NBER WORKING PAPER SERIES

INTERPRETING SPECTRAL ANALYSES IN TERMS OF TIME-DOMAIN MODELS

Robert F. Engle*

Working Paper No. 37

COMPUTER RESEARCH CENTER FOR ECONOMICS AND MANAGEMENT SCIENCE
National Bureau of Economic Research, Inc.
575 Technology Square
Cambridge, Massachusetts 02139

April 1974

Preliminary: not for quotation

*NBER Computer Research Center and Massachusetts Institute of Technology, Department of Economics. Research supported in part by National Science Foundation Grant GJ-1154X3 to the National Bureau of Economic Research, Inc.

Abstract

This paper derives relationships between frequency-domain and standard time-domain distributed-lag and autoregressive moving-average models. These relations are well known in the literature but are presented here in a pedogogic form in order to facilitate interpretation of spectral and cross-spectral analyses. In addition, the paper employs the conventions and discusses the estimation procedures used in TROLL. Some aspects of these estimation procedures are new and have not been discussed in the literature.

Contents

Figures

1. Introduction

Although there are many good reference works describing the details of spectral analysis, these do not focus on the fundamental relationships between spectral methods and such standard techniques as distributed lags, autoregressive processes or "naive" models. Granger (6)* is perhaps the easiest to read, while Jenkins and Watts (10) is the most comprehensive. Fishman (7) focuses on some economic estimation problems and Dhrymes (3, 4) extends this direction with somewhat more mathematics. Hannan (8) gives a very rigorous treatment of the whole area. Relatively short and simple early expositions of the theory and practice are in Jenkins (9) and Parzen (13) with a good application to economics in Nerlove (12). The book which is recommended as a companion to this system is Cooley, Lewis, and Welch (1) which is more application oriented and which describes, in Chapters 5 and 7, the basic concepts used in designing this system. Also see Cooley, Lewis, and Welch (2).

The purpose of the first section of this paper is to clarify these relationships and thereby enable a user without substantial knowledge of spectral theory to carefully and accurately interpret spectral analyses in terms of concepts which are familiar. When viewed in this way, spectral analysis provides a way of simply comparing a great many types of models so that the data can suggest which is most appropriate.

The second section of the paper will discuss the estimation procedure as implemented in the TROLL system from a conceptual

*Parenthesized numerals refer to entries in the Reference section, p. 23.

point of view. As the system uses the method of periodogram averaging which has become popular since the rediscovery of the fast Fourier transform, there are several issues which are not adequately covered in the literature. Furthermore, a general understanding of the estimation procedure is important for sensible interpretation of the results.

2. The Spectrum

Many data series can be considered successive chance observations over time called stochastic processes. Possibly, each observation is independent of the preceeding ones. However, for most applications, there is some suspected dependence between the observations. Both spectral analysis (frequency domain) and the more familiar time domain analysis are ways to characterize this dependence. High correlations between neighboring observations or seasonal components might be important forms of this dependence. Once we have characterized the stochastic process we may be able to forecast its values, improve the efficiency of a regression where this is the disturbance, or infer some information about the economic model which produced such a variable.

Both frequency domain and time domain analysis begin with stochastic processes which are covariance stationary. This means that the covariance between an observation now and one a few periods later depends only on the time interval, not the dates themselves. Mathematically this can be expressed as

$$(1) \quad \gamma(s) = E(x_{t+s} - \mu)\,(x_t - \mu)$$

where γ is the autocovariance function and μ is the mean. The important assumption is that neither depend upon t. While this assumption may seem strong, it is only because of this condition that information from the past can be used to describe the present or future behavior.

Many economic time series appear to violate this assumption, particularly those with pronounced trends. It is generally possible, however, to create an approximately stationary series by taking first

differences, or extracting a trend, thus leaving the series with a constant mean of zero. There may also be trends in variance which can often be removed by first taking logs of the series.

In the time domain the most common models are the autoregressive moving average models (ARMA). These may be purely autoregressive, purely moving average, or mixed.

A pth order autoregressive and a qth order moving average are shown in equations (2) and (3), respectively while (4) is an ARMA (p,q).

(2) $x_t = a_1 x_{t-1} + a_2 x_{t-2} + \ldots + a_p x_{t-p} + e_t$

(2') $A(L)\, x_t = e$

(3) $x_t = b_1 e_{t-1} + b_2 e_{t-2} + \ldots + b_q e_{t-q} + e_t$

(3') $x = B(L)e$

(4) $x_t = a_1 x_{t-1} + \ldots + a_p x_{t-p} + b_1 e_{t-1} + \ldots + b_q e_{t-q} + e_t$;

(4') $A(L)x = B(L)e$

In these equations e is a series of independent random variables where e_t is independent of x_{t-i} for all i greater than zero; L is the lag operator and $A(L)$ and $B(L)$ are polynomials. These classifications are not unique since one type of process can, in general, be transformed into one of the others. Nevertheless, they provide useful, simple models of time series which can be tested with data or used for analysis.

The spectrum provides another way of characterizing time series. In this case we think of a series as being made up of a great number of sine and cosine waves of different frequencies which have just the right (random) amplitudes to make up the original series. Thus the list of how much of each frequency component was necessary is also a full description of the time series. The spectrum is a plot of the squared amplitude of each component against the frequency of that component. It is continuous and always greater than zero as long as we have no deterministic elements (that is no exactly repeating components, or components which can be predicted exactly on the basis of the past).

Although this seems like an unusual way to think of a stochastic process, it is very general since any covariance stationary process can be uniquely described in this fashion.

