DECOMPOSITION OF TIME SERIES INTO DETERMINISTIC AND INDETERMINISTIC COMPONENTS

by

Charles T. Brodnax

Technical Report No. 105
Department of Statistics ONR Contract

April 19, 1971

Research sponsored by the Office of Naval Research
Contract N00014-68-A-0515
Project NR 042-260

DEPARTMENT OF STATISTICS
Southern Methodist University

Brodnax, Charles Troy

B.S., University of Houston, 1960
B.S.E.E., University of Houston, 1960
M.E.E., New York University, 1962

Decomposition of Time Series Into Deterministic and Indeterministic Components

Advisers: Visiting Industrial Assistant Professor R. E. Kromer
and Associate Professor C. H. Kapadia

Doctor of Philosophy degree conferred May 17, 1971

Dissertation completed April 19, 1971

Most of the procedures employed for detection of hidden periodicity require the spectrum be estimated for a set of samples. This sampled spectrum is plotted and observed for peaks indicative of periodic components. Various tests are then applied at the frequencies at which these peaks occur. Observations and decisions by the investigator is the key part in deciding where periodic components occur.

This dissertation presents a systematic approach for detection of periodic components and the consequent decomposition of the spectral distribution function. The proposed procedure is readily adaptable to a general purpose computer. Theorem 2.2, a major contribution of this dissertation, gives the conditions under which this procedure can be applied. These conditions are very general and most stationary processes of practical interest are included. An estimator for the frequency of periodicity is

derived as a result of theorem 2.1. Though not original theorem 2.1 and its proof are included for completeness.

Two tests for periodicity are presented in Chapter III. The first is a well known chi-square test based on a known constant power density spectrum under the null hypothesis H_0: no periodic component is present in the data. The second test does not depend on a known constant power density spectrum but rather uses some of the available data to estimate the value of the spectrum under H_0. This second test is intended as a large sample test and is believed to be original since it could not be found in the literature.

A computer program has been written using the procedures proposed in this dissertation for spectrum decomposition. Several test cases of data to estimate the value of the spectrum under H_0. This second test is intended as a large sample test and is believed to be an original contribution.

ACKNOWLEDGEMENTS

The author wishes to express his gratitude to his dissertation committee, Dr. C.H. Kapadia, Dr. A.M. Kshirsagar, Dr. Donald B. Owen and Dr. R.E. Kromer. A special thanks is extended to Dr. Kromer for his continued technical direction and encouragement and to Dr. Kapadia for his advice and assistance, both administrative and technical.

Thanks are due to the author's employer, Texas Instruments Incorporated, for continued support both financial and in time.

A very special word of appreciation go to my wife and son for their love and understanding throughout the years of this educational endeavor.

Finally my thanks are extended to the faculty and staff of the Statistics Department of Southern Methodist University for the countless assistances they have given throughout our association.

TABLE OF CONTENTS

Page

LIST OF ILLUSTRATIONS

CHAPTER I

STATEMENT OF THE PROBLEM

1.1 Introduction

"Prediction theory" is considered somewhere between witchcraft and "show biz" by many persons not familiar with the field, though they may be intelligent and educated. Yet these same persons will predict the paths of the planets based on Newton's Laws of Motion, or will predict current densities in a solid state device using Maxwell's Equations. Why the willingness to accept some predictions while rejecting the general field? The difference is a physical model or models with parameters related to familiar phenomena in nature. Models based on Newton's Laws of Motion, Maxwell's Equations, etc., are said to be deterministic and have been used and generally accepted almost since the beginning of time. These deterministic models use past values of variables to accurately predict future performance.

A second type of model is a stochastic model, which differs from the deterministic model in that past values of variables only partly determine the future of the process. Probabalistic effects enter a model because either they are unknown, because one is interested in the average or "typical" behavior, or because they represent a host of minor deterministic effects which could not be accounted for even if one wanted to. With the advent of large computing machines there is an ever increasing emphasis on the use of models and simulation as analytical tools. Many processes of interest such as radar detection, manufacture of integrated circuits, etc., can only be described with stochastic models. These needs have given great impetus to the development of stochastic models.

Significant attention has been paid to stochastic models in the literature. Since the late 19th Century, Autoregressive (AR) and Moving Average (MA) models have received more and more attention. Some of the more recent publications covering autoregression and moving average processes are Akutowicz[1], Cox and Miller[3], Jenkins and Watts[4], Parzen[7] and Whittle[10]. Kromer[5] shows that any discrete process whose power spectral density is absolutely continuous, bounded from above and away from zero can be represented as an Autoregressive process. This type of process is of particular interest since it is very useful in modeling human response. Measurement of human response is normally in the presence of mechanical or electrical equipment which introduce periodic components due to gear backlash, bandwidth limitations and other resonances. Jenkins and Watts[4] and Parzen[7] discuss statistical tests for hidden periodicity in the estimated spectrum of an AR or MA process. What is needed in order to apply these tests is an estimator for the frequencies of periodicity that can be applied in a systematic manner.

This dissertation presents a systematic approach for estimating frequency of periodicity, testing for periodicity, removing periodic components from data and fitting the remaining data to an autoregressive model. This procedure requires that:

1. The sample autocovariance function be computed from the sample data. This is

$$R_{XX}(\tau) = \frac{1}{T} \sum_{t=1}^{T-|\tau|} X_t X_t \qquad 1.1.1$$

2. Autoregressive model of order M be fitted to the sample autocovariance function. Two methods for determining the order M are suggested and referenced. One is based on the residual sum of squares and the second on confidence intervals of the terms of order greater than M. The model for the process is

$$X_t = \sum_{i=1}^{M} \hat{\alpha}_i X_{t-i} + N_t \qquad 1.1.2$$

where $N_t \sim \text{iid}(0, \sigma_N^2)$

and $\hat{\alpha}_i$ are the estimated autoregressive coefficients.

3. The function $f(z; \hat{\alpha}_1, \hat{\alpha}_2, \ldots) = 1 - \sum_{i=1}^{M} \hat{\alpha}_i z^{-i}$ be factored. The zeroes of $f(z; \hat{\alpha}_1, \hat{\alpha}_2, \ldots)$ nearest the unit circle are used to estimate the frequency of periodicity. This estimator will be shown to be

$$\hat{f}_j = \frac{1}{2\Pi\Delta t} \tan^{-1} \frac{Im(z_j)}{Re(z_j)} \qquad 1.1.3$$

if Δt = time between samples

z_j = zero of $f(z; \hat{\alpha}_1, \hat{\alpha}_2, \ldots)$ nearest unit circle

and $I_m(z_j)$, $Re(z_j)$ = imaginary and real components of z_j respectively.

4. The data be tested for existence of a periodic component at $\hat{f}_j$. Both a chi-square and an F-test are presented for this

application.

5. The data be reduced by the component at $\hat{f}_j$ when the test indicates presence of periodicity. Fourier coefficients are used to estimate the amplitude of the periodic component. These are

$$A_T(f_j) = \frac{2}{T}\sum_{t=1}^{T} X_t \ \mathrm{Cos}\ 2\Pi f_j t \qquad 1.1.4$$

$$B_T(f_j) = \frac{2}{T}\sum_{t=1}^{T} X_t \ \mathrm{Sin}\ 2\Pi f_j t \qquad 1.1.5$$

The periodic component is estimated to be

$$P_{f_j}(t) = A_T(f_j)\ \mathrm{Cos}\ 2\Pi f_j t + B_T(f_j)\ \mathrm{Sin}\ 2\Pi f_j t \qquad 1.1.6$$

6. Steps 1 through 5 be repeated on the reduced data until no additional periodic components are found.

Theorem 2.2 justifies the application of this procedure to data containing periodic components. Kromer[5], as stated above, proved that processes with non zero, bounded from above, and absolutely continuous power spectral density have autoregressive representations. Theorem 2.2 extends this result to include processes with a finite number of step discontinuities in the power distribution spectrum. The proof of this theorem and the resulting application of the procedure to decomposition of a power distribution spectrum is the major contribution of this thesis.

1.2 Processes With Continuous Spectra

Consider a linear system described by:

$$\sum_{i=1}^{\infty} a_i \frac{d^i X}{dt^i} + (X(t)-\mu) = N(t) \qquad 1.2.1$$

where N(t) is the input to the system and X(t) is the response of the system to that input. There are numerous examples of systems that are described by Equation 1.2.1. Most feedback control systems are analyzed by use of the model of Equation 1.2.1 (possibly with $a_i = 0$ for $i >$ some m.) Many chemical processes and heat transfer relationships can be described by Equation 1.2.1. Quite often a nonlinear process, such as the equations of a missile or projectile can be represented by Equation 1.2.1 over a limited region. There are sufficient applications of Equation 1.2.1 to justify development of a set of analysis for studying such systems.

In discrete time the analogue to Equation 1.2.1 is:

$$\sum_{i=1}^{\infty} \alpha_i (X_{t-i} - \mu) + (X_t-\mu) = N_t \qquad 1.2.2$$

where again N_t is the value of the driving function at the t^{th} sample time and X_t is the associated response. Systems described by Equation 1.2.2 are sometimes called sampled data systems. While sampled data systems have been studied for years the recent acceptance and application of digital computers as system controllers has greatly expanded the sampled data system technology. When the N_t are uncorrelated samples of a random function developing in time, X_t becomes a stochastic process and is referred to as an Autoregressive Process.

This dissertation will be primarily concerned with Autoregressive Processes. A good physical example of the processes of interest here is sampled values of human response to a random perturbation. The autocovariance function of an Autoregressive Process is:

$$\sum_{i=1}^{\infty} \alpha_i \, \Gamma_{XX}(k+i) + \Gamma_{XX}(k) = E[N_t(X_{t+k} - \mu)] \qquad 1.2.3$$

which is derived from

$$\Gamma_{XX}(k) = E(X_t-\mu)(X_{t+k}-\mu) \qquad 1.2.4$$

$$\Gamma_{XX}(-k) = \Gamma_{XX}(k) \qquad 1.2.5$$

If the N_t are uncorrelated random variables with variance σ_N^2 , then:

$$\Gamma_{XX}(k) + \sum_{i=1}^{\infty} \alpha_i \, \Gamma_{XX}(k+i) = \begin{cases} 0 & k < 0 \\ \sigma_N^2 & k = 0 \end{cases} \qquad 1.2.6$$

According to the Wold Decomposition Theorem (see Whittle[10], p. 23 or Wold[11]) any stationary process, X_t, can be uniquely represented as the sum of two mutually uncorrelated processes

$$X_t = V_t + W_t \qquad 1.2.7$$

where V_t is deterministic and W_t can be represented as the one sided moving average of a stationary uncorrelated sequence:

$$W_t = \sum_{i=0}^{\infty} b_i \, N_{t-i} \qquad 1.2.8$$

The W_t component is referred to as purely non-deterministic: it has limiting prediction error equal to σ_N^2. The deterministic property of V_t refers to linear determinism and means that V_t can be predicted by a linear combination of V_{t-1}, V_{t-2}, . . . with zero mean square error.

Akutowicz[1] proves that the necessary and sufficient conditions for the existence of

$$\sum_{i=0}^{\infty} a_i W_{t-i} = N_t \qquad 1.2.9$$

that

$$\phi(z) = \sum_{i=0}^{\infty} b_i z^i \qquad 1.2.10$$

be analytic inside the unit circle provided the b_i is a set of constants such that

$$\sum_{i=0}^{\infty} |b_i|^2 < \infty \qquad 1.2.11$$

In this case

$$\sum_{i=0}^{\infty} a_i z^i = \frac{1}{\sum_{i=0}^{\infty} b_i z^i} \qquad 1.2.12$$

The Wold Decomposition Theorem and the fact that many processes can be fitted rather easily to the autoregressive model is the justification for studying this model.

Suppose that the deterministic component, V_t, in Equation 1.2.8 is zero. Then,

$$\sum_{i=0}^{\infty} a_i X_{t-i} = N_t \qquad 1.2.13$$

is the model for the given process. Of particular interest are processes that can be fitted to.

$$\sum_{i=0}^{m} a_i X_{t-i} = N_t \qquad 1.2.14$$

for a reasonable number of terms m. Corresponding to Equation 1.2.14, the autocovariance function satisfies

$$\sum_{n=0}^{m} a_n \Gamma_{XX} (k+n) = \begin{cases} 0 & k<0 \\ \sigma_N^2 & k=0 \end{cases} \qquad 1.2.15$$

which leads to the power density spectrum

$$f_x(\omega) = \frac{\sigma_N^2}{2\Pi \left| \sum_{n=0}^{m} a_n e^{j(n\omega)} \right|^2} \quad - \Pi \leq \omega < \Pi \qquad 1.2.16$$

where

$$f_x(\omega) = \frac{1}{2\Pi} \sum_{k=-\infty}^{\infty} \Gamma_{XX} (\tau) e^{j\omega k} \qquad 1.2.17$$

and

$$j = \sqrt{-1} \qquad 1.2.18$$

The function $f_x(\omega)$ given in 1.2.16 is continuous on $-\Pi \leq \omega < \Pi$ and describes the spectrum of the indeterministic component of a stationary process.

