then $r^{-n} = 1$ and, for k = 0, 1, 2, ..., n,

$$\lambda_k = \sum_{j=0}^{n-1} c_j W^{jk} \tag{72}$$

and

$$x_j^{(k)} = W^{jk}. \tag{73}$$

Therefore, for circulant matrices the eigenvalues, Equation (72), and the eigenvectors, Equation (73), are given explicitly. Moreover, the eigenvalues are the discrete Fourier transform of the elements of the first row of C and the eigenvectors are the same for all nxn circulant matrices. The set of n eigenvectors form an orthogonal set of n-tuples.

Since the eigenvalues and eigenvectors of circulant matrices are known we might try embedding a nxn Toeplitz matrix in a more tractable circulant matrix. There are two recent articles (Gray; Pearl) which take this approach and show that the circulant matrix is "asymptotically equivalent" to the Toeplitz matrix. Consequently, the distribution of eigenvalues of the Toeplitz matrix can be inferred from the distribution of the eigenvalues of the circulant matrix, as the order of the matrices increases without bound. Grenander and Szego give the definitive treatment of the distribution of the eigenvalues of Toeplitz matrices.

There are several ways to embed a nxn Toeplitz matrix in a circulant matrix. We will only indicate two ways. Gray and Pearl give similar embeddings.

For any nxn Toeplitz matrix T, we determine a nxn Toeplitz matrix U such that

$$C_{2n} = \begin{bmatrix} T & U \\ U & T \end{bmatrix} \tag{74}$$

is a 2nx2n circulant matrix. That is, for i, j = 1, 2, ..., n, let

$$(T)_{ij} = t_{i-j} \tag{75}$$

then

$$(U)_{ij} = \begin{cases} t_{i-j+n} & , \ i \leq j \\ t_{i-j-n} & , \ i > j. \end{cases} \tag{76}$$

Since T has at most 2n-1 distinct elements, U is completely determined except for the arbitrary element t_n.

C_{2n} has 2n eigenvalues, λ_k, and 2n eigenvectors, $z^{(k)}$, where, for k = 1, 2, ..., 2n,

$$C_{2n}z^{(k)} = \lambda_k z^{(k)}, \tag{77}$$

$$\lambda_k = \sum_{j=-n+1}^{n} t_j W_{2n}^{-jk} = \sum_{j=0}^{n-1} (t_{-j}+(-1)^k t_{n-j}) W_{2n}^{jk}, \tag{78}$$

and

$$z^{(k)} = \begin{bmatrix} x^{(k)} \\ y^{(k)} \end{bmatrix} \tag{79}$$

where $W_{2n}^k = \exp(i2\pi k/2n)$, $y^{(k)} = (-1)^k x^{(k)}$, and

$$x^{(k)} = [1, W_{2n}^k, \ldots, W_{2n}^{k(n-1)}]'. \tag{80}$$

These results follow from Equation (68), (72) and (73). Equation (77),

$$\begin{bmatrix} T & U \\ U & T \end{bmatrix} \begin{bmatrix} x^{(k)} \\ y^{(k)} \end{bmatrix} = \lambda_k \begin{bmatrix} x^{(k)} \\ y^{(k)} \end{bmatrix},$$

leads to

$$\begin{aligned} Tx^{(k)} + Uy^{(k)} &= \lambda_k x^{(k)} \\ Ux^{(k)} + Ty^{(k)} &= \lambda_k y^{(k)}. \end{aligned} \tag{81}$$

We see that λ_k and $x^{(k)}$ are the eigenvalues and eigenvectors of the nxn Toeplitz matrix T+U for k = 2, 4, 6, ..., 2n since $y^{(k)} = (-1)^k x^{(k)}$ and

$$(T+U)x^{(k)} = \lambda_k x^{(k)}. \tag{82}$$

λ_k and $x^{(k)}$ are the eigenvalues and eigenvectors of the nxn Toeplitz matrix T-U for k = 1, 3, 5, ..., 2n-1 and

$$(T-U)x^{(k)} = \lambda_k x^{(k)}. \tag{83}$$

In addition, T+U is actually a circulant matrix and T-U is a "skew-circulant" matrix. These results were anticipated in section 3.1 by Equation (25).

The results contained in Equations (81), (82), and (83) at first looked quite promising. However, Hannan's example and Lemma 2.8 in Chapter Two indicate that the covariance matrix for a periodic process of period N is a NxN circulant matrix. It is unlikely that non-periodic processes would result in circulant covariance matrices.