The spectral density function is defined as the Fourier transform of the autocovariance function

$$(5)\quad f(\theta) = \gamma(0) + 2\sum_{s=1}^{\infty}\gamma(s)\cos(2\pi\theta s) = \sum_{s=-\infty}^{\infty}\gamma(s)e^{-2\pi i\theta s} \qquad 0\leq\theta\leq 1$$

when $i = \sqrt{-1}$, $e^{i\theta} = \cos(\theta) + i\sin(\theta)$, and the last equality follows from $\gamma(s) = \gamma(-s)$. There are several important features of this definition. First, although we have used complex notation, the spectrum is real valued since all the imaginary sine terms cancel exactly. Second, since the cosine is symmetric $f(\theta) = f(1-\theta)$, only the frequencies from 0 to 1/2 are needed to describe the spectrum. Third, if we integrate equation (5) from 0 to 1, we will find that the area under the spectrum is equal to $\gamma(0)$, the variance. We can show that the spectrum is a decomposition of the variance into the components contributed by each frequency. Fourth, since f is measured in cycles per period, it appears that we have no component from less than one cycle every two periods, (the Nyquist frequency). The reason for this becomes clear upon reflection. When we observe monthly data, weekly fluctuations will be undistinguishable from longer oscillations which have the same value at the moment the observation is taken. The weekly component will therefore be counted with these lower frequencies.

The important implication of the definition, which we will not prove, is that each spectral point can be interpreted as the variance of the component with that frequency so that the spectrum is always positive. It is not difficult to give an intuitive proof that the value of the spectrum at each frequency is just the squared amplitude of a sine wave of that frequency and therefore non-negative, but the basic theorem, called the spectral representation theorem, is quite difficult. From this result we have the standard interpretation of the spectrum, as a decomposition of the variance by frequency.

To clarify the interpretation of a spectrum and help with the notion of frequency components, let us interpret the spectrum in Figure 1 which has been estimated from quarterly data.

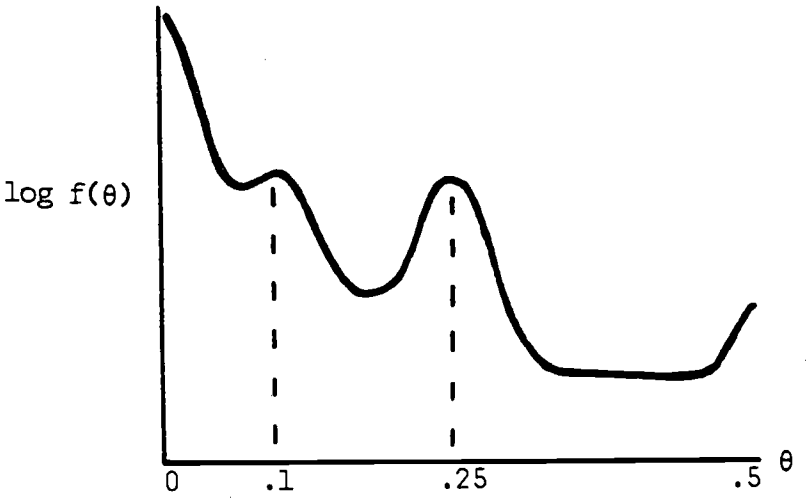


FIGURE 1. TYPICAL SPECTRUM

We only look at the first half of the spectrum and therefore the highest frequency oscillation we can distinguish is 0.5 cycle per period. At this frequency, it takes two quarters to complete a cycle so there are two cycles per year. There is a peak at 0.25 cycles which corresponds to a four-quarter, or annual cycle. This is most likely a seasonal component. Similarly, the peak at 0.5 also probably indicates a seasonal component since it has an even number of cycles per year. The peak at 0.1 corresponds to a two and a half year oscillation. This might be a business cycle and therefore economically interesting if it is significantly above its neighboring points. Generally, economic time series show behavior much like that of Figure 1.

In this paper we wish to emphasize the relationship between these concepts of frequency domain analysis and the more conventional time domain analysis. The first order serial correlation coefficient is easily calculated in the time domain and is generally large and positive for economic time series. We can translate this finding into the frequency domain as well. If we multiply the spectrum by

$\cos(2\pi\theta)$ and integrate, we obtain from equation (5) just $\gamma(1)$, the first order serial covariance. Roughly, this amounts to multiplying low frequencies by a positive number, high frequencies by a negative number, and adding. If the result is positive, there is positive first order serial correlation. Thus data series with generally downward sloping spectra have positive first order serial correlation while those with upward sloping spectra have negative serial correlation. Very important is the observation that spectra which are roughly symmetric about 0.25 will show no first order serial correlation.

A useful application of this analysis is found in interpretation of regression results. The assumption of no serial correlation in the disturbance is equivalent to the assumption that its spectrum is constant. The Durbin-Watson statistic gives us a test against the possibility that there is first order serial correlation. However, we now recognize this as a test against a general slope of the spectrum of the disturbance, whereas we would like to test against all forms of variation. In particular, notice that if the seasonality in Figure 1 were more severe, the spectrum might easily have no first order serial correlation but be far from constant. Durbin (5) formulated such a test based upon the spectrum of the residuals which is easily computed within TROLL. In general, examination of the residual spectrum gives very useful information about the validity of the regression assumptions.

The link between time domain and frequency domain is completed by a derivation of the spectrum corresponding to the ARMA models of equations (2)-(4). The basic result is quite simple but will be established in the appendix.

LEMMA 1: If x is a stochastic process generated by the model

$$A(L)x_t = B(L)\, e_t$$

where e_t is a series of independent identically distributed random variables with variance σ^2, and the polynomial $A(L)$ has all roots outside the unit circle, then the spectrum of x is given by

$$(6)\quad f_x(\theta) = \sigma^2|B(z)|^2/|A(z)|^2, \quad z = e^{(-2\pi i\theta)}$$

Notice that z is a complex function of θ.

Several examples should help to illustrate the usefulness of this result. First, notice that the spectrum of the very simple (white noise) process which has no time dependence, is just a constant. It has equal contributions from all frequencies.