1.3 Processes With Discrete and Mixed Spectra

The simplest form of deterministic stationary process satisfying V_t of Equation 1.2.8 is,

$$V_t = a \cos (\omega_0 t + \phi) \qquad t = 1, 2, \ldots \qquad 1.3.1$$

where ϕ is uniformly distributed over $(-\Pi, \Pi)$. A more general process is,

$$V_t = \sum_{i=1}^{p} a_i \text{ Cos } (t\ \omega_{0i} + \psi_i) \qquad 1.3.2$$

with $a_1, \ldots, a_p$ and $\omega_{0_1}, \ldots, \omega_{0p}$ constants and $\psi_1, \ldots, \psi_p$ identically and independently distributed uniformly on $(-\Pi, \Pi)$. The autocovariance function of 1.3.2 is,

$$\bar{\Gamma}_{VV}(k) = \sum_{i=1}^{p} a_i^2 \text{ Cos } (k\ \omega_{0i}) \qquad 1.3.3$$

and the autocorrelation function is,

$$\bar{\rho}_{VV}(k) = \sum_{i=1}^{p} \frac{a_i^2 \text{ Cos } (k\omega_{0i})}{\sum_{i=1}^{p} a_i^2} \qquad 1.3.4$$

The spectral density function of $\Gamma_{VV}(k)$ does not exist in the usual sense so it is customary to work with the spectral distribution function. In terms of the spectral distribution function, $F_V(\omega)$,

$$\bar{\rho}_{VV}(k) = \int_{-\infty}^{\infty} e^{ik\omega}\, dF_V(\omega) \qquad 1.3.5$$

Corresponding to;

$$V_{t_i} = {}_i \text{ Cos } (t\ \omega_{0i} + \psi_i) \qquad 1.3.6$$

the spectral distribution function is,

$$\bar{\rho}_{Vi}(\omega) = \begin{cases} 0 & \omega \leq -\omega_{0i} \\ \frac{a_i}{2} & -\omega_{0i} < \omega \leq \omega_{0i} \\ a_i & \omega_{0i} < \omega \end{cases} \qquad 1.3.7$$

so that corresponding to V_t one can write,

$$\bar{\underline{\rho}}_V(\omega) = \frac{\sum_{i=1}^{p} a_i^2 \; F_{Vi}(\omega)}{\sum_{i=1}^{p} a_i^2} \qquad 1.3.8$$

The process V_t is said to possess a discrete or line spectrum. Its spectral distribution function changes value at a finite, or at most countable, number of points. A great many of the deterministic processes encountered in practice can be represented by V_t as given in Equation 1.3.2. However, rarely do these processes exist, nor are they measurable, by themselves but usually coexist with an indeterministic component W_t as expressed in Equation 1.2.9. In this case the spectral distribution function $\bar{\underline{\rho}}_X(\omega)$ is,

$$\bar{\underline{\rho}}_X(\omega) = \bar{\underline{\rho}}_V(\omega) + \bar{\underline{\rho}}_W(\omega) \qquad 1.3.9$$

where $\bar{\underline{\rho}}_V(\omega)$ is discrete and $\bar{\underline{\rho}}_W(\omega)$ is absolutely continuous. $\bar{\underline{\rho}}_X(\omega)$ is said to be a mixed spectrum or is said to contain hidden periodicities. This thesis is directed toward the separation of $\bar{\underline{\rho}}_X(\omega)$ into $\bar{\underline{\rho}}_V(\omega)$ and $\bar{\underline{\rho}}_W(\omega)$, which will be accomplished by predicting frequencies of periodic components, testing for hidden periodicity at these frequencies and removing these components from $\bar{\underline{\rho}}_X(\omega)$. The necessary theorem to justify this procedure will be proven.

CHAPTER II

ESTIMATION OF FREQUENCY OF PERIODIC COMPONENTS

This chapter will derive an estimator for the frequency of a periodic component in a time series. In this derivation, extensive use will be made of transform analysis and techniques. Since the derivation of transform techniques is well covered in the literature only definitions will be presented here. A theorem relating the transform of the ACF to periodic components will be presented. An estimator based on this relationship is proposed for use in the power distribution spectrum. The second theorem in this chapter justifies fitting the data with an autoregressive model.

2.1 Convergence of Sample Autocovariance Function

Extensive use will be made of the sample autocovariance function (ACVF) in this and the following chapters. The sample ACVF will be used to estimate the population ACVF and a process model derived from this estimator.

The usual estimator for the ACVF is

$$R_{XX}(v) = \frac{1}{T} \sum_{t=1}^{T-|v|} x_t \, x_{t+v}, \quad |v| = 0, 1, \ldots, T-1 \qquad 2.1.1$$

for a process of known zero mean, and

$$R_{XX}(v) = \frac{1}{T} \sum_{t=1}^{T-|v|} (x_t-\bar{x})(x_{t+v}-\bar{x}), \quad |v| = 0, 1, \ldots, T-1 \qquad 2.1.2$$

for a process of unknown mean, where $\bar{x}$ is the sample mean. The estimator $R_{XX}(v)$ has been widely used because of an intuitive appeal but as discussed by Jenkins and Watts[4] is not based on any maximum liklihood or least square error criteria. In fact such criteria are extremely difficult if not impossible to apply except in the simplest of cases. For that reason $R_{XX}(v)$ has been selected as the population ACVF estimator. The properties of this estimator will be discussed in the following paragraphs.

The first and second moment properties of $R_{XX}(v)$ are derived by Jenkins and Watts[4] for a signal $x(t)$ that is a realization of a stationary stochastic process $X(t)$ which has properties

$$E\,[X(t)] = 0 \qquad 2.1.3$$

$$\mathrm{Cov}\,[X(t),\, X(t+u)] = \gamma_{XX}(u) \qquad 2.1.4$$

$$\begin{aligned}\mathrm{Cov}\,[X(t)X(t+u_1),\, X(v)X(v+u_2)] &= \gamma_{XX}(v-t)\gamma_{XX}(v-t+u_2-u_1) \\ &+ \gamma_{XX}(v-t+u_2)\,\gamma_{XX}(v-t-u_1) \\ &+ \kappa_4(v-t,\, u_1,\, u_2)\end{aligned} \qquad 2.1.5$$

The function $\kappa_4(v-t, u_1, u_2)$ is the fourth joint cumulate of the stochastic process $X(t)$, and is zero if $X(t)$ is distributed normal. Bartlett[2] has shown that for other processes the contribution of this term can be neglected when deriving properties of the ACVF estimator.

With these assumptions on the process, the mean of $R_{XX}(v)$ is

$$E[R_{XX}(v)] = E\left[\frac{1}{T}\sum_{t=1}^{T-|v|} X_t X_{t+v}\right] \qquad 2.1.6$$

$$= \begin{cases} \gamma_{XX}(v)\ (1-\frac{v}{t}), & |v| = 0, 1, \ldots, T-1 \\ 0 & , |v| \geq T \end{cases}$$

$R_{XX}(v)$ is a biased estimator of the population ACVF for finite sample size. An unbiased estimator can be obtained by replacing T by $T-|v|$ in the denominator of the sum. The reason for not making this change will become apparent when the estimator variance is derived. Notice that as T tends to infinity, $R_{XX}(v)$ is asymptotically unbiased.

To obtain an expression for the variance of $R_{XX}(v)$, the covariance between two estimates $R_{XX}(v_1)$ and $R_{XX}(v_2)$ with $v_2 \geq v_1 \geq 0$ is derived.

$$\text{Cov}\ (R_{XX}(v_1),\ R_{XX}(v_2)) = \frac{1}{T^2}\sum_{t=1}^{T-|v_1|}\sum_{s=1}^{T-|v_2|} \text{Cov}\ [X_t X_{t+v_1},\ X_s X_{s+v_2}] \qquad 2.1.7$$

Assuming the K_4 term in 2.1.5 to be negligible

$$\text{Cov}\ (R_{XX}(v_1),\ R_{XX}(v_2)) = \frac{1}{T^2}\sum_{t=1}^{T-v_1}\sum_{s=1}^{T-v_2}\Big[\gamma_{XX}(s-t)\ \gamma_{XX}(s-t+v_2-v_1) + \gamma_{XX}\ (s-t+v_2)\ \gamma_{XX}\ (s-t-v_1)\Big] \qquad 2.1.8$$

Now let s-t=r and t=u and equation (2.1.8) can be written

$$\text{Cov}\ [R_{XX}(v_1),\ R_{XX}(v_2)] = \frac{1}{T^2}\sum_{r=-T+v_1}^{T-u_2}\{\gamma_{XX}(r)\ \gamma_{XX}(r+v_2-v_1) + \gamma_{XX}\ (r+v_2)\gamma_{XX}(r-v_1)\}\sum_u (1) \qquad 2.1.9$$

where

$$\sum_{u} (1) = \phi(r) = \begin{cases} T - v_2 - r & r \geq 0 \\ T - v_2 & -(v_2-v_1) \leq r < 0 \\ T - v_1 + r & -(T-v_1) \leq r < -(v_2-v_1) \end{cases} \qquad 2.1.10$$

Substituting (2.1.10) into (2.1.9) gives

$$\text{Cov}\,[R_{XX}(v_1),\, R_{XX}(v_2)] = \frac{1}{T^2} \sum_{r=-(T-v_1)}^{T-v_2} \phi(r)\ \{\gamma_{XX}(r)\ \gamma_{XX}(r+v_2-v_1) + \gamma_{XX}(r+v_2)\ \gamma_{XX}(r-v_1)\} \qquad 2.1.11$$

so that with $v_1 = v_2 = v$, the variance of $R_{XX}(v)$ is

$$\text{Var}[R_{XX}(v)] = \frac{1}{T^2} \sum_{r=-(T-v)}^{T-v} \phi(r)\ \{\gamma^2_{XX}(r) + \gamma_{XX}(r+v)\gamma_{XX}(r-v)\} \qquad 2.1.12$$

If the unbiased estimator, say $R'_{XX}(v)$, where

$$R'_{XX}(v) = \frac{T}{T-v}\ R_{XX}(v) \qquad 2.1.13$$

is selected, its variance will be

$$\text{Var}\ [R_{XX}(v)] = \frac{T^2}{(T-|v|)^2}\ \text{Var}[R_{XX}(v)] \qquad 2.1.14$$

For zero lag the variance of these two estimators is the same.As the lag increases, the variance of the unbiased estimator will be greater than that of $R_{XX}(v)$. As $|v|$ approaches its maximum allowable value of T-1,

$$\text{Var}\ [R'_{XX}(v)] \rightarrow T^2\ \text{Var}\ [R_{XX}(v)]$$

which can become quite large for large T unless the variance of $R_{XX}(v)$ is zero, the trivial case. Hence, even though $R'_{XX}(v)$ is unbiased, $R_{XX}(v)$ is usually preferred because it has smaller variance.

Referring back to Equation 2.1.6 it is seen that the expected value of $R_{XX}(v)$ is $\gamma_{XX}(v)$ for large T. From 2.1.12 the variance of $R_{XX}(v)$ is bounded above by a function proportional to 1/T. The covariance function $E[X(t)X(t+v)]$ can be estimated with arbitrarily small error from a sufficiently long record. Time averages over an infinite record and ensemble averages are equivalent for the covariance function of a stationary process, demonstrating the ergodic property of such a covariance function.

The bias demonstrated in 2.1.6 was based on a zero mean X(t). If such is not the case correction must be made for the mean, in which case $R_{XX}(v)$ can be written

$$R_{XX}(v) = \frac{1}{T} \sum_{t=1}^{T-|\gamma|} [(X_t-\mu)(X_{t+v}-\mu)] - (1 - \frac{|v|}{T})(\bar{X}-\mu)^2 \qquad 2.1.15$$

where $E[X(t)] = \mu$ and $\bar{X} = \frac{1}{T}\sum_{t=1}^{T} x_t$ is an estimator of μ. The expected value of $R_{XX}(v)$ from 2.1.15 is

$$E[R_{XX}(v)] = (1 - \frac{|v|}{T})\gamma_{XX}(v) - (1 - \frac{|v|}{T}) \operatorname{Var} \bar{X} \qquad 2.1.16$$

and since

$$\operatorname{Var}(\bar{X}) = \frac{1}{T} \sum_{v=-(r-1)}^{T-1} (1 - \frac{|v|}{T})\gamma_{XX}(v) \qquad 2.1.17$$

the effect of correcting for the mean is to increase the bias by terms of order $(\frac{1}{T})$ and higher.

Equation 2.1.12 shows that the mean square error of $R_{XX}(v)$ as the sample ACVF estimator is of order 1/T and hence its distribution tends to be clustered about $\gamma_{XX}(v)$ as T tends to infinity. $R_{XX}(v)$ is then a consistent estimator of $\gamma_{XX}(v)$. The ergodic property of $R_{XX}(v)$ does not imply that it hold for its Fourier transform, the power density spectrum estimator. It is usually true that the ergodic property does not hold for the Fourier transform of $R_{XX}(v)$ as a spectrum estimator. For a more detailed discussion of this last point see Jenkins and Watts,[4] page 222.

2.2 Process Sample Characteristic Function in Z Domain

Consider an autoregressive process given by

$$X_t = \sum_{k=1}^{\infty} \alpha_i X_{t-k} + N_t \qquad 2.2.1$$

For all except a few simple cases the model in 2.2.1 is awkward to work with in the time domain. Hence it is convenient to resort to transform methods for analysis of the above model. A particularly usable transform which will be reviewed in the succeeding paragraphs. A complete coverage of transform methods will not be attempted here. Only definitions of the transforms used later will be covered. For more complete development of transform methods the reader is referred to Zemanian[12] for the Fourier and Laplace transforms and to Kuo[6] for Z-transforms.