This observation does however suggest that the "asymptotically equivalent" approach might be fruitful because any discrete stochastic process can be approximated, in norm, by a sequence of periodic processes of increasingly large periods.

A slightly different embedding results if we eliminate the t_n terms in Equations (74) and (76). The resulting circulant matrix, of order 2n-1, is uniquely determined by its first row:

$$[t_0\ t_{-1}\ \cdots\ t_{-n+1}\ t_{n-1}\ \cdots\ t_1]. \tag{84}$$

The eigenvalues and eigenvectors are given by

$$\lambda_k = \sum_{j=1-n}^{n-1} t_j W_{2n-1}^{-jk} \tag{85}$$

and

$$z_k = [1,\ W_{2n-1}^{k},\ \ldots,\ W_{2n-1}^{k(2n-2)}]', \tag{86}$$

respectively.

Chapter 4. Conclusions

In this chapter we will comment on the discrete (numerical) analogues that we have defined and developed in this thesis and discuss where we feel further work is needed.

While most engineers and scientists working in the communication and information fields are aware of the dichotomy between time- and band-limited functions, we have not found a completely satisfactory illustration of this fact for the discrete functions. In fact, while the concepts of discrete time- and band-limited functions may be intuitively well known, we have not been able to find explicit definitions given anywhere in the literature. Most authors state their results for the continuous variable case and then assume that the discrete variable analogues follow directly without any difficulties. We believe that such conclusions should be established on rigorous as well as intuitive grounds.

Therefore we have defined TL(N), the space of discrete time-limited functions, and BL(N), the space of band-limited functions, as two N-dimensional complex vector spaces. Because of the mathematical isomorphism between TL(N) and BL(N), it is not surprising that discrete time- and band-limited functions are used interchangeably without regard for the consequences. This situation is most apparent when we attempt to define a discrete convolution operation as we have already mention in the introduction. Extensive use of the discrete convolution operation suggests that further work is needed in this area.

While the approach to the discrete Fourier transform (DFT) by way of generalized functions (distributions) has been rigorously established, the inverse problem has largely been ignored. That is, since there are many functions which possess the same discrete sample function, how do we interpret the discrete Fourier spectrum, the DFT of the discrete sample function, with this fact in mind?

Our interpretation is that the DFT of a discrete time-limited function is a discrete band-limited function; that is, we are only able to determine N distinct frequencies in the discrete Fourier spectrum of N sample values and this includes the constant ("DC") term.

We also observe that the sampling interval, the interval between successive sample values, does not explicitly appear in the DFT, only the number of samples.

What happens if the sought after frequency does not correspond to one of the N distinct frequencies in the spectrum? That is, does the DFT of the sample values of a function represent the sample values of the Fourier transform of the function? The generalized function approach clearly shows that the answer is yes.

In researching this thesis, we have from time to time conducted a number of numerical experiments on electronic computers to test various features of the discrete analogues developed in this thesis. With respect to the question posed above, our numerical examples suggest that the answer is yes.

A more difficult problem is to find the relationship between the discrete Fourier spectrums of the same function transformed (DFT) using

two different values of N. The early fast Fourier transform (FFT) algorithms required the number of data values, being transformed, to be an integral power of two ($\log_2 N$ is a positive integer). When the number of data values was not an integral power of two, the advice given in the literature at that time was to add a sufficient number of zeros as data values so that the FFT algorithm could be used. To our knowledge, no one has explained the consequence those zeros have on the resulting discrete Fourier spectrum.

Our numerical experimentation shows that the general level of "noise" (error) increases as the number of zeros is increased but the effect on individual terms is largely unpredictable. The power-of-two restriction has been removed from the FFT algorithm but two discrete Fourier spectrums, for different values of N, are still difficult to compare.

We must always remember that the DFT of any finite set of complex numbers, considered as a truncated sample path of some process, only represents an estimate of the frequency spectrum of this process. Consequently our interpretation of this frequency spectrum is determined by the significance we place on the various frequencies making up this spectrum. For example, what significance should we place on the frequency spectrum of the DFT of a set of randomly chosen complex numbers?

A good deal of work has been done for the continuous variable case concerning data contaminated with random errors ("noise") generated by a stationary process having a normal distribution. More work is needed in the discrete case for non-normal distributions.

Another area of further work is to investigate various limits of the DFT of N terms of a sequence in l^2 (the space of complex square-summable sequences) as N increases without bound (the limit of a sequence whose Nth term is an element of BL(N)). This limit would be the discrete analogue of a Fourier transform. We investigated several possible limits without much success. As far as we were able to determine, there is no discrete analogue of Plancherel's Theorem on l^2.