Now consider the first order moving average process with parameter p, $x_t = e_t + pe_{t-1}$. From equation (6) the spectrum of x is

$$f_x(\theta) = |1 + pe^{-2\pi i\theta}|^2\sigma^2 = \{1 + p^2 + 2p\cos(2\pi\theta)\}\sigma^2$$

Evaluating this for θ in the range $(0, \frac{1}{2})$, gives a smooth spectrum which begins at $(1 + p)^2$ and ends at $(1 - p)^2$. If p is positive, this has the typical spectral shape which is common to most economic time series, and which implies a positive serial correlation coefficient, $p/(1+p^2)$. The first order autoregressive case is very similar but gives a somewhat steeper spectrum at low frequencies.

A very simple autoregressive model which captures the behavior of purely seasonal stochastic processes for monthly data is

$$x_t = px_{t-12} + e_t$$

From equation (6) the spectrum of this seasonal process is given by

$$f_x(\theta) = \sigma^2/|1-pe^{-24\pi i\theta}|^2 = \sigma^2/(1+p^2 - 2p\cos(24\pi\theta))$$

which is plotted in Figure 2. There are peaks at all the harmonic frequencies: $\theta = 1/12, 2/12, 3/12, 4/12, 5/12, 6/12$, and all are equally important.

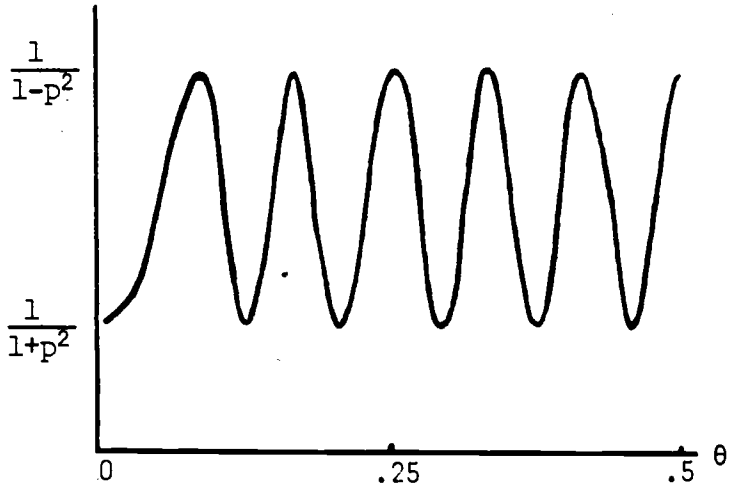


FIGURE 2. SPECTRUM OF PURE SEASONAL

3. The Cross Spectrum

The techniques used above can also be used to describe the relations between two jointly covariance-stationary time series. Both the individual behavior and the interrelations can be decomposed into basic sinusoidal elements.

First, we define the cross covariance function which is a direct analogue of the autocovariance function. For two series with mean zero this is simply defined as:

$$(7) \quad \gamma_{xy}(s) = E(x_{t+s}\, y_t)$$

Again, notice that it does not depend on t. The cross spectrum is similarly defined as:

$$(8) \quad f_{xy}(\theta) = \sum_{s=-\infty}^{\infty} \gamma_{xy}(s) e^{-2\pi i \theta s}$$

Because γ_{xy} is no longer symmetric the cross spectrum is not a real valued function of θ but rather a complex valued function.

Although the cross spectrum summarizes all the information we need, we cannot plot it directly. Instead, we commonly look at what are called "coherence squared", "gain" (or "transfer"), and "phase". We will define these measures here and give a rather extended interpretation below, connecting these concepts with the ideas of distributed lag regression models.

The coherence squared (COH) is like a correlation coefficient and is defined as:

$$(9) \quad \mathrm{COH}(\theta) = |f_{xy}(\theta)|^2 / f_x(\theta) f_y(\theta)$$

which is clearly between 0 and 1.

The gain (G) indicates how much the spectrum of x has been amplified to approximate that component of y. It is therefore like a regression coefficient.

$$(10) \quad G_{xy}(\theta) = |f_{xy}(\theta)| / f_x(\theta)$$

This expression can clearly never be negative. However, if it is small, it indicates that at frequency θ, x has little effect on y.

The phase (PH) is a measure of the timing between the series. It is measured in the fractions of a cycle that y lags x.

$$(11)\quad PH(\theta) = \frac{1}{2\pi} \arctan\left(\frac{-Im(f_{xy}(\theta))}{Re(f_{xy}(\theta))}\right)$$

where Im and Re are the imaginary and real parts of the cross spectrum.* There is a natural ambiguity about the phase since adding or subtracting 1 whole cycle from an angle will not change its tangent. Thus the phase is known only up to adding or subtracting an integer and therefore even the lead-lag relation is not known for sure. The plot of the phase is designed to emphasize this fact. It is possible to combine the phase and the gain in a simple expression

$$(12)\quad f_{xy}(\theta)/f_x(\theta) = G_{xy}(\theta)e^{-2\pi i PH(\theta)}$$

Two other potentially useful measures of the cross spectrum are its amplitude, which is merely its absolute value, and its time lag. The latter describes the phase in terms of the number of periods y lags x rather than the fraction of a cycle. Although this seems like a useful measure, the natural ambiguity of the phase also makes the time lag ambiguous and difficult to interpret. This may not be the case at low frequencies, where these difficulties are less likely to be important.

A natural and very general way for economists to think about the relations between two time series is in terms of a bivariate distributed lag model, such as

$$(13)\quad y_t = \sum_{j=p}^{-p} w_j x_{t-j} + u_t.$$

This is often rewritten in terms of the lag operator L as

$$(14)\quad y_t = \sum_{j=p}^{-p} w_j L^j x_t + u_t = w(L)x_t + u_t$$

where we have, for generality, allowed for leads as well as lags and interpreted L^{-1} as a lead operator. Using the same techniques required for equations (5) and (6), we can establish frequency domain interpretations of equation (14).