The z-transform for discrete functions can be defined as

$$\underline{\bar{X}}(z) = \sum_{k=0}^{\infty} X_{k.} z^{-k} \qquad 2.2.2$$

where X_k is the value of a time series at $t = k\Delta t$ and z is a complex variable. Since from the defining equation 2.2.2

$$X_{r-k} = z^{-k} X_r \qquad 2.2.3$$

z is sometimes regarded as a shift operator. The Fourier transform of the sequence X_k is (if $X_k = 0$ for $k < 0$)

$$\underline{\bar{X}}(\omega) = \sum_{k=0}^{\infty} X_k \, e^{-j\omega\Delta tk} \qquad 2.2.4$$

and the Laplace transform is

$$\underline{\bar{X}}(s) = \sum_{k=0}^{\infty} X_k \, e^{-s\Delta tk} \qquad 2.2.5$$

$$s = \sigma + j\omega$$

Under these definition

$$\underline{\bar{X}}(\omega) = \lim_{\sigma \to 0} \underline{\bar{X}}(s)$$

and

$$\underline{\bar{X}}(\omega) = \lim_{|z| \to 1} \underline{\bar{X}}(z)$$

More generally than Equation 2.2.4, $\underline{\bar{X}}(\omega)$ is defined as

$$\underline{\bar{X}}(\omega) = \sum_{k=-\infty}^{\infty} X_k \, e^{-j\omega\Delta tk} \qquad 2.2.6$$

2.3 Frequency Estimators Based on Z Transform

In Chapter I the spectra of stochastic processes were discussed and it was pointed out that the indeterministic component of a process yielded a continuous spectrum. Step discontinuities in the spectrum, on the other hand, result from periodic components. When modeling man-machine systems these periodic components are of much interest since they quite often disclose some information concerning the machine response. Tests for hidden periodic components have been reported by Whittle[10], Cox[3] and Miller and Parzen[7] as well as others. These tests assume that the frequency of occurence (f_0) is known before the test can be applied. If f_0 is not known some attempt must be made to estimate it. Estimation of f_0 will be the objective of this section. Discussion of tests for periodicity at f_o are postponed until Chapter III.

Assume that samples x_t, t = 1, 2, ..., T of a stationary time series have been measured. According to Section 2.1 a good estimator for the autocovariance function is

$$R_T(v) = \frac{1}{T} \sum_{t=1}^{T-|v|} (X_t-\bar{X})(X_{t+v}-\bar{X}) \qquad |v| < T \tag{2.3.1}$$

It is desired to predict X_{T+1} given the process x_t, t=1, ..., T. Consideration will be restricted to linear predictors, those for which X_{T+1} is predicted from a linear function of the known values of X_t. In this case the predictor $\hat{x}_{T+1}$ can be written

$$\hat{x}_T = \sum_{j=1}^{\infty} \alpha_j \, x_{T-j} \tag{2.3.2}$$

The α_j will be selected so that

$$E(\delta^2) = E(\hat{x}_T - X_T)^2$$

will be a minimum. When the α_j have been so selected, $\hat{x}_T$ is known as the linear least squares ($\ell.\ell.$s.) predictor. If $\delta = (\hat{x}_T - x_T)$ then

$$E(\delta^2) = E\left(\sum_{j=1}^{\infty} \alpha_j \quad x_{T-j} - x_T\right)^2 \qquad 2.3.3$$

$$= \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \alpha_j \quad \alpha_k \quad \gamma_{k-j} - 2\sum_{j=1}^{M} \alpha_j \, \gamma_j \; + \gamma_0$$

Now let $\sigma_X^2 = \gamma_0$ 2.3.4

so that

$$E(\delta^2) = \sigma_X^2 \left[1 - 2\sum_{j=1}^{\infty} \alpha_j \quad \rho_j \; + \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \alpha_j \, \alpha_k \, \rho_{k-j}\right] \qquad 2.3.5$$

where $\rho_k = \gamma_k / \gamma_0$ 2.3.6

$$= E\, X_t X_{t+k} \Big/ E X_t^2$$

Let the prediction be based on a process of order M so that differentiating Equation 2.3.5 with respect to α_j and equating the result to zero gives

$$-2\,\rho_j + 2\sum_{k=1}^{M} \hat{\alpha}_{kM}\, \rho_{k-j} = 0, \; j=1,\, 2,\, \ldots\, M \qquad 2.3.7$$

or

$$\hat{\alpha}_{1M} + \hat{\alpha}_{2M}\rho_1 + \hat{\alpha}_{3M}\rho_2 + \ldots + \hat{\alpha}_{MM}\rho_{M-1} = \rho_1 \qquad 2.3.8$$

$$\hat{\alpha}_{1M}\rho_1 + \hat{\alpha}_{2M} + \hat{\alpha}_{3M}\rho_1 + \ldots + \hat{\alpha}_{1MM}\rho_{M-2} = \rho_2$$

$$\hat{\alpha}_{1M}\rho_2 + \hat{\alpha}_{2M}\rho_1 + \hat{\alpha}_{3M} + \ldots + \hat{\alpha}_{MM}\rho_{M-3} = \rho_3$$

$$\vdots$$

$$\hat{\alpha}_{1M}\rho_{M-1} + \hat{\alpha}_{2M}\rho_{M-2} + \hat{\alpha}_{3M}\rho_{M-3} + \ldots + \hat{\alpha}_{MM} = \rho_M$$

which in matrix notation is

$$\begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{M-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{M-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{M-3} \\ \vdots & & & & \\ \rho_{M-1} & \rho_{M-2} & \rho_{M-3} & \cdots & 1 \end{pmatrix} \begin{pmatrix} \hat{\alpha}_{1M} \\ \alpha_{2M} \\ \hat{\alpha}_{3M} \\ \vdots \\ \hat{\alpha}_{MM} \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_M \end{pmatrix}$$

Solving this equation for the values of $\hat{\alpha}_{jM}$ satisfying the least squares Equations 2.3.8, the $(\hat{\alpha}_M)$ vector is

$$\begin{pmatrix} \hat{\alpha}_{1M} \\ \hat{\alpha}_{2M} \\ \vdots \\ \hat{\alpha}_{MM} \end{pmatrix} = \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{M-1} \\ \rho_1 & 1 & \cdots & \rho_{M-2} \\ \vdots & & & \vdots \\ \rho_{M-1} & \rho_{M-2} & \cdots & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_M \end{pmatrix} \qquad 2.3.9$$

if the inverse of the correlation matrix exists. The vector $(\hat{\alpha}_M)$ is the set of coefficients of x_{T-1}, x_{T-2},... that predict x_T with the smallest expected squared error. The equation

$$x_T = \hat{\alpha}_{1M}x_{T-1} + \hat{\alpha}_{2M}x_{T-2} + \cdots + \hat{\alpha}_{MM}x_{T-M} \qquad 2.3.10$$

is the linear regression of x_T on x_{T-1}, ..., x_{T-M}. Equation 2.3.10 is the model of particular interest here, that is the M^{th} order Autoregression Model. This is a fit of an M^{th} order process to the system at hand. A process of infinite order would be fitted by

$$x_T = \lim_{M \to \infty} \sum_{j=1}^{M} \alpha_{jM} x_{T-j} + n_t \qquad 2.3.11$$

Due to physical limitations in inverting the correlation matrix a finite M is usually chosen. A later section of this chapter will deal with the selection of M.

The predictor for the frequency of the periodic component will be derived by use of transform methods discussed earlier in this chapter. Consider the process in 2.3.11 where M may be countably infinite. Then population autocovariance function is

$$\Gamma_{XX}(k) = E[x_T x_{T+k}] \qquad 2.3.12$$

or

$$\Gamma_{XX}(k) = \begin{cases} \sum_{j=1}^{\infty} \alpha_j \, \Gamma(k-j), & k>0 \\ \sum_{j=1}^{\infty} \alpha_j \, \Gamma(j) + \sigma_N^2, & k=0 \end{cases} \qquad 2.3.13$$

Multiply both sides of Equation 2.3.13, $k \neq 0$, by z^{-k}

$$z^{-k} \, \Gamma_{XX}(k) = \sum_{j=1}^{\infty} \alpha_j \, z^{-k} \, \Gamma_{XX}(k-j) \qquad 2.3.14$$

which is good for all $k > 0$. Hence

$$\sum_{k=1}^{\infty} z^{-k} \, \Gamma_{XX}(k) = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \alpha_j z^{-k} \, \Gamma_{XX}(k-j) \qquad 2.3.15$$

$$= \sum_{j=1}^{\infty} \alpha_j \sum_{k=1}^{\infty} z^{-(k-1)} \, \Gamma_{XX}(k-1-j+1) z^{-1}$$

Using z as the shift operator in 2.2.3 gives

$$\sum_{k=1}^{\infty} z^{-k} \Gamma_{XX}(k) = \sum_{j=1}^{\infty} \alpha_j z^{-j} \sum_{k'=0}^{\infty} z^{-k'} \Gamma_{XX}(k') \qquad 2.3.16$$

which becomes, upon application of 2.2.2 to the right hand side

$$\sum_{k=1}^{\infty} z^{-k} \Gamma_{XX}(k) = \sum_{j=1}^{\infty} \alpha_j z^{-j} \overline{\underline{\Gamma_{XX}}}(z) \qquad 2.3.17$$

But

$$\sum_{k=0}^{\infty} z^{-k} \Gamma_{XX}(k) = \Gamma_{XX}(0) + \sum_{k=1}^{\infty} z^{-k} \Gamma_{XX}(k) \qquad 2.3.18$$

so

$$\overline{\underline{\Gamma_{XX}}}(z) - \Gamma_{XX}(0) = \overline{\underline{\Gamma_{XX}}}(z) \sum_{j=1}^{\infty} \alpha_j z^{-j} \qquad 2.3.19$$

or

$$\overline{\underline{\Gamma_{XX}}}(z) = \frac{\Gamma_{XX}(0)}{1 - \sum_{j=1}^{\infty} \alpha_j z^{-j}} \qquad 2.3.20$$

In terms of the sample autocovariance function Equation 2.3.20 becomes

$$\overline{\underline{R_{XX}}}(z) = \frac{R_{XX}(0)}{1 - \sum_{j=1}^{\infty} \hat{\alpha}_j z^{-j}} \qquad 2.3.21$$

where

$$R_{XX}(0) = \frac{1}{T} \sum_{t=1}^{T} X_t^2 \qquad 2.3.22$$

Equation 2.3.21 is the basis for the procedure to follow for decomposing the power distribution spectrum. The first theorem presented, theorem 2.1 relates the components of the time series X_t to the position of the poles of 2.3.21. Then the question of which time series satisfy Equation 2.3.21 is answered leading to the proof of theorem 2.2, an original theorem.

Theorem 2.1

A necessary and sufficient condition for the existence of a periodic component in a stationary time series, $x(t)$, is that the z-transform of $R(v) = E[x(t)x(t+v)]$

$$\underline{\overline{\Gamma}}(z) = \mathfrak{z}[\Gamma(v)] \qquad 2.3.23$$

have point singularities, or poles, on the unit circle, $|z| = 1$. All poles of $\underline{\overline{\Gamma}}(z)$ inside the unit circle is necessary and sufficient for $x(t)$ to be indeterministic.

Proof:

According to the Wold Decomposition Theorem

$$x_t = x_t^{(1)} + x_t^{(2)} \qquad 2.3.24$$

where $x_t^{(1)}$ is an indeterministic or regular time series and $x_t^{(2)}$ is deterministic or singular. By a regular time series it is meant that an $f_1(\omega)$ exists such that

$$\Gamma^{(1)}(v) = \int_{-\infty}^{\infty} e^{j\omega v} f_1(\omega)\, d\omega \qquad 2.3.25$$

and a singular time series implies that its frequency distribution function, $F_2(\omega)$, has a countable (at most) number of step discontinuities and is unvaring except at these points. For $f_1(\omega)$ in Equation 2.3.25 to exist there must be an $F_1(s)$, where $s = \sigma + j\omega$, that converges to

$F_1(\omega) = \int_{-\infty}^{\omega} f_1(x)dx$ as $s \to j\omega$. But $F_1(\omega)$ is bounded and absolutely continuous for all ω, hence $F_1(s)$, and consequently $f_2(s)$, is analytic in at least a narrow strip along the $j\omega$ axis. The transformation $z=e^{s\tau}$ is a conformal mapping from the s-plane into the z-plane and since the $j\omega$ axis in the s-plane maps into the unit circle in the z-plane,

$$\Gamma^{(1)}(z) = \mathcal{Z}[\Gamma^{(1)}_{(v)}] \qquad 2.3.26$$

in a small region around the unit circle.

The singular sequence defined by $x_t^{(2)}$ has a spectral distribution function $F_2(\omega)$ which is composed of a countable number of step discontinuities. Since

$$\overset{(2)}{\Gamma}(v) = \int_{-\infty}^{\infty} e^{j\omega v}\, d\, F_2(\omega) \qquad 2.3.27$$

then

$$\overset{(2)}{\Gamma}(v) = \sum_{\ell=1}^{\infty} A_\ell\, e^{j\omega_\ell v} \qquad 2.3.28$$

where the ω_ℓ are the values of ω at which the step discontinuities occur, and A_ℓ is the size of the step at ω_ℓ. The La Place Transform of $\Gamma^{(2)}_{(v)}$ is

$$\underline{\overline{\Gamma}}^{(2)}(s) = \int_0^{\infty} \sum_{\ell=1}^{\infty} A_\ell\, e^{j\omega_\ell t}\, e^{-st}\, dt \qquad 2.3.29$$

$$= \sum_{\ell=1}^{\infty} A_\ell \int_0^{\infty} e^{-(s-j\omega_\ell)t}\, dt$$

or

$$\underline{\overline{\Gamma}}^{(2)}(s) = \sum_{\ell=1}^{\infty} A_\ell/(s-j\omega_\ell) \qquad 2.3.30$$

Since $F_2(\omega)$ is a spectral distribution function

$$F_2\ (\infty) = \sum_{\ell=1}^{\infty} A_\ell = R(0) < \infty \qquad 2.3.31$$

All the A are positive so the A_ℓ must be finite. Therefore the poles of $\underline{\overline{\Gamma}}^{(2)}(s)$ occur at, and only at,

$$s_\ell = j\omega_\ell \qquad 2.3.32$$

which is on the j ω axis. The poles of $\underline{\overline{\Gamma}}^{(2)}(z)$ will occur at

$$\ell n\ z_\ell = j\omega_\ell \qquad 2.3.33$$

but

$$\ell n\ z_\ell = \ell n\ |z_\ell| + j\tan^{-1} \frac{I_m(z_\ell)}{Re(z_\ell)} \qquad 2.3.34$$

so

$$|z_\ell| = 1 \qquad 2.3.35$$

and

$$f_\ell = \frac{\omega_\ell}{2\Pi} = \frac{1}{2\Pi} \tan^{-1} \frac{I_m(z_\ell)}{Re(z_\ell)} \qquad 2.3.36$$

The z-transform of $\Gamma(v)$ can be written as the sum of the transforms of $\Gamma^{(1)}(v)$ and $\Gamma^{(2)}(v)$, or

$$\underline{\overline{\Gamma}}(z) = \underline{\overline{\Gamma}}^{(1)}(z) + \underline{\overline{\Gamma}}^{(2)}(z) \qquad 2.3.37$$

which demonstrates that R(z) has simple poles on the unit circle in direct correspondence to the periodic components in x(t). Each periodic component in x(t) contributes a complex conjugate pair of pole to $\underline{\overline{\Gamma}}(z)$ which are on the unit circle and at an angle whose tangent is $\pm 2\Pi f_\ell$ where f_ℓ is the frequency of the corresponding periodic component.