There are very few instances where natural phenomena are truly periodic, for all t. Consequently, there is a great need to analyze observations whose amplitude diminishes in time. For example, seismic waves, including the free oscillations of the earth, decay in amplitude as they propagate further from the source so that the use of "dampened periodic" functions would probably be more realistic than periodic functions. We might, for example, investigate the use of discrete analogues of Hankel transforms. In addition, the application of Z-transforms and Walsh transforms could be investigated.

When investigating several expansions using basis elements of C(N), the N-dimensional complex vector space, we didn't consider the discrete analogue of (sin x)/x functions which are encountered in sampling theorems of non-discrete band-limited functions.

The discrete Karhunen-Loeve (DKL) expansion represents a generalization of the DFT expansion in that the coefficients in the DKL expansion reduce to the coefficients of the DFT expansion when the original stochastic process is circular or periodic.

The main difficulty in using the DKL expansion lies in the frequent lack of physical interpretation of the uncorrelated random variables of the new stochastic process. These random variables are the eigenvectors of the covariance matrix for the original stochastic process and the covariance matrix is a Toeplitz matrix when the process is stationary in the wide sense. This difficulty is partly due to the lack of explicit expressions for the eigenvalues and eigenvectors of Toeplitz matrices, except in the special case when the process is circular or periodic (the Toeplitz matrix is then a circulant matrix).

We searched unsuccessfully for explicit expressions for the eigenvalues and eigenvectors of Toeplitz matrices. As we pointed out in the introduction, the existence of the DKL expansion is equivalent to the existence of a unitary transformation which transforms the original Hermitian, Toeplitz matrix into a diagonal matrix. It is known that such unitary transformations do exist for Hermitian matrices. Our work in Chapter Two on the DKL expansion shows that the eigenvectors of the covariance (Toeplitz) matrix form an orthonormal basis for the N-dimensional vector space determined by the process. Therefore the DKL expansion is determined by an orthonormal basis possessing some (as yet unknown) property.

The continuous Karhunen-Loeve expansion possesses an optimal property with respect to the minimum number of terms in the expansion needed to represent or approximate the original stochastic process (Davenport and Root; Watanabe). We did not find a discrete analogue for this optimal property but we believe that if a discrete analogue

were found then this analogue could lead to the unknown property mentioned in the previous paragraph.

The partial results that we did obtain, in Chapter Three, are of some value in the cases where the realizations of the random variables become negligibly small or approximately periodic for large N. For small positive integers N, up to 30 or so, the numerical calculations of the eigenvalues and eigenvectors of Hermitian (Toeplitz) matrices, and hence the DKL expansion, can be found using either the Jacobi method or the Householder method.

While the DKL expansion may be difficult to interpret, it may also be computationally efficient to use. The DKL expansion is determined by the covariance matrix for the stochastic process and the covariance matrix can be estimated from a single realization of a process. A process, having many random variables, may have only a few significant terms in its DKL expansion. That is, a large number of correlated random variables may be represented by a fewer number of uncorrelated random variables in its DKL expansion. This seems to be the situation when the Karhunen-Loeve expansion is applied to the problems of pattern recognition and speech communication (Chien and Fu; Watanabe).

Some work is currently being done on defining a measure of the informational content ("entropy") of a set of observations; this measure involves the use of Parseval's theorem and covariance matrix of the process. This approach constructs discrete functions which maximize the entropy in a given set of observations and as such offers an alternative to the use of discrete time- and band-limited functions.

Bibliography

Aitken, A.C., 1958, Determinants and Matrices: Edinburgh, Oliver and Boyd, 144p.

Akaike, Hirotugu, 1973, Block Toeplitz Matrix Inversion: SIAM J. Appl. Math., v. 24, no. 2, p. 234-241.

Apostol, Tom M., 1957, Mathematical analysis: Readings, Mass., Addison-Wesley, 553p.

Barber, N.F., 1966, Fourier methods in geophysics, in Runcorn, S.K., Methods and techniques in geophysics, v. 2: London, Interscience Publishers, 314p.

Box, George E.P., and Jenkins, Gwilym M., 1970, Time series analysis, forecasing and control: San Francisco, Holden-Day, 553p.

Carslaw, H.S., 1950, An introduction to the theory of Fourier series and integrals: New York, Dover Publications, 368p.