*The appropriate quadrant for PH is chosen on the basis of the signs of the real and imaginary components of the cross spectrum.

LEMMA 2: If y is generated by a distributed lag model

$$y = w(L)\, x + u$$

where x and u are uncorrelated covariance stationary processes, then

(15) $f_{xy}(\theta) = w(z^{-1})\, f_x(\theta)$, and

(16) $f_y(\theta) = |w(z)|^2 f_x(\theta) + f_u(\theta)$

where $z = e^{(-2\pi i\theta)}$.

We notice from (16) that the variance of y is broken into two parts; one which is the variation due to x modified by the lag distribution and the other due to the disturbance. From (15) we see that f_{xy}/f_x is an estimator of $w(z)$ which is just a function of the lag coefficients. In fact, it is the Fourier transform of the lag distribution; and therefore, once $w(z)$ is known, all the lag coefficients can be found by merely taking the inverse Fourier transform. This is the basis of a very useful type of distributed lag estimation which is often called Hannan's inefficient method.* This is available in TROLL.

Suppose that we now consider running a regression of one component of y against the same frequency component of x. The regression coefficient would be the ratio of the covariance of x and y to the variance of x. In our notation this would be just $f_{xy}(\theta)/f_x(\theta)$. The R-squared of this regression is one minus the unexplained variance over the total variance. Substituting this into equation (9) we see that the coherence squared is just the R-squared of this regression. Similarly, from equation (12) we see a very intimate relation between this regression and the gain and phase estimators. Let us explore this further.

If the model is (1) really a static model, which means that only w_0 is non-zero, or (2) a very simple dynamic model (often called a delay model), in which only one w_j is non-zero, then

(17) $w(z^{-1}) = w_j e^{+2\pi i\theta j}$

*It is inefficient because it does not use the properties of the disturbance to construct an estimator with the smallest possible variance.

and therefore the gain should be constant and equal to w_j. For more complicated models we would not expect the gain to be constant. At low frequencies, z is close to unity and therefore the gain should be close to the long run multiplier. At other frequencies the gain is the effective multiplier which a sine wave of that frequency would experience.

In the figures at the end of this paper, the gain for several common simple distributed lag models are computed and plotted. Two basic patterns appear. When the gain is a declining function of frequency, the lag distribution tends to emphasize low frequencies and eliminate, or smooth out, higher frequencies. This is the typical pattern for moving average type of lag distributions. The other pattern emphasizes high frequencies and it arises from lag distributions which depend on the rate of change of the right hand variable such as first differences or accelerator models. Of course many more complicated models are possible, yielding a great variety of patterns.

The phase estimator incorporates all the information about leads and lags. If the model is a static model, there should be no phase shift. If there is a simple delay, the phase is just a straight line with slope equal to the number of periods of the delay. For more complicated models, the slope of the phase near zero frequency can be shown to be equal to the average or mean lag of the lag distribution (as long as the gain is non-zero). For several examples of this see the figures at the end of the chapter. Notice the dashed lines which indicate the ambiguity between leads and lags. This is in exactly the same form as the computer printout.

The series of gain and phase plots (called Bode plots) presented at the end of this paper will never be exactly the same as any estimated from data. Nevertheless, there may be one close enough that it is possible to infer a form of time domain lag distribution. If not, two facilities are available. First, the definitions of gain and phase imply that if the lag distribution is a product of two lag distributions, the composite gain will be the product of the individual gains and the composite phase will be the sum of the two phases.

Second, you can construct any other diagram for comparison by means of a TROLL program. Thus by trial and error it may be possible to understand what type of lag distribution would have a spectrum similar to that estimated. Similarly, it is possible to know what gains and phase are implied by a particular lag distribution.

4. Spectrum Estimation

There are several distinct methods for estimating spectra and cross-spectra. The advantages and disadvantages of each have been extensively discussed. In particular, see Cooley, Lewis, and Welch (1) and Parzen (14). Since the rediscovery of the fast Fourier transform, computational considerations suggest that periodogram averaging may be the most efficient method for spectrum estimation. In addition, it is conceptually simplest and leads to great versatility in the estimation procedures. Finally, the usefulness of the periodogram in regression and various test procedures makes it sensible to compute this as a first step. See also Jones (11) and Tick (15).

The periodogram is defined as the square of the absolute value of the Fourier transform of the series at each frequency, all divided by m, the number of observations. The formula for the periodogram is

$$(18)\quad I_x(\theta_j) = \frac{1}{m}\left|\sum_{t=0}^{m-1} x_t e^{-2\pi i t\theta_j}\right|^2$$

where $\theta_j = j/m$ and $j=0,1,2,\ldots,m$. This quantity is an estimator of the spectrum, but it is not a very good one. The expected value of the periodogram is

$$(19)\quad E(I_x(\theta_j)) = \sum_{v=-m+1}^{m-1} \frac{(m-|v|)}{m}\,\gamma(v)e^{-2\pi i v\theta_j}$$

For large values of m this estimator is an unbiased estimator of the spectrum, since $\gamma(v)$ is small for large v. Unfortunately, however, it is not a consistent estimator since the variance does not decrease as the sample approaches infinity. In fact, the spectral estimator at each frequency is approximately proportional to a chi squared random variable with two degrees of freedom, regardless of the number

of observations. An intuitive explanation for this unusual circumstance is that as the sample becomes larger, more and more frequency points are estimated rather than obtaining better estimates of a fixed number of parameters. This explanation also suggests the solution. Since neighboring points are independent (if the original series was normal), the average of a few should give a better estimate of the spectrum in that neighborhood (assuming that it is not changing too rapidly). Thus we must use smoothing procedures to obtain consistent spectrum estimators.

5. Smoothing Windows and Confidence Intervals

Two averaging or smoothing procedures, called "windows", are currently available in this system. A rectangular moving average gives the minimum variance for smoothing over a flat spectrum using only a certain number of points. However, when there are peaks in the spectrum, the rectangular window will lead to considerable bias and broadening of the peaks. An alternative window is a triangular window which gives the spectrum a much smoother appearance and is probably better at describing the shape of peaks.