By the uniqueness of the La Place Transform pair the sufficiency part of the theorem follows immediately.

If the estimator R(v) is used for $\Gamma(v)$ the zeroes of $\underline{\overline{R}}(z)$ can be used to estimate f_ℓ. In this case

$$\hat{f}_\ell = \frac{1}{2\Pi} \tan^{-1} [\mathrm{Im}(\hat{z}_\ell) \;/\; \mathrm{Re}(\hat{z}_\ell)] \qquad 2.3.38$$

where $\hat{z}_\ell$ is the zero of $\underline{\bar{R}}$ (z) sufficiently near the unit circle.

Just how general are the sequences in Equation 2.3.24? First, $x^{(1)}(t)$ is a regular sequence with absolutely continuous spectrum, is bounded from above and away from zero so $x^{(1)}(t)$ includes all sequences of the form

$$x^{(1)}(t) = \sum_{n=1}^{\infty} \alpha_n x^{(1)}(t-n\Delta) + n_t \qquad 2.3.39$$

which are indeterministic. Next, it was stated that $x^{(2)}(t)$ was such that its spectral distribution function consisted of a countable number of finit step discontinuities. From the restriction on physical realizability it is necessary that for each frequency f_n for which a step occurs a like step must occur at $-f_n$. These two steps combine to form a $\text{Cos } 2\Pi f_n t$ term in $x^{(2)}(t)$. Hence $x^{(2)}(t)$ has representation of the form

$$x^{(2)}(t) = \sum_{n=1}^{N/2} A_n \text{ Cos } (2\Pi f_n\, t + \phi_n) \qquad 2.3.40$$

where N represents the number of discontinuities in $F_2(\omega)$. N may be either finite or countable infinite. If one of the f_n occurs at f=0, it will be assumed that there are two steps of size An/2 at that f_n.

Now consider the case where N in Equation 2.3.40 is finite and equal to 2M. In such a case $x^{(2)}(t)$ is

$$x^{(2)}(t) = \sum_{n=1}^{M} A_n \operatorname{Cos}\,(2\Pi f_n t + \phi_n) \qquad 2.3.41$$

The z-transform of $x^{(2)}(t)$ is

$$\underline{\overline{X}}^{(2)}(z) = \sum_{n=1}^{M} \frac{C_n z^{-1} + d_n}{1-2\rho_n z^{-1}+z^{-2}} \qquad 2.3.42$$

$$= \frac{\sum_{n=1}^{M} \prod_{\substack{i=1 \\ i \neq n}}^{M} \left(C_i z^{-1} + d_i\right)}{\prod_{n=1}^{M} (1-2\rho_n z^{-1} + z^{-2})}$$

But the numerator of 2.3.42 is

$$g(z) = \sum_{n=1}^{M} \prod_{\substack{i=1 \\ i \neq n}}^{M} \left(C_i z^{-1} + d_i\right) \qquad 2.3.43$$

a polynomial in z^{-1} of order M-1 which is bounded on $|z| = 1$. So

$$h(z) = 1/g\,(z) \qquad 2.3.44$$

does not equal zero on the unit circle, else g(z) would be unbounded at that point. Now $\underline{\overline{X}}^{(2)}(z)$ can be written as

$$\underline{\overline{X}}^{(2)}(z) = \frac{\sum_{n=0}^{M-1} b_n z^{-n}}{\sum_{n=0}^{2M} a_n z^{-n}} \qquad 2.3.45$$

The autocovariance function of $x^{(2)}(t)$ is

$$\gamma_2(\tau) = \sum_{n=1}^{M} A_n^2 \operatorname{Cos}\, 2\Pi\, f_n\, \tau \qquad 2.3.46$$

with z-transform

$$\underline{\overline{\Gamma_2}}(z) = \frac{\sum_{n=0}^{M-1} r_n \, z^{-n}}{\sum_{n=0}^{2M} s_n \, z^{-n}} \qquad 2.3.47$$

From 2.3.39 the autocovariance function of $x^{(1)}(t)$ is

$$\gamma_1(\tau) = \sum_{n=1}^{\infty} \alpha_n \, \gamma_1(\tau - n\Delta) \qquad 2.3.48$$

and

$$\gamma_1(0) = \sum_{n=1}^{\infty} \alpha_n \, \gamma_1(n) + \sigma_N^2 \qquad 2.3.49$$

with z-transform

$$\underline{\overline{\Gamma}}_1(z) = \frac{\Gamma_1(0)}{1 - \sum_{n=1}^{\infty} \alpha_n z^{-n}} \qquad 2.3.50$$

Since $x^{(1)}(t)$ and $x^{(2)}(t)$ are uncorrelated

$$\gamma_x(\tau) = \gamma_1(\tau) + \gamma_2(\tau) \qquad 2.3.51$$

and by the linearity of the z-transform

$$\underline{\overline{\Gamma_x}}(z) = \underline{\overline{\Gamma}}_1(z) + \underline{\overline{\Gamma}}_2(z)$$

or

$$\underline{\overline{\Gamma_x}}(z) = \frac{R(0)}{1 - \sum_{n=1}^{\infty} \alpha_n z^{-n}} + \frac{\sum_{n=0}^{M-1} r_n \, z^{-n}}{\sum_{n=0}^{2M} s_n \, z^{-n}} \qquad 2.3.52$$

Combining the terms on the right $\underline{\overline{\Gamma_x}}(z)$ may be written

$$\underline{\overline{\Gamma}_x}(z) = \frac{R(0) \sum_{n=0}^{2M} s_n z^{-n} + (1 - \sum_{n=1}^{\infty} \alpha_n z^{-n}) (\sum_{n=0}^{M-1} r_n z^{-n})}{(1 - \sum_{n=1}^{\infty} \alpha_n z^{-n}) \ (\sum_{n=0}^{2M} s_n z^{-n})} \qquad 2.3.53$$

The terms $\sum_{n=0}^{2M} s_n z^{-n}$ and $\sum_{n=0}^{M-1} r_n z^{-n}$ are polynomials hence they are analytic over the entire z-plane except at z = 0. Also, $(1 - \sum_{n=1}^{\infty} \alpha_n z^{-n})$ is analytic over the region on and outside the unit circle. Therefore the numerator is analytic on and outside the unit circle and likewise the reciprocal of the numerator is analytic over the same region except at points where the numerator is zero. Let

$$p(z) = R(0) \sum_{n=0}^{2M} s_n z^{-n} + (1 - \sum_{n=1}^{\infty} \alpha_n z^{-n}) (\sum_{n=0}^{M-1} r_n z^{-n}) \qquad 2.3.54$$

which will be analytic on and outside the unit circle. If its reciprocal is

$$q(z) = 1/p(z) \qquad 2.3.55$$

q(z) will be analytic over this same region except at isolated points where p(z) is zero. Over its region of analyticity q(z) can be expanded into a Laurent's series as

$$q(z) = \sum_{n=0}^{\infty} q_i z^{-i} \qquad 2.3.56$$

which has no zeroes on or outside the unit circle. Still assuming p(z) to have no zeroes on the unit circle, the terms $(\sum_{n=0}^{\infty} q_i z^{-i})$, $(1-\sum_{n=1}^{\infty} \alpha_n z^{-n})$, and $(\sum_{n=0}^{2M} s_n z^{-n})$ can be multiplied together to give

$$W(z) = w_0 (1 - \sum_{n=1}^{\infty} \frac{w_i}{w_0} \ z^{-i}) \qquad 2.3.57$$

so that

$$\Gamma_x(z) = \frac{1/w_0}{1 - \sum_{n=1}^{\infty} \frac{w_i}{w_0} z^{-i}} \qquad 2.3.58$$

The statement was made that p(z) in Equation 2.3.54 can be inverted on and outside the unit circle except where it was equal to zero. The zeroes of p(z) are the zeroes of $\Gamma_x(z)$ since p(z) is the numerator of $\Gamma_x(z)$. But

$$\lim_{|z|\to 1} \overline{\Gamma_x}(z) = \lim_{|z|\to 1} \overline{\Gamma^{(1)}}(z) + \lim_{s\to j\omega} \sum_{n=1}^{M} \frac{A_n\, s}{s^2 + (2\Pi f_n)^2} \qquad 2.3.59$$

or

$$\lim_{|z|\to 1} \Gamma_x(z) = \lim_{|z|\to 1} \overline{\Gamma^{(1)}}(z) + \sum_{n=1}^{M} \frac{A_n\, j\omega}{\omega_n^2 - \omega^2} \qquad 2.3.60$$

so that for p(z) to be zero on the unit circle it would be required that

$$\lim_{|z|\to 1} \overline{\Gamma^{(1)}}(z) = -\sum_{n=1}^{M} \frac{A_n\, j\omega}{\omega_n^2 - \omega^2} \qquad 2.3.61$$

Assume that Equation 2.3.61 is true and a contradiction will follow. Truth of 2.3.61 requires that $\lim_{|z|\to 1} \overline{\Gamma^{(1)}}(z)$ be an imaginary function at all z for which 2.3.61 holds. But

$$\lim_{|z|\to 1} \overline{\Gamma^{(1)}}(z) = \frac{1}{2\pi} \int_0^{\infty} \Gamma^{(1)}(\tau) e^{-j\omega\tau} d\tau \qquad 2.3.62$$

while the power distribution spectrum corresponding to $X^{(1)}(t)$ is

$$f^{(1)}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Gamma^{(1)}(\tau)\ e^{-j\omega\tau}\ d\tau \qquad 2.3.63$$

which can be written

$$f^{(1)}(\omega) = \frac{1}{2\pi}\int_{\infty}^{0} \Gamma^{(1)}(\tau)\ e^{-j\omega\tau}\ d\tau + \frac{1}{2\pi}\int_{0}^{\infty} \Gamma^{(1)}(\tau)\ e^{-j\omega\tau}\ d\tau \qquad 2.3.64$$

or due to the symmetry of $\Gamma^{(1)}(\tau)$

$$f^{(1)}(\omega) = \frac{1}{2\pi}\int_{0}^{\infty} \Gamma^{(1)}(\tau)\ e^{j\omega\tau} d\tau + \frac{1}{2\pi}\int_{0}^{\infty} \Gamma^{(1)}(\tau)\ e^{-j\omega\tau} d\tau \qquad 2.3.65$$

which must be real. If Equations 2.3.64 and 2.3.61 are both true, then

$$\lim_{|z|\to 1} \overline{\Gamma^{(1)}}(z) = 0 \qquad 2.3.65$$

and

$$f^{(1)}(\omega) = 0 \qquad 2.3.66$$

But by hypotheses

$$f^{(1)}(\omega) > 0$$

so Equation 2.3.61 cannot be true and p(z) cannot equal zero on the unit circle. Therefore Equation 2.3.58 is true in at least an annulus about the unit circle. The following theorem can now be stated.

Theorem 2.2

Let x(t), t=0, ± 1, ± 2, ..., be a time series whose spectral distribution is absolutely continuous except at a finite number of points where jump discontinuities exist. Furthermore, let the absolutely continuous portion of the process be such that it has an autoregressive representation (finite or infinite). Then the entire process has an autoregressive representation.

Proof:

The time series meeting these requirements is

$$x(t) = x^1(t) + x^2(t) \qquad 2.3.67$$

with $x^1(t)$ being regular and

$$x^2(t) = \sum_{n=1}^{M} A_n \operatorname{Cos} (\omega_n t+\phi_n) \qquad 2.3.68$$

According to Equation 2.3.58 $x(t)$ has autocovariance function

$$\Gamma_x(z) = \frac{1/w_0}{1 - \sum_{i=1}^{\infty} \frac{w_i}{w_0} z^{-i}} \qquad 2.3.69$$

but this is the autocovariance function of

$$x(t) = \sum_{i=1}^{\infty} \frac{\omega_i}{\omega_0} x(t-i\Delta) + N(t) \qquad 2.3.70$$

where $N(t) \sim$ uncorrelated id $(0, \frac{1}{0} - \sum_{n=1}^{\infty} \frac{w_n}{w_0} R_x(n))$ which is an infinite autoregressive process.

2.4 Determination of the Order of an Autoregressive Process

Estimation of the frequency of periodicity described in the previous section was based on the properties of the $\hat{\alpha}_{iM}$'s in the model

$$X_t = \sum_{i=1}^{M} \hat{\alpha}_{iM} X_{t-i} \qquad 2.4.1$$

The order of the autoregressive process, or the value of M, must be determined before the $\hat{\alpha}_{jM}$ can be evaluated. If a value of M that is too large is chosen, zeroes near the unit circle with small residue may be introduced. While the components corresponding to these zeroes may eventually be eliminated from the time series, an excessive amount of work will be introduced by including terms that should be set to zero. On the other hand too few terms can cause the periodic component to be missed completely and the residual sum of squares to be unduly high. A method of determining the correct order of the process is needed.