Chien, Y.T., and Fu, K.S., 1967, On the generalized Karhunen-Loeve expansion: IEEE Trans. on Inform. Theory, v. IT-13, no. 4, p. 518-520.

Cooley, J.W., and Tukey, J.W., 1965, An algorithm for machine computation of complex Fourier series: Math. Comp., v. 19, p. 297-301.

Davenport, Wilber B. Jr., and Root, W.L., 1965, Introduction to random signals and noise: New York, McGraw-Hill, 393p.

Gray, Robert M., 1972, On the asymptotic eigenvalue distribution of Toeplitz matrices: IEEE Trans. on Inform. Theory, v. IT-18, no. 6, p. 725-730.

Grenander, Ulf, and Szego, Gabor, 1958, Toeplitz forms and their applications: Berkeley, California, Univ. of Calif. Press, 245p.

Hadsell, Frank A., 1972, Linear systems theory I: unpublished course notes, Colorado School of Mines, Golden, Colorado.

Halmos, Paul R., 1958, Finite dimensional vector spaces: Princeton, N.J., Van Nostrand Reinhold Co., 199p.

Hannan, E.J., 1967, Time series analysis: London, Methuen, 152p.

Hoog, Robert V., and Craig, Allen T., 1970, Introduction to mathematical statistics: London, The Macmillan Co., 415p.

Jenkins, Gwilym M., and Watts, Donald G., 1968, Spectral analysis and its applications: San Francisco, Holden-Day, 525p.

Landau, H.J., and Pollak, H.O., 1961, Prolate spheroidal wave functions, Fourier analysis and uncertainty - II: Bell Systems Tech. J., v. 40, p. 65-84.

Lee, Joseph R., 1972, Advanced calculus with linear algebra: Readings, Mass., Addison-Wesley, 218p.

Levinson, Norman, 1947, The Wiener rms (root mean square) error criterion in filter design and prediction: J. of Math. and Phys., v. 25, no. 4, p. 261-278.

Mullen, James A., 1968, On the reducibility of Toeplitz eigenvalue equations: IEEE Trans. on Inform. Theory, v. IT-14, no. 1, p. 162-163.

Papoulis, Athanasios, 1962, The Fourier integral and its applications: New York, McGraw-Hill, 318p.

Parzen, Emanuel, 1962, Stochastic processes: San Francisco, Holden-Day, 324p.

Pearl, Judea, 1973, On coding and filtering stationary signals by discrete Fourier transformations: IEEE Trans. on Inform. Theory, v. IT-19, no. 2, p. 229-232.

Schoenberg, I.J., 1950, Finite Fourier series and elementary geometry: Amer. Math. Monthly, v. 57, no. 6, p. 390-404.

Selin, Ivan, 1965, Detection theory: Princeton, N.J., Princeton Univ. Press, 128p.

Shilov, Georgi E., 1961, An introduction to the theory of linear spaces: Englewood Cliffs, N.J., Prentice-Hall, 310p.

Slepian, D., and Pollak, H.O., 1961, Prolate spheroidal wave functions, Fourier analysis and uncertainty - I: Bell Systems Tech. J., v. 40, p. 43-63.

Smylie, D.E., Clarke, G.K.C., and Ulrych, T.J., 1973, Analysis of irregularities in the earth's rotation, in Methods of computational physics, p. 391-430.

Trench, William F., 1964, An algorithm for the inversion of finite Toeplitz matrices: SIAM J. Appl. Math., v. 12, no. 3, p. 515-522.

Vinograde, Bernard, 1967, Linear and matrix algebra: Boston, Heath, 252p.

Watanabe, Satosi, 1965, Karhunen-Loeve expansion and factor analysis - theoretical remarks and applications: Proc. 4th Prague Conf. on Information Theory, p. 635-660.

Widom, Harold, 1965, Toeplitz matrices, in Hirschman I.I. Jr., ed., Studies in real and complex analysis, v. 3: Mathematical Association of America.

Wiener, Norbert, 1949, Extrapolation, interpolation, and smoothing of stationary time series: Cambridge, Mass., The M.I.T. Press, 163p.

Wiener, Norbert, 1958, The Fourier integral and certain of its applications: New York, Dover Publications, 201p.

Wiggins, Ralph A., and Robinson, Enders A., 1965, Recursive solutions to the multichannel filtering problem: J. of Geophys. Res., v. 70, no. 8, p. 1885-1891.

Zohar, Shalov, 1969, Toeplitz matrix inversion, the algorithm of W.F. Trench: J. of Assoc. Comp. Mach., v. 16, no. 4, p. 592-601.