Clearly, the width of the window is an important parameter in the estimation. The wider the window, the smaller is the variance of the resulting estimate; yet, the wider the window, the more serious may be the bias of smoothing over non-smooth portions of the spectrum. Two measures of width are used to describe the windows in the TROLL spectral package, the bandwidth and the range. The bandwidth is the half-power width of the window. It is measured in frequency units, i.e., it is a fractional number of cycles per period. If, for example, the bandwidth is specified as 0.1, there will be five separate "bands" since the frequencies range from 0 to 0.5. For many purposes, spectral estimates separated by more than one bandwidth are considered to be independent.

The second measure is the range. This is merely the number of spectral points used in each moving average; it therefore gives the separation between which two points are known to be completely

independent. If the effective sample is 200 observations (implying 100 points in the spectrum) and the range is 20, there will be five separate window widths in the estimation. A sensible value for the range is $\sqrt{m}$, where m is the number of observations.

Near the endpoints of the spectrum, the smoothing procedures must be modified. The choice followed here is to decrease the range so that the window does not overlap the endpoints. Because the variance increases as the window becomes narrower, the variance increases markedly at very low or very high frequencies and one must be very cautious in interpreting low frequency peaks or troughs. The user who wishes to construct his own window or to modify the endpoint procedure can easily do so within the structure of the system.

The resulting spectral estimator is approximately proportional to another chi squared random variable, this time with more degrees of freedom. The equivalent degrees of freedom are just equal to

(20) E.D.F. = $B \cdot m$

where B is the bandwidth. This allows us to compute a confidence interval for the spectrum. On the spectral plot a 95% confidence interval is given for each frequency separately.

Estimates of the cross spectrum are accomplished in exactly the same manner. The finite Fourier transform of one series is multiplied times the complex conjugate of the Fourier transform of the other to form the cross periodogram. The real and complex parts of this are then smoothed individually, just as for the periodogram. The sampling distributions for the various measures derived from the cross spectrum also depend only on the equivalent degrees of freedom of the estimate. With the coherence plot, the critical point for a 5% test of the hypothesis of zero coherence is given. Approximate 50% confidence intervals for the gain are plotted with the output. These depend on the sample coherence; where the coherence is small, the confidence interval is large.

6. Prewhitening and Fourier Transforms

When using a wide window, peaks tend to be spread out. For many series we know *a priori* where these peaks will be, either because

the series is typical in having strong low frequencies or because it has important seasonality. In these cases "prewhitening" is often recommended. This amounts to dividing the raw periodogram by the expected or typical shape, smoothing this "prewhitened" periodogram which no longer has the large peaks, and then "recoloring" by multiplying by the typical spectral shape. A seasonal and non-seasonal version of the prewhitening filter are available, but the user can easily construct his own. Prewhitening can be done in connection with either spectrum or cross spectrum estimation.

A second characteristic which is likely to make the smoothing procedure badly biased in cross spectral estimation is misalignment of the series. When one series lags another by several periods, there is a peak in the cross covariance function which is not at zero. This leads to a regular oscillation in the amplitude of the cross periodogram. Smoothing this will obscure this particular bit of information as well as distorting other results. The recommended procedure is to first divide the cross periodogram by an aligning series, smooth the cross periodogram, and then remultiply it by the aligning series. The program to construct the alignment series first computes the inverse Fourier transform of the cross periodogram which is exactly the cross covariance function. This could have been computed from the data directly, but such a method is apparently inferior to the computation of Fourier and inverse Fourier transforms. Searching the cross covariance function for the maximum yields the information needed to construct the alignment series. If this procedure were applied to the estimate of the spectrum or a cross spectrum which was already aligned, the maximum covariance would be the zeroth estimate and thus the alignment series would be unity and would have no effect.

The Fourier transform algorithm used in these computations is the Cooley-Tukey fast Fourier transform. In its basic form it expects a series with 2^n elements and thus each series is padded out to this length with zeros. The number of spectral points estimated is therefore 2^{n-1} which are evenly spaced between the frequencies 0 and

1/2 cycle per basic time unit of the data. If the user wishes to construct the spectrum at particular points or wishes not to pad with zeros, he is provided an option of choosing another integer, r, so that the series is padded to $N=2^s r$. Choosing $r=3$, for example, would insure factors of 12 which would be required in order to have exact seasonal points with monthly data. If unspecified by the user, r receives the default value of 1.*

7. Cases of Interest

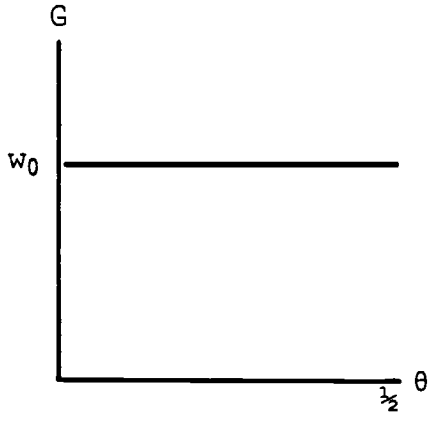
1. Simple Static Model

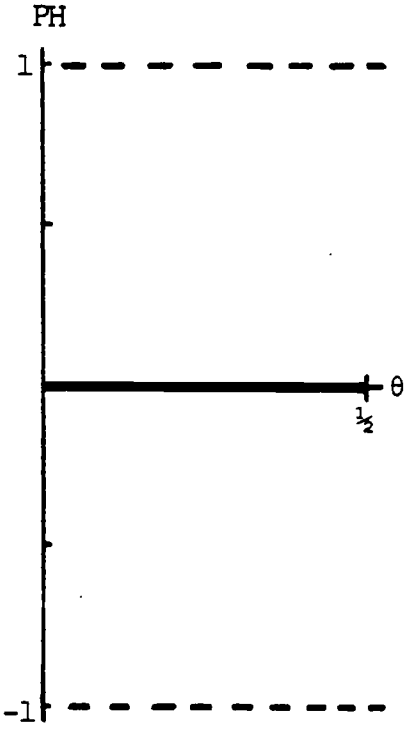
$w(L) = w_0$

$w(z^{-1}) = w_0$

$G(\theta) = w_0$

$PH(\theta) = 0$





*All output and statistics are corrected so that they depend on m, not N.