The first method for determining the order of the process is based on the residual sum of squares. Consider the process estimated by Equation 2.3.10,

$$x_T = \sum_{i=1}^{M} \hat{\alpha}_{iM} \, x_{T-i} \qquad 2.4.2$$

The residual sum of squares, based on N observations, is

$$\sigma_N^2(M) = \frac{(N-M)}{N-2M-1}\{R_{XX}(0) - \sum_{i=1}^{M} \hat{\alpha}_i \, R_{XX}(i)\} \qquad 2.4.3$$

Jenkins and Watts[4] suggests that the plot of $\hat{\sigma}_N^2$ (M) vs. M will flatten out or show a minimum at the proper value of M. This minimum may be rather flat and not very sharp. In this case the proper value of M is not indicated strongly enough. For that reason a more decisive procedure is discussed next.

A more sensitive criterion is obtained by determining the confidence internal for $\hat{\alpha}_M$, the estimate of the last coefficient in the fitted model. An approximate confidence region for the estimated vector $(\hat{\alpha})$ is

$$(\alpha-\hat{\alpha})' \; R \; (\alpha-\hat{\alpha}) \leq \frac{M}{N-2M-1} \; f_{M,N-2M-1}(1-\alpha) \; \hat{\sigma}_N^2(M) \qquad 2.4.4$$

where

$$(\alpha-\hat{\alpha})' = (\alpha_1-\hat{\alpha}_1,\ \alpha_2-\hat{\alpha}_2,\ \ldots,\ \alpha_{M-1}-\hat{\alpha}_{M-1},\ \alpha_M-\hat{\alpha}_M) \qquad 2.4.5$$

and

$$R = \begin{pmatrix} R_{XX}(0) & R_{XX}(1) & \ldots & R_{XX}(M-1) \\ R_{XX}(1) & R_{XX}(0) & \ldots & R_{XX}(M-2) \\ \vdots & & & \\ R_{XX}(M-1) & R_{XX}(M-2) & \ldots & R_{XX}(0) \end{pmatrix} \qquad 2.4.6$$

$f^{(1-\alpha)}_{M,N-2M-1}$ = Probability ($f \leq 1-\alpha$) where f has the F-distribution with M, 2N-M-1 degrees of freedom. Under the assumption that $\alpha_i = \hat{\alpha}_i$, for $i < M$, the confidence interval of α_M about zero is given by

$$\alpha_M \leq \left(\frac{M\, f^{(1-\alpha)}_{M,\ N-2M-1}}{R_{XX}(0)\,(M+ N-2M-1)}\ \hat{\sigma}^2_N\ (M) \right)^{\frac{1}{2}} \qquad 2.4.7$$

The value of $\hat{\alpha}_M$ and $\sigma^2_N(M)$ is computed for successive values of M and plotted along with the confidence interval according to 2.4.7. From this plot a value of M can be determined beyond which all values $\hat{\alpha}_M$ are within the confidence limits. This value of M should be selected as the order of the process.

CHAPTER III

TEST FOR PERIODIC COMPONENT

In the preceding chapter an estimator for the frequency of a periodic component in a time series was developed. Tests for the significance of the existence of such periodic components will be the subject of this section. Two tests will be presented. The first is a test of historical interest with some rather simplifying assumptions. The second will be a more exact test applicable to large samples. Properties of these tests will be discussed.

3.1 Test for Periodic Components in the Presence of Uncorrelated Noise

Suppose that a time series satisfying

$$X_t = n_t + \sum_{j=1}^{r} H_j \operatorname{Cos} (\omega_j t - \phi_j) \qquad 3.1.1$$

$$n_t \sim \text{nid}\ (0, \sigma_n^2)$$

is observed and that samples X_t, t=1, 2, ..., T are recorded. Quite often much information about a physical system can be deduced if H_j, ω_j and ϕ_j can be estimated. The preceding chapter presented an estimator for ω_j. Knowing ω_j, H_j and ϕ_j, or equivalently A_j and B_j where

$$A_j = H_j \operatorname{Cos} \phi_j \qquad 3.1.2$$

$$B_j = H_j \operatorname{Sin} \phi_j \qquad 3.1.3$$

can be estimated as Fourier Coefficients. These estimators are

$$A_T(\omega_j) = \frac{2}{T} \sum_{t=1}^{T} X_t \text{ Cos } \omega_j t \qquad 3.1.4$$

$$B_T(\omega_j) = \frac{2}{T} \sum_{t=1}^{T} X_t \text{ Sin } \omega_j t \qquad 3.1.5$$

and the corresponding sample spectrum estimator is

$$C_T(\omega_j) = [A_T^2(\omega_j) + B_T^2(\omega_j)] \qquad 3.1.6$$

Under the null hypotheses, H_0, $A_T(\omega_j)$ and $B_T(\omega_j)$ are both zero. This is equivalent to the hypotheses that $C_T(\omega_j)$ equals zero.

By the central limit theorem $A_T(\omega_j)$ and $B_T(\omega_j)$ are asymptotically normally distributed. Furthermore, they are assymptotically independent since they are orthogonal linear combinations of the X_T. Under the null hypotheses

$$\text{Var}[A_T(\omega_j)] = \text{Var}[B_T(\omega_j)] = \frac{2}{T} \sigma_n^2 \qquad 3.1.7$$

for large values of T. In any case, Equation 3.1.7 is true for ω_j an harmonic of the sampling frequency. That is, if

$$\omega_j = \frac{2\Pi j}{T\Delta t} \qquad 3.1.8$$

Δt = time between samples

Equation 3.1.7 is true for all T. If σ_n^2 is known then $2A_T^2(\omega_j)/T\sigma_n^2$ and $2B_T^2(\omega_j)/T\sigma_n^2$ are each χ_1^2 variables and are independent so that

$$\frac{C_T(\omega_j)}{\frac{2}{T}\sigma_n^2} = \frac{A_T^2(\omega_j) + B_T^2(\omega_j)}{\frac{2}{T}\sigma_n^2} \sim \chi_2^2 \qquad 3.1.9$$

is a chi square with two degrees of freedom. The null hypothesis, there is no periodic component at ω_j, would be rejected whenever

$$C_T(\omega_j) > 6(\frac{2}{T}\sigma_n^2) \qquad 3.1.10$$

at the 95% significance level.

There are two principle faults with this test. First it assumes that σ_n^2 is known which is rarely the case. In general σ_n^2 must be estimated from the data which means the estimate, say $\hat{\sigma}_n^2$, is a random variable. The test statistic $TC_T(\omega_j)/2\hat{\sigma}_n^2$ is no longer a χ_2^2 variable. The second criticism concerns the model assumed in 3.1.1. Rarely is the process being observed this simple. The continuous portion of the spectrum is often colored and contains peaks rather than being flat as assumed in 3.1.7. The estimator $\frac{2}{T}\hat{\sigma}_n^2$ underestimates these peaks so that the null hypothesis may be rejected when it is in fact true. The test presented in the next section does not suffer from these criticisms.

3.2 Test for Periodic Component in the Presence of Correlated Noise

Let samples X_t, t=1, 2, ..., T be taken from a process satisfying

$$X_t = \sum_{i=0}^{m} \alpha_i n_{t-i} + \sum_{i=1}^{r} (A_i \mathrm{Cos}\omega_i t + B_i \mathrm{Sin}\omega_i t) \qquad 3.2.1$$

$$n_t \sim \mathrm{iid}(0, \sigma_n^2)$$

and assume that T is quite large. In particular, let T be much larger than the largest lag for which it is assumed the samples are correlated in the absence of deterministic components. Under the null hypothesis, H_o: $A_i = B_i = 0$, let a value of lag, say v, be given such that

$$E[X_t X_{t+\tau}] << E\, X_t^2 \qquad 3.2.2$$

for all $\tau > v$. This requirement is not unusual when estimating the autocovariance function since, for large τ, there are fewer products of points to be summed over. Typically lags on the order of 3-5% of the record length are considered as maximum usable value. Of course, this would require T, the record length, to be 20-30 times the largest lag for which there is significant correlation in the process.

Under the foregoing conditions on T and v, divide the X_T into vectors (U_j) such that

$$\begin{aligned} (U_1') &= (X_1, \quad X_2, \quad \ldots, \quad X_\ell\ ) \qquad 3.2.3 \\ (U_2') &= (X_{\ell+1}, \quad X_{\ell+2}, \quad \ldots, \quad X_{2\ell}) \\ &\vdots \\ (U_k') &= (X_{(k-1)\ell+1}, X_{(k-1)\ell+2}, \ldots, X_{K\ell}) \end{aligned}$$

where $\ell \geq v$ in Equation 3.2.2 and

$$k\ell \leq T < (k+1)\ell \qquad 3.2.4$$

The matrix of the (U_j) vectors

$$(U) = (\ (U_1)(U_2)(U_3)\ \ldots\ (U_k)\ ) \qquad 3.2.5$$

is an ℓXk matrix made up of k independent vectors $(U_1) \ldots (U_k)$. Also under H_0 each of the (U_i) are multivariate normal with mean zero and covariance matrix (Σ). By the ergodic requirement, the (U_i) are identically distributed and by the assumption on the relationship between v, ℓ and T they may be considered independent. The maximum likelihood estimates for the mean vector and dispersion matrix are

$$(\bar{U}) = \frac{1}{k}\sum_{i=1}^{k} (U_i) \qquad 3.2.5$$

and

$$(S) = \frac{1}{k-1}\sum_{i=1}^{k} \left((U_i)-(\bar{U}) \right)\left((U_i)-(\bar{U}) \right)' \qquad 3.2.6$$

where $(\bar{U})$ is the mean vector estimate and (S) is the dispersion matrix estimate.

Rao[9] shows that the sample mean $(\bar{U})$ and dispersion matrix (S) are independent. Further linear combinations of the mean and dispersion matrix are independent also. These combinations and their distributions are

$$(L')(\bar{U}) \sim n\left((L')(u), \frac{1}{k}(L')(\Sigma)(L) \right) \qquad 3.2.7$$

where (u) and (Σ) are the population mean vector and dispersion matrix respectively of each of the (U_i), and

$$\frac{(L')(S)(L)}{(L')(\Sigma)(L)} \sim \chi^2_{k-1} \qquad 3.2.8$$

since

$$(S) \sim W_p(k-1, (\Sigma) \qquad 3.2.7$$

has a central Wishart distribution with parameters k-1 and (Σ).

Now suppose that an estimator for the frequency of periodicity (ω_j) has been found from the procedure in Chapter II. It is desired to test whether or not a periodic component actually exists at this frequency. Let the vectors (L) be given by

$$(L_1') = (\mathrm{Cos}\omega_j, \mathrm{Cos}2\omega_j, \ldots, \mathrm{Cos}\ell\omega_j) \qquad 3.2.10$$

and

$$(L_2') = (\mathrm{Sin}\omega_j, \mathrm{Sin}2\omega_j, \ldots, \mathrm{Sin}\ell\omega_j) \qquad 3.2.11$$

An estimator for A_j is

$$A_\ell(\omega_j) = \frac{2}{\ell} \sum_{h=1}^{\ell} \mathrm{Cos}\ \omega_j X_{i\ell+h} \qquad 3.2.12$$

where $0 \leq i \leq k-1$ with an improved estimator of A_j being

$$A_T(\omega_j) = \frac{2}{\ell k} \sum_{i=0}^{k} \sum_{h=1}^{\ell} \mathrm{Cos}\omega_j\ X_{i\ell+h} \qquad 3.2.13$$

But this is nothing more than

$$A_T(\omega_j) = \frac{2}{\ell} (L_1')(\overline{U}) \qquad 3.2.14$$

which has distribution

$$\frac{\ell}{2} A_T(\omega_j) \sim n(0, \frac{1}{k}(L_1')(\Sigma)(L_1) \qquad 3.2.15$$

and

$$\frac{\ell^2 k}{4} \quad \frac{A_T^2(\omega_j)}{(L')(\Sigma)(L_1)} \sim \chi_1^2 \qquad 3.2.16$$

From Equation 3.2.8

$$\frac{(L_1')(S)(L_1)}{(L_1')(\Sigma)(L_1)} \sim \chi_{k-1}^2 \qquad 3.2.17$$

so that

$$\frac{\ell^2 k^2 (k-1)}{4} \quad \frac{A_T^2(\omega_j)}{(L_1')(S)(L_1)} \sim F_{1,\ k-1} \qquad 3.2.18$$

due to the independence of the mean and dispersion matrix. Then H_0 i $A_j=0$ would be rejected if

$$\frac{A_T^2\ (\omega_j)}{(L_1')(S)(L_1)} > F_{1,k-1}(1-\alpha)\ \frac{4}{\ell^2 k(k-1)} \tag{3.2.19}$$

and likewise the hypothesis $B_j = 0$ would be rejected when

$$\frac{[(L_2)(\bar{U})]^2}{(L_2')(S)(L_2)} > F_{1,k-1}(1-\alpha)\ \frac{4}{k\ell^2(k-1)} \tag{3.2.20}$$

But

$$(L_1')(\Sigma)(L_1) \not\equiv_{\ell} (L_2')(\Sigma)(L_2) \tag{3.2.21}$$

since (Σ) is a covariance matrix of a stationary process. In particular

$$(\Sigma) = \begin{pmatrix} \sigma_0 & \sigma_1 & \cdots & \sigma_{\ell-1} \\ \sigma_1 & \sigma_0 & \cdots & \sigma_{\ell-2} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \sigma_{\ell-1} & \sigma_{\ell-2} & \cdots & \sigma_0 \end{pmatrix} \tag{3.2.22}$$

With this (Σ)

$$\frac{1}{\ell}(L_1')(\Sigma)(L_1) = \sum_{h=1}^{\ell}\sum_{i=1}^{\ell} \sigma_{|h-i|}\ \mathrm{Cos}h\omega_j\ \mathrm{Cos}i\omega_j \tag{3.2.23}$$

$$= \sum_{h=1}^{\ell} \sigma_0^2\ \mathrm{Cos}^2 h\omega_j + 2\sum_{h=2}^{\ell}\sum_{i=1}^{h-1} \sigma_{h-i}\mathrm{Cos}h\omega_j\ \mathrm{Cos}i\omega_j$$

and

$$\frac{1}{\ell}(L_2')(\Sigma)(L_2) = \sum_{h=1}^{\ell}\sum_{i=1}^{\ell} \sigma_{|h-i|}\mathrm{Sin}h\omega_j\ \mathrm{Sin}i\omega_j \tag{3.2.24}$$

$$= \sum_{h=1}^{\ell} \sigma_0^2\ \mathrm{Sin}^2 h\omega_j + 2\sum_{h=2}^{\ell}\sum_{i=1}^{h-1} \sigma_{h-i}\ \mathrm{Sin}h\omega_j\ \mathrm{Sin}i'\omega_j$$

Or

$$\frac{1}{\ell}(L_1')(\Sigma)(L_1) = \frac{1}{2\ell}\sum_{h=1}^{\ell} \sigma_0^2(1+\mathrm{Cos}\,2h\omega_j)+\frac{1}{\ell}\sum_{h=1}^{\ell-1} \sigma_h \mathrm{Cos}\,h\omega_j \sum_{i=h+1}^{\ell-1} \mathrm{Cos}\,(2i-h)\omega_j \tag{3.2.25}$$

while

$$\frac{1}{\ell}(L_2')(\Sigma)(L_2) = \frac{1}{2\ell}\sum_{h=1}^{\ell} \sigma_0^2(1-\mathrm{Cos}\,2h\omega_j)+\frac{1}{\ell}\sum_{h=1}^{\ell-1} \sigma_h \mathrm{Sin}\,h\omega_j \sum_{i=h+1}^{\ell-1} \mathrm{Sin}\,(2i-h)\omega_j \tag{3.2.26}$$

so that

$$\frac{1}{\ell}(L_1')(\Sigma)(\ell_1) \underset{\ell}{\rightarrow} \frac{1}{\ell}(L_2')(\Sigma)(L_2) \tag{3.2.27}$$

Let σ_1^2 denote $(L_1')(\Sigma)(L_1)$ and σ_2^2 be $(L_1')(\Sigma)(L_2)$. As ℓ increase

$$\sigma_1^2 \underset{\ell}{\rightarrow} \sigma_2^2 \underset{\ell}{\rightarrow} \sigma \tag{3.2.28}$$

and

$$\frac{[(L_1')(\bar{U})]^2}{\sigma_1^2} + \frac{[(L_2')(\bar{U})]^2}{\sigma_2^2} \underset{\ell}{\rightarrow} \frac{[(L_1')(\bar{U})]^2 + [(L_2')(\bar{U})]^2}{\sigma^2} \tag{3.2.29}$$

But

$$\frac{[(L_1')(\bar{U})]^2 + [(L_2')(\bar{U})]^2}{2\sigma^2} \sim \chi_2^2 \tag{3.2.30}$$

under the hypothesis H_0. Both $(L_1')(S)(L_1)$ and $(L_2')(S)(L_2)$ are estimators of σ^2 with equal variance so they can be averaged to obtain

$$\hat{\sigma}^2 = \frac{1}{2}\{[(L_1')(S)(L_1)] + [(L_2')(S)(L_2)]\} \tag{3.2.31}$$

which has expected value $k\sigma^2$ and k-1 degrees of freedom. Then

$$\frac{\hat{\sigma}^2}{K_{,}\sigma^2} \sim \chi_{k-1}^2 \tag{3.2.32}$$

Using this averaged estimator of σ^2 which is independent of the χ_2^2 in 3.2.30 an F statistic an be written as

$$\frac{k(k-1)\{[(L_1')(\bar{U})]^2+[(L_2')(\bar{U})]^2\}}{[(L_1')(S)(L_1)+(L_2')(S)(L_2)]} \sim F_{2,k-1} \qquad 3.2.33$$

and a test at the $(1-\alpha)$ level would be reject $H_0 : A_i = B_i = 0$ if

$$\frac{k(k-1)\{[(L_1')(\bar{U})]^2+[(L_2')(\bar{U})]^2\}}{[(L_1')(S)(L_1)+(L_2')(S)(L_2)]} > F_{2,k-1}(1-\alpha) \qquad 3.2.34$$

which is an improved test over the individual tests of 3.2.19 and 3.2.20.

3.3 Properties of the Test Statistics

In the previous sections two tests have been proposed for testing the hypothesis that a periodic component is not present at a given frequency. This section will review the properties of those tests and present some recommendations about when each of the tests should be applied. The first test was reject H_0 when (for the 95% level)

$$\left(\frac{2}{T}\sum_{t=1}^{T} X_t \text{Cos}\omega_j t\right)^2 + \left(\frac{2}{T}\sum_{t=1}^{T} X_t \text{Sin}\omega_j t\right)^2 > 6T\sigma_n^2 \qquad 3.3.1$$

where X_t was a sample of size T and σ_n^2 was assumed known. This test requires the indeterministic portion of the time series to be white noise with uniform spectrum. Seldom is this true since most noise is the result of some physical system which tends to filter or color it. At the peaks in the a continuous spectrum H_0 could be rejected when it is in fact true. The test is for deviations in the power spectra from some average level and H_0 is rejected when this deviation exceeds a given amount in the positive direction.

The second shortcoming of this test concerns the knowledge of σ_n^2. Seldom is this known a priori and usually must be estimated from the data. But if σ_n^2 is estimated from the data, that estimator becomes a random variable not necessarily independent of the numerator. The test statistic in 3.1.9 is no longer distributed chi-square. On the plus side the test is fairly simple to apply and while it, like most tests, improves with the amount of data available, it does not depend on a large T. In fact its asymptotic properties will not be improved if the peak being tested is a peak in the continuous spectrum.

The second test developed in Section 3.2 was to reject H_0 when (at the 95% level)

$$\frac{k(k-1)\{[(L_1')(\overline{U})]^2 + [(L_2')(\overline{U})]^2\}}{[(L_1')(S)(L_1) + (L_2')(S)(L_2)]} > F_{2,k-1}(0.95) \qquad 3.3.3$$

and is intended primarily as a large sample test. This test uses part of the data available to estimate the level of the spectrum at ω_j when there is no periodic component available, and the remainder to estimate the total sample spectra at ω_j. If only limited data is available the estimator for the sample spectra will be some what cruder than that in the previous test. However, as the record length increases the sample spectra estimator for the two tests should converge to the same value and the estimator of the continuous portion of the spectra will be much improved in the second test, in addition to being independent of the numerator. The F-test is more cumbersome to apply and will be primarily useful for computer applications.

Neither of these tests treat the problem of smoothing the side lobes of the sampling window. Parzen[5] along with numerous authors treat the problem of applying spectral windows in an attempt to filter out or suppress the frequency components due to sampling a continuous process. The tests presented in this paper use only a rectangular window although other windows could be applied.

In summary, Test 3.3.1 is simple to apply, is straight forward and should be used when limited data is available. Test 3.3.2 attempts to overcome some of the shortcomings of the simpler test and should be applied when there is a sufficiently long record, especially if a computer is available. The later test appears similar to the test given by Bartlett[2] on page 319. The denominators are the same when v in 3.3.2 is equal to n. However the numerators are different because the tests are for different hypothesis. Bartlett's test is for the null hypothesis that the spectrum is uniform forming a linear power distribution spectrum. Test 3.3.2 is for significant power at a given ω_i independent of the remainder of the spectrum. Since they test different hypothesis, numerical comparisons of the two tests is not applicable.

CHAPTER IV

DECOMPOSITION OF POWER DISTRIBUTION SPECTRUM

The preceding portion of this dissertation has been directed at showing conditions for a time series to have an autoregressive representation, the detection of hidden periodicities and the estimation of the periodic components. This chapter will demonstrate these principles through application to numerical examples and will show how this knowledge can be applied to improve time series modelling.

4.1 Numerical Examples

The first numerical example selected for analysis was a third order autoregressive process with a single tone added. Expressed mathematically this generating function was

$$X_t = X_t^{(1)} + X_t^{(2)} \qquad 4.1.1$$

with

$$X_t^{(1)} = 0.043\ X_{t-1}^{(1)} + 0.028025\ X_{t-2}^{(2)} + 0.1875\ X_{t-3}^{(2)} + n_t \qquad 4.1.2$$

$$X_t^{(2)} = C\ \mathrm{Cos}(2\Pi Ft + \phi) \qquad 4.1.3$$

where

$$n_t \sim \mathrm{nid}(0,1) \qquad 4.1.4$$

and

$$\phi \sim \text{uniformly } (-\Pi,\Pi) \qquad 4.1.5$$

The $X_t^{(1)}$ in Equation 4.1.2 is an approximate model for a humans' ability to point or aim a device in the presence of uncorrelated perturbations. Values of α_i are determined by the mass and response time of the system, amount of rate aiding, operator training, etc. but in general it has been possible to fit actual tracking data to third order autoregressive processes. The periodic term $X_t^{(2)}$ represents periodicity due to gear backlash, undamped control system responses, etc. In summary $X_t^{(1)}$ is the human response while $X_t^{(2)}$ is system response. Both terms may be of interest to the experimenter and so an accurate model is desired.

Another physical interpretation of 4.1.1 is the signal detection problem in a noise background. Examples of this type of problem are acoustical signals in underwater work and telemetry to and from space satellites. In either case the information is contained in either the aplitude or frequency, or both, of $X_t^{(2)}$, and consequently an accurate model of $X_t^{(2)}$ in the absence of $X_t^{(1)}$ is desired.

A time series of 1024 samples of model 4.1.1 were generated with a sampling interval (Δ) of 1. The first 48 samples of this time series are plotted in figure 4.1-a with

$$C = \sqrt{2}$$

$$F = 0.3875$$

Figure 4.1-b is a plot of the sample autocovariance function based on all 1024 samples of X_t calculated according to Equation 2.1.1. Equation 2.3.9 was used to estimate $\hat{\alpha}_{iM}$ for various values of M. Tabulated in figure 4.2-a are the estimated $\hat{\alpha}_{iM}$ for M ranging from 1 to 12. Corresponding to each M, the residual sum of squares

$$\hat{\sigma}_n^2 = R_{XX}(0) - \sum_{i=1}^{M} \hat{\alpha}_{iM} R_{XX}(i) \qquad 4.1.6$$

and is plotted in figure 4.2-b. Using the criteria of Chapter II for selecting M based on $\hat{\sigma}_n^2$ an M of 3 or at most 4 would be selected.

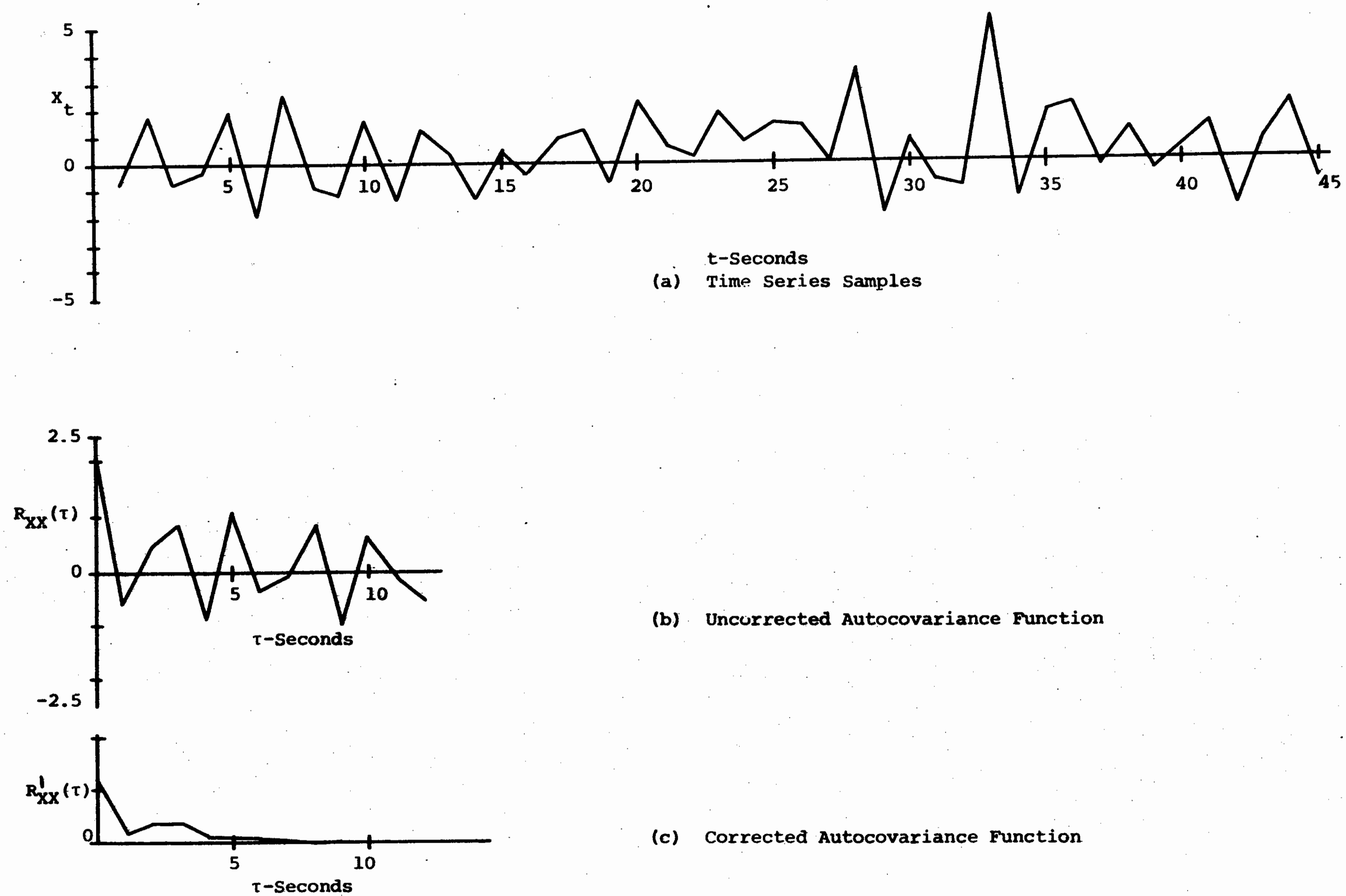
The periodic component would be missed if an M of 4 were selected. Figure 4.3-a summarizes some of the results of the search for periodicity for M between 1 and 12 inclusive. The first row labelled $|z_j|$ is the magnitude of the root of