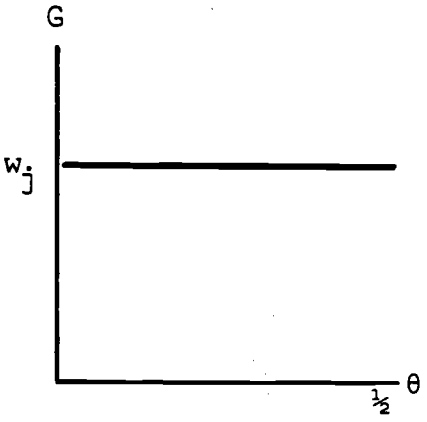
2. Simple Delay

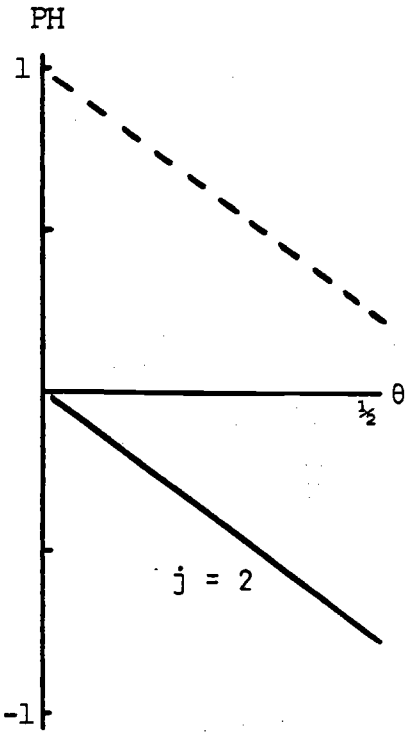
$w(\theta) \quad w_j L^j$

$w(z^{-1}) = w_j e^{-2\pi i j\theta}$

$G(\theta) = w_j$

$PH(\theta) = -j\theta$





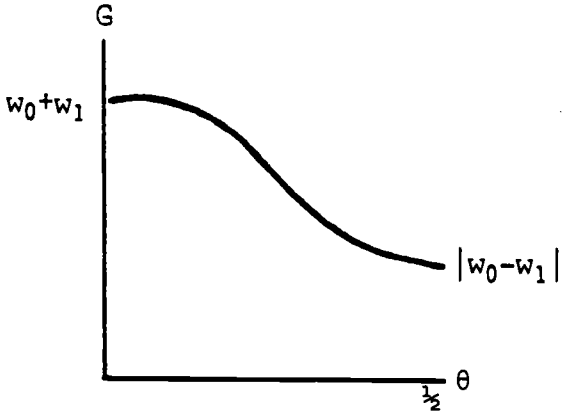
3. One Period Lag

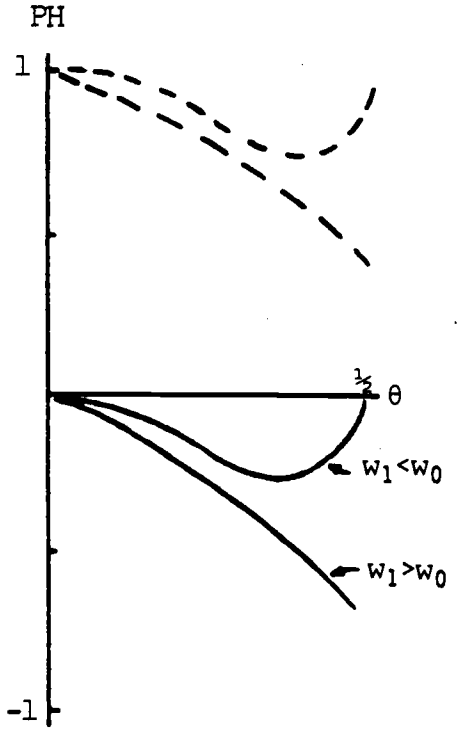
$w(L) = w_0+w_1L \qquad w_0, w_1>0$

$w(z^{-1}) = w_0+w_1e^{2\pi i\theta}$

$G(\theta) = \sqrt{w_0^2+w_1^2+2w_1w_0\cos 2\pi\theta}$

$PH(\theta) = \arctan\left(-\frac{w_1\sin 2\pi\theta}{w_0+w_1\cos 2\pi\theta}\right)$





4. Geometric Lag

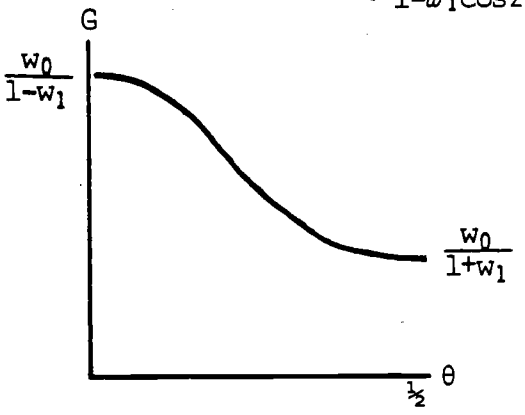
$$w(L) = \frac{w_0}{1-w_1 L} = w_0\ (1+w_1 L+w_1{}^2 L^2+\ldots)$$