$$1 - \sum_{i=1}^{M} \hat{\alpha}_i z^{-i} = 0 \qquad 4.1.7$$

nearest the unit circle. Θ is the corresponding angle of this root measured from the positive real axis while $\hat{F}$ is the frequency estimator from Chapter II based on this root. χ_2^2 is the test statistic from Chapter III for a chi-square test at $\hat{F}$ while $\hat{C}$ is the Fourier coefficient estimator of the amplitude of the periodic component at $\hat{F}$. For M of 4 or greater the error in estimating $\hat{F}$ was less than 1.5% but rejection of the null hypothesis based on the chi-square test requires about an order of magnitude reduction in this error. Accuracy in estimating $\hat{C}$ requires still another order of magnitude improvement in the error of estimating $\hat{F}$.

As M increases $|Z_j|$ appears to approach 1 while $\hat{F}$ appears to converge on F as seen in figure 4.4-b and a respectively. Figure 4.5 is a plot of χ_2^2 versus M and demonstrates a much more erractic behavior in the range of M of interest.

Conclusions from the analysis are that:



(a) Time Series Samples

(b) Uncorrected Autocovariance Function

(c) Corrected Autocovariance Function

Figure 4.1 Time Samples and Autocovariance Function

	M 1	2	3	4	5	6	7	8	9	10	11	12
$\hat{\alpha}_{1M}$	-.262	-.218	-.306	-.162	-.115	-.102	-.107	-.100	-.075	-.060	-.061	-.059
$\hat{\alpha}_{2M}$		.168	.282	.359	.238	.265	.275	.285	.274	0.267	.270	.279
$\hat{\alpha}_{3M}$			.523	.440	.378	.407	.393	.372	.348	.347	.346	.321
$\hat{\alpha}_{4M}$				-.273	-.245	-.224	-.198	-.175	-.147	-.144	-.144	-.136
$\hat{\alpha}_{5M}$					.172	.163	.180	.134	.097	.090	.090	.089
$\hat{\alpha}_{6M}$						-.076	-.083	-.115	-.036	-.026	-.027	-.030
$\hat{\alpha}_{7M}$							-.063	-.050	.009	-.016	-.014	-.003
$\hat{\alpha}_{8M}$								.117	.096	.076	.070	.053
$\hat{\alpha}_{9M}$									-.210	-.205	-.209	-.168
$\hat{\alpha}_{10M}$										.073	.074	.106
$\hat{\alpha}_{11M}$											.016	.009
$\hat{\alpha}_{12M}$												-.119

(a) Estimated Coefficients

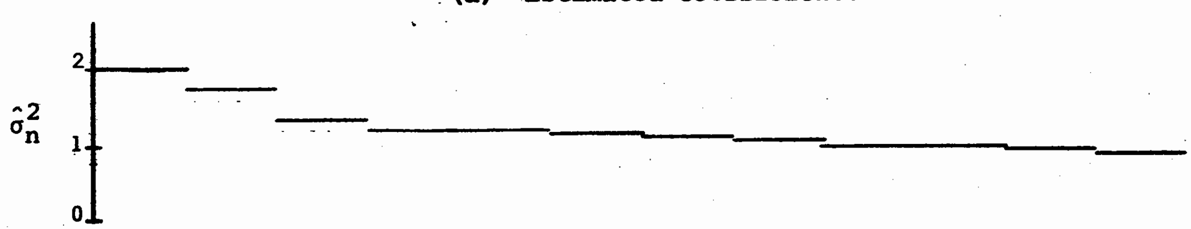


Figure 4.2 (b) Estimated Residual Sum of Squares
Estimated Autoregressive Models

	M											
	1	2	3	4	5	6	7	8	9	10	11	12
$\lvert z_j \rvert$	.262	.533	.819	.885	.912	.917	.922	.940	.962	.966	.966	.972
θ	180°	180°	0	137.5°	139.3°	140.2	139.5	139.0	139.5	139.8	139.7	139.4
$\hat{F}$	.5	.5	0	.382	.3870	.389	.3874	.3862	.3876	.3883	.3880	.3871
χ_2^2	.001	.001	.001	.755	89.3	.363	232.9	11.0	233.9	8.63	53.4	141.7
$\hat{C}$	.003	.003	.003	.080	.866	.055	1.399	.305	1.402	.269	.670	1.091

(a) Signal To Noise = 2

$\lvert z_j \rvert$	.028	.582	.482	.518	.514	.508	.802	.839	.897	.907	.907	.925
θ	0	0	134.0	136.7	139.7	140.5	138.1	138.0	139.8	140.5	139.6	138.90
$\hat{F}$	0	0	.3721	.3797	.3882	.3904	.3836	.3835	.3883	.3902	.3879	.3858
χ_2^2	.058	.058	.334	.700	20.89	.087	.003	.004	10.68	.862	63.75	2.127
$\hat{C}$	.018	.018	.044	.064	.347	.022	.004	.005	.248	.071	.606	.111

(b) Signal To Noise = $\frac{1}{2}$

Figure 4.3 Summary Autoregressive Analysis

1. The periodic component is detected for $M = 5$ and $M > 6$.
2. Errors of 1.4% or less in the prediction of F are achieved for $M \geq 4$.
3. Extremely accurate $\hat{F}$ is required for prediction of $\hat{C}$.

In an attempt to investigate the dependence of the results on the signal to noise ratio, the amplitude of the periodic component was reduced by a factor of 2. Figure 4.3-b summarizes the results of the analysis on the new time series. Errors in predicting F for M equal to or greater than 4 is increased to 2%. Figure 4-4 shows $\hat{F}$ to be converging to F much more slowly than in the previous case and correspondingly $|Z_j|$ approaches 1 more slowly. According to figure 4-5 the periodic component is detected in this case only for M equal to 5, 9 or 11. As one would suspect the signal detection difficulty has increased as the signal to noise power decreased.

A third example was analyzed. This model was

$$X_t = \sum_{n=1}^{5} C_n \operatorname{Cos}(2\Pi F_n t + \phi_n) \qquad 4.1.8$$

with $\phi_n \sim u\ (-\Pi, \Pi)$, and

n	C_n	F_n
1	1	0.08333...
2	1	0.16666...
3	1	0.25
4	1	0.3333...
5	1	0.416666...

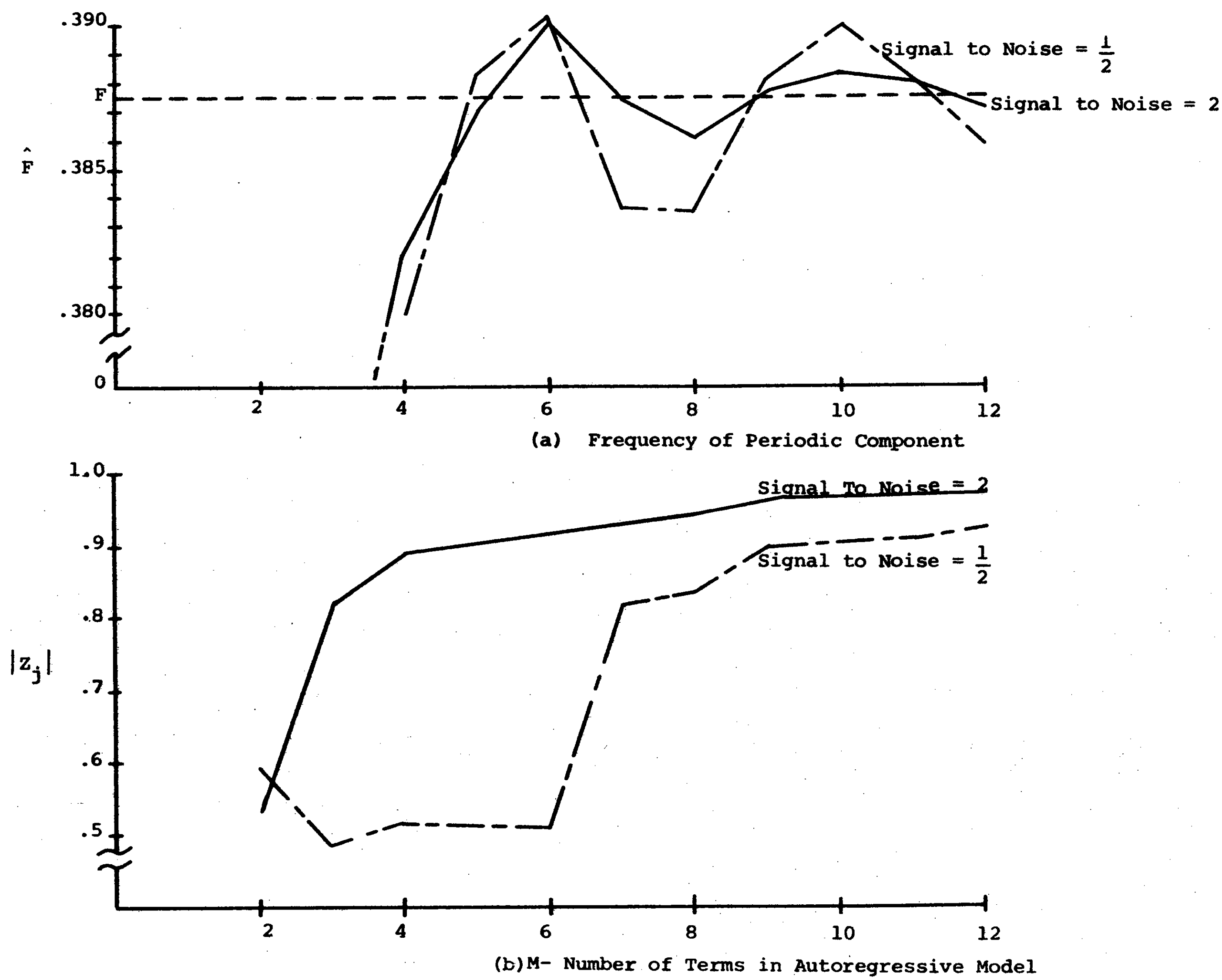


Figure 4.4 Estimated Periodic Component

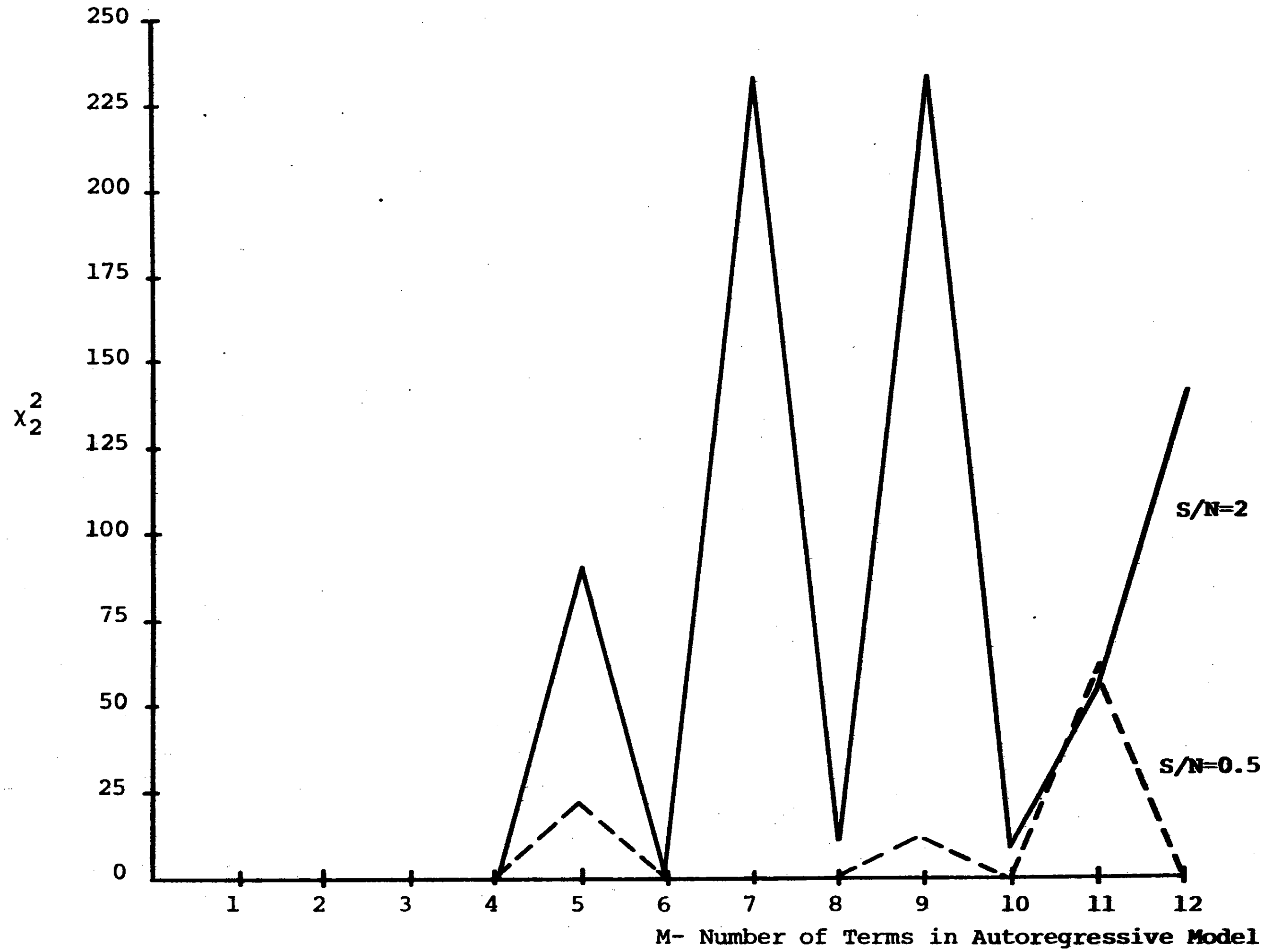


Figure 4.5 χ_2^2 — Test Statistic

Since each periodic component of a distinct frequency contributes two zeroes to Equation 4.1.6, 10 is the minimum value of M that can contain all information of the process. When a tenth order autoregressive process was fitted to the data the resulting $\hat{\alpha}$-vector was

$$(\hat{\alpha}_{iM})' = (-0.001, -0.988, 0, -0.99, -0.002, -0.99, 0.001, -0.988, -0.001, -0.985) \qquad 4.1.9$$

which lead to the following estimates.