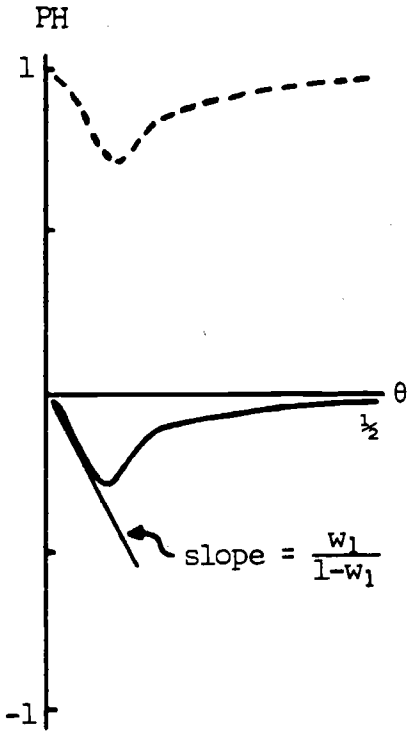
$$0 \leq w_1 < 1$$

$$w(z^{-1}) = \frac{w_0}{1-w_1 e^{2\pi i\theta}}$$

$$G(\theta) = w_0\sqrt{1+w_1^2-2w_1\cos 2\pi\theta}$$

$$PH(\theta) = \arctan\left(\frac{-w_1\sin 2\pi\theta}{1-w_1\cos 2\pi\theta}\right)$$





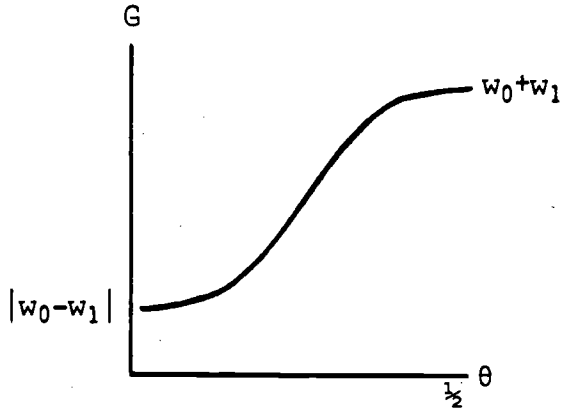
5. First Differences

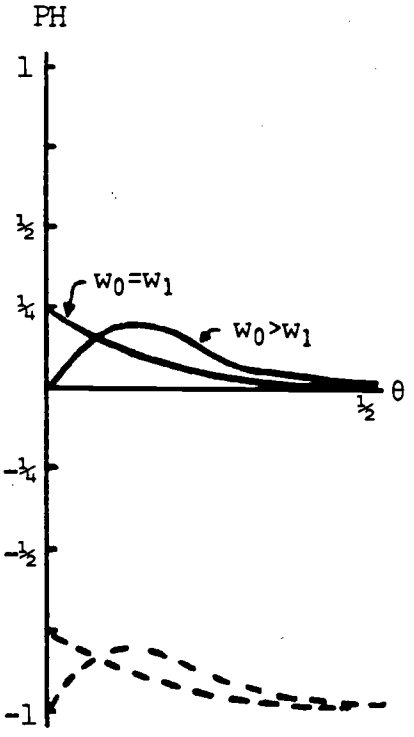
$$w(L) = w_0 - w_1 L \qquad w_0, w_1 > 0$$

$$w(z^{-1}) = w_0 - w_1 e^{2\pi i\theta}$$

$$G(\theta) = \sqrt{w_0^2+w_1^2-2w_0 w_1\cos 2\pi\theta}$$

$$PH(\theta) = \arctan\left(\frac{-w_1\sin 2\pi\theta}{w_0-w_1\cos 2\pi\theta}\right)$$





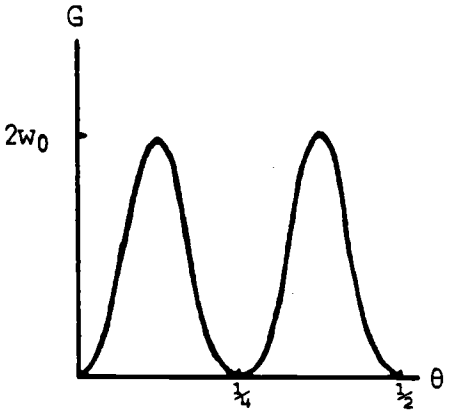
6. Four Period Differences

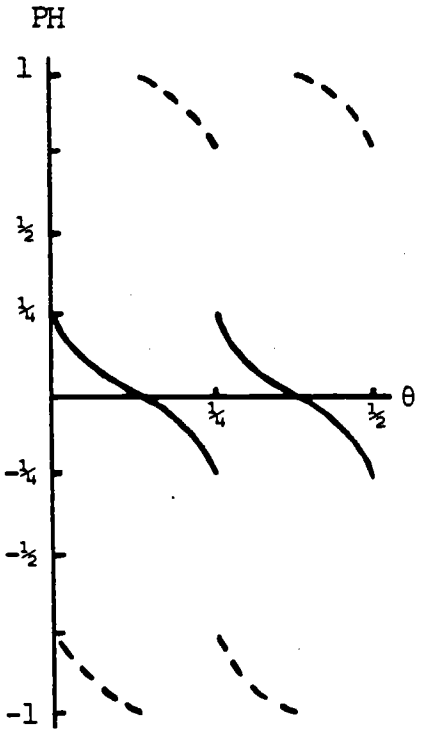
$$w(L) = w_0(1-L^4)$$

$$w(z^{-1}) = w_0(1-e^{8\pi i\theta})$$

$$G(\theta) = w_0\sqrt{2-2\cos(8\pi\theta)}$$

$$PH(\theta) = \arctan\left(\frac{-\sin 8\pi\theta}{1-\cos 8\pi\theta}\right)$$





APPENDIX

LEMMA A.1: If x is a stochastic process generated by the model

$$A(L)\, x_t = B(L)\, e_t$$

where e_t is a series of independent identically distributed random variables with variance σ^2, and $A(L)$ has all roots outside the unit circle, then the spectrum of x is given by

$$f_x(\theta) = \sigma^2 |B(z)|^2 / |A(z)|^2$$

where $z = e^{(-2\pi i\theta)}$

Proof: Consider the moving average process

$$x_t = \sum_{j=0}^{q} b_j\, e_{t-j}$$

where the e are all independent. Then

$$\gamma(s) = E x_{t+s} x_t = E \sum_{j=0}^{q} b_j\, e_{t+s-j} \sum_{k=0}^{q} b_k\, e_{t-k}$$

where the expectation on the right only has non-zero values where $k = s-j$ and $0 \leq k \leq q$. Therefore for $q \geq s \geq 0$

$$\gamma(s) = \sigma^2 \sum_{j=0}^{s} b_j\, b_{s-j}$$

for 0 otherwise. The spectrum of x is defined using equation (5) and the symmetry of γ by

$$f(\theta) = \sum_{s=-\infty}^{\infty} \gamma(s) z^s = \sum_{s=1}^{q} \gamma(s)\, (z^s + z^{-s}) + \gamma(0)$$

$$= \sigma^2 \sum_{s=1}^{q} \sum_{j=0}^{s} b_j\, b_{s-j}\, (z^s + z^{-s}) + \sigma^2 \sum_{j=0}^{q} b_j^2$$

which can be written

$$f(\theta) = \sigma^2 \sum_{j=0}^{q} b_j z^j \sum_{k=0}^{q} b_k z^{-k} = \sigma^2 \left| \sum_{j=0}^{q} b_j z^j \right|^2$$

There is nothing in this proof which requires that q be finite. Therefore, since every stable ARMA process has a (possibly infinite dimensional) moving average representation, the result is true for any ARMA process.

LEMMA A.2: If y is generated by a distributed lag model

$$y = w(L)\, x + u$$

where x and u are uncorrelated covariance stationary processes, then

$$f_{xy}(\theta) = w(z^{-1})\, f_x(\theta)$$

and

$$f_y(\theta) = |w(z)|^2\, f_x(\theta) + f_u(\theta)$$

where $z = e^{(2\pi i\theta)}$

Proof:

Without loss of generality take both x and y to have mean zero

$$\gamma_{xy}(s) = Ex_{t+s}\, y_t$$

$$= E\Sigma w_j x_{t-j} x_{t+s} + Ex_{t+s} u_t$$

$$\gamma_{xy}(s) = \Sigma w_j \gamma(s+j)$$

$$f_{xy}(\theta) = \sum_{s=-\infty}^{\infty} \gamma_{xy}(s) z^s = \Sigma_j \Sigma_s w_j \gamma(s+j) z^{s+j} z^{-j}$$

$$= \Sigma w_j z^{-j} f_x(\theta)$$

$$f_{xy}(\theta) = w(z^{-1}) f_x(\theta).$$

And

$$\gamma_y(s) = Ey_{t+s} y_t$$

$$= E(\Sigma w_j x_{t+s-j} + u_{t+s})(\Sigma w_k x_{t-k} + u_t)$$

$$\gamma_y(s) = \Sigma_j \Sigma_k w_j w_k \gamma_x(s-j+k) + \gamma_u(s)$$

$$f_y(s) = \sum_s \gamma_y(s) z^s$$

$$= \sum_{sjk}\sum\sum \omega_j \omega_k \gamma_x(s-j+k) z^{s-j+k} z^{-k} z^j + \sum_s \gamma_u(s) z^s$$

$$= \sum \omega_j z^j \sum \omega_k z^{-k} f_x(\theta) + f_u(\theta)$$

$$f_y(s) = |w(z)|^2 f_x(\theta) + f_u(\theta).$$

REFERENCES

(1) Cooley, J.W., P.A.W. Lewis and P.D. Welch. *The Fast Fourier Transform Algorithm and its Applications*. IBM Research Paper RC-1743, 1967.

(2) _____ "The Application of the Fast Fourier Transform Algorithm to the Estimation of Spectra and Cross Spectra", *Journal of Sound Vibrations*, 1970.

(3) Dhrymes, P.J. *Econometrics: Statistical Foundations and Applications*. New York: Harper and Row, 1970.

(4) _____ *Distributed Lags; Problems of Estimation and Formulation*. San Francisco: Holden Day, 1971.

(5) Durbin, J. "Tests for Serial Correlation in Regression Analysis Based on the Periodogram of Least Squares Residuals", *Biometrica*, 56, (1969), pp. 1-14.

(6) Granger, C.W.J. in association with M. Hatanaka. *Spectral Analysis of Economic Time Series*. Princeton: Princeton University Press, 1964.

(7) Fishman, G.S. *Spectral Methods in Econometrics*. Cambridge: Harvard University Press, 1969.

(8) Hannan, E.J. *Multiple Time Series*. New York: J. Wiley, 1970.

(9) Jenkins, G.M. "General Considerations in the Analysis of Spectra", *Technometrics*, 3, 2, (1961), pp. 133-166.

(10) Jenkins, G.M. and D.G. Watts. *Spectral Analysis and Its Applications*. San Francisco: Holden Day, 1969.

(11) Jones, R.H., "A Reappraisal of the Periodogram in Spectral Analysis", *Technometrics*, 7, (1965), pp. 531-542.

(12) Nerlove, M., "Spectral Analysis of Seasonal Adjustment Procedures", *Econometrica*, 32, (July 1964), pp. 241-286.

(13) Parzen, E., "Mathematical Considerations in the Estimation of Spectra", *Technometrics*, 3, 2, (1961), pp. 167-190.

(14) _____ "Multiple Time Series Modeling" in *Multivariate Analysis*. New York: Academic Press, 1969.

(15) Tick, L.H., Letter to the Editor, *Technometrics*, 8, (1966), pp. 559-561.