n	$\hat{C}_n$	$\hat{F}_n$	χ^2_2
1	.989	.08326	100.38
2	.955	.16651	93.68
3	.971	.25013	96.78
4	.974	.33344	97.43
5	.990	.41674	100.69

By almost any realistic criteria, these components were seperated with a minimum number of terms. This is not unexpected since the data being analyzed were deterministic.

4.2 Decomposition of the Spectrum

Let X_t be given by

$$X_t = X_t^{(1)} + X_t^{(2)} \qquad 4.2.1$$

with

$$X_t^{(1)} = \sum_{i=1}^{\infty} \alpha_i \, X_{t-i}^{(1)} + n_t \qquad 4.2.2$$

a regular time series, and

$$X_t^{(2)} = \sum_{\ell=1}^{L} C_\ell \, \mathrm{Cos}(\omega_\ell \, t + \phi_\ell) \qquad 4.2.3$$

If $\Gamma_{XX}^{(1)}(v)$ is the autocovariance function of $X_t^{(1)}$, then

$$\Gamma_{XX}(v) = \Gamma_{XX}^{(1)}(v) + \sum_{\ell=1}^{L} C_\ell^2 \operatorname{Cos}(\omega_\ell v) \qquad 4.2.4$$

Having estimated $\hat{C}_\ell$ and $\hat{F}_\ell$ from

$$R_{XX}(v) = \frac{1}{T} \sum_{t=1}^{T-|v|} x_t x_{t+v} \qquad 4.2.5$$

one can estimate $\Gamma_{XX}^{(1)}(v)$ as

$$R_{XX}^{(1)}(v) = R_{XX}(v) - \sum_{\ell=1}^{L} \hat{C}_\ell \operatorname{Cos}(2\Pi \hat{F}_\ell v) \qquad 4.2.6$$

The component of error in $R_{XX}^{(1)}(v)$ due to error in estimating C_ℓ and F_ℓ is

$$\begin{aligned} \xi_\ell(v) &= C_\ell^2 \operatorname{Cos}(2\Pi F_\ell v) - \hat{C}_\ell^2 \operatorname{Cos}(2\Pi \hat{F}_\ell v) \qquad 4.2.7 \\ &= (C_\ell^2 - \hat{C}_\ell^2) \operatorname{Cos}(2\Pi F_\ell v) + \hat{C}_\ell^2 [\operatorname{Cos}(2\Pi F_\ell v) \\ &\qquad - \operatorname{Cos}(2\Pi \hat{F}_\ell v)] \end{aligned}$$

or

$$\xi_\ell(v) = (C_\ell^2 - \hat{C}_\ell^2) \operatorname{Cos}(2\Pi F_\ell v) - 2\hat{C}_\ell^2 \operatorname{Sin} 2\Pi v \frac{F_\ell - \hat{F}_\ell}{2} \operatorname{Sin} 2\Pi v \frac{F_\ell + \hat{F}_\ell}{2} \qquad 4.2.8$$

as $\hat{C}_\ell^2 \to C_\ell^2$ and $\hat{F}_\ell \to F_\ell$

$$\xi_\ell(v) \to \xi_{C_\ell} C_\ell \operatorname{Cos}(2\Pi F_\ell v) - 2\Pi v \xi_{F_\ell} C_\ell^2 \operatorname{Sin} 2\Pi v F_\ell \qquad 4.2.9$$

where

$$\xi_{C_\ell} = C_\ell - \hat{C}_\ell \qquad 4.2.10$$

and

$$\xi_{F_\ell} = F_\ell - \hat{F}_\ell \qquad 4.2.11$$

Let

$$S_C = C_\ell \operatorname{Cos} 2\Pi v F_\ell \qquad 4.2.12$$

and

$$S_F = -2\Pi v C_\ell^2 \operatorname{Sin} 2\Pi v F_\ell \qquad 4.2.13$$

so that

$$\xi_\ell(v) = S_C\, \xi_{C_\ell} + S_F\, \xi_{F_\ell} \qquad 4.2.14$$

and

$$\text{Var}\ [\xi_\ell(v)] = S_C^2\, \sigma_{C_\ell}^2 + S_F^2\, \sigma_{F_\ell}^2 + 2S_C S_F \text{Cov}(\xi_{C_\ell} \xi_{F_\ell}) \qquad 4.2.15$$

Equations 4.2.9 and 4.2.15 indicate the strong dependence of the residual error on the accuracy of estimating $\hat{F}_\ell$. The numerical examples in the previous section indicated that the covariance of ξ_{C_ℓ} and ξ_{F_ℓ} was quite high.

The first example from the previous section was corrected for the estimated periodic component. This correction was performed by subtracting the periodic component from the sample autocovariance function. For comparison purposes the corrected Sample Autocovariance function is plotted in figure 4.1-c. The periodic component so obvious in 4.1-b no longer appears to be present in 4.1-c. Analysis of the resulting Sample ACF bears this observation out.

The estimated coefficients were

$$(\hat{\alpha}) = (\ 0.058,\ 0.251,\ 0.217) \qquad 4.2.16$$

$$(\hat{\sigma}_n^2) = 1.016$$

when the correction made was for

$$\hat{F} = 0.38741 \qquad 4.2.17$$

$$\hat{C} = 1.3991 \qquad 4.2.18$$

which was the resultant periodic component from fitting a seventh order process to the original data. Comparing $(\hat{\alpha})$ to the real

$$(\alpha) = (0.043,\ 0.28025,\ 0.1875) \qquad 4.2.19$$

and

$$\sigma_n^2 = 1 \qquad 4.2.20$$

shows that the spectrum has been decomposed.

4.3 Frequency Estimate Improvement

The examples earlier in this chapter indicate that accuracies within 2% or better can be achieved when estimating the frequencies of periodic components with autoregressive models. Even with these accuracies, periodic components are sometime missed or the residual error still contains a high portion of the component. The minimum variance estimator occurs at the maximum of the power distribution spectrum. The estimator of Chapter II is near this maximum but it can be improved by iterative techniques. One such technique is discussed below.

The value of the peak in the spectrum is proportional to the square of C in Equation 4.1.3. $\hat{C}_\ell^2$ is a measure of the value of the estimated spectrum at $\hat{F}_\ell$. By choosing $\hat{F}_\ell$ to maximize $\hat{C}_\ell^2$ the estimate of F_ℓ can be improved. Since

$$\hat{C}_\ell^2 = \hat{A}_\ell^2 + \hat{B}_\ell^2 \qquad 4.3.1$$

$$\hat{A}_\ell = \left(\frac{2}{T}\sum_{t=1}^{T} X_t \operatorname{Cos}2\Pi\hat{F}_\ell t\right) \qquad 4.3.2$$

and

$$\hat{B}_\ell = \left(\frac{2}{T}\sum_{t=1}^{T} X_t \operatorname{Sin}2\Pi\hat{F}_\ell t\right) \qquad 4.3.3$$

$\hat{C}_\ell^2$ is differentiable with respect to $\hat{F}_\ell$. Taking this derivative gives

$$\frac{d\hat{C}_\ell^2}{dF_\ell} = 2\hat{A}_\ell \hat{A}_\ell' + 2\hat{B}_\ell \hat{B}_\ell' \qquad 4.3.4$$

and

$$\frac{d^2\hat{C}_\ell^2}{d\,\hat{F}_\ell^2} = 2[(\hat{A}_\ell')^2 + \hat{A}_\ell'' + (\hat{B}_\ell)^2 + \hat{B}_\ell''] \qquad 4.3.5$$

where

$$\hat{A}_\ell' = -\frac{4\Pi}{T}\sum_{t=1}^{T} tX_t \text{Sin} 2\Pi\hat{F}_\ell t \qquad 4.3.6$$

$$\hat{A}_\ell'' = \frac{-8\Pi^2}{T}\sum_{t=1}^{T} t^2X_t \text{Cos} 2\Pi\hat{F}_\ell t \qquad 4.3.7$$

$$\hat{B}_\ell' = \frac{4\Pi}{T}\sum_{t=1}^{T} tX_t \text{Cos} 2\Pi\hat{F}_\ell t \qquad 4.3.8$$

$$\hat{B}_\ell'' = \frac{-8\Pi^2}{T}\sum_{t=1}^{T} t^2X_t \text{Sin} 2\Pi\hat{F}_\ell t \qquad 4.3.9$$

The maximum of $\hat{C}_\ell^2$ occurs where

$$\frac{d\hat{C}_\ell^2}{d\hat{F}_\ell} = 0 \qquad 4.3.10$$

which can be estimated iteratively according to

$$(\hat{F}_\ell)_i = (\hat{F}_\ell)_{i-1} - \frac{\hat{A}_\ell\hat{A}_\ell' + \hat{B}_\ell \hat{B}_\ell'}{(\hat{A}_\ell')^2 + \hat{A}_\ell'' + (\hat{B}_\ell')^2 + \hat{B}_\ell''} \qquad 4.3.11$$

with $(\hat{F}_\ell)_0$ the frequency corresponding to the root of Equation 4.1.6 nearest the unit circle.

The estimator derived from Equation 4.3.11 will converge to a local maximum. To insure convergence to the proper maximum, the original estimate should be near that maximum. The estimator ($\hat{F}_{\ell}$) from above has been shown to possess this property.

LIST OF REFERENCES

1. Akutowicz, E.J., "On An Explicit Formula in Linear Least Squares Prediction," Mathematica Scandinavia, 5, (1957).
2. Bartlett, M.S., "On Theoretical Specification and Sampling Properties of Autocorrelated Time Series," Journal Royal Statistics Society, B8, 27, 1946.
3. Cox, D.R., and Miller, H.D., "The Theory of Stochastic Processes," John Wiley & Sons Inc., New York, 1968.
4. Jenkins, G.M. and Watts, D.G., "Spectral Analysis and Its Applications," Holden-Day, San Francisco, 1968.
5. Kromer, R.E., "Asymptotic Properties of the Autoregressive Spectral Estimators," PhD Dissertation, Stanford University, December 1969.
6. Kuo, B.C., "Analysis and Synthesis of Sampled-Data Control Systems," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.
7. Parzen, E., "Statistical Spectral Analysis (Single Channel Case) in 1968," TR-11 Stanford University, June 10, 1968.
8. Parzen, E. "Stochatic Processes," Holden-Day, San Francisco, August, 1967.
9. Rao, C.R., "Linear Statistical Inference and Its Applications," John Wiley & Sons, Inc., New York, New York, 1967.

10. Whittle, P., "Prediction and Regulation," D. Van Norstrand Co., Inc., Princeton, N.J., 1963.

11. Wold, H. (1938, 2nd ed., 1954) "The Analysis of Stationary Time Series" Almquist and Wicksell, Uppsala.

12. Zemanina, A., "Distribution Theory and Transform Analysis," McGraw-Hill, Inc., 1965.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*

Southern Methodist University
Department of Statistics
Dallas, Texas 75222

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

Unclassified

3. REPORT TITLE

Decomposition of Time Series into Deterministic and Indeterministic Components

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Technical Report

5. AUTHOR(S) *(First name, middle initial, last name)*

C.T. Brodnax

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
April 19, 1971	68	12

8a. CONTRACT OR GRANT NO.

N00014-68-A-0515

b. PROJECT NO.

NR 042-260

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

105

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale, its distribution is unlimited.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

Office of Naval Research
Department of the Navy
Washington, D.C.

13. ABSTRACT

Samples of a time series contain deterministic and indeterministic components. Frequency of occurrence of deterministic components can be estimated by fitting the data to an autoregression model and applying Z-transform theory of this model. A test is presented for testing the hypothesis that a periodic component is contained. By extraction of periodic components and repeating the procedure, the spectrum of the sampled data can be whitened. This report discusses this procedure, presents a computer program that implements this procedure and gives results of this program applied to test data.

DD FORM 1 NOV 65 1473

UNCLASSIFIED

Security Classification