The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms

PETER D. WELCH

Abstract—The use of the fast Fourier transform in power spectrum analysis is described. Principal advantages of this method are a reduction in the number of computations and in required core storage, and convenient application in nonstationarity tests. The method involves sectioning the record and averaging modified periodograms of the sections.

INTRODUCTION

THIS PAPER outlines a method for the application of the fast Fourier transform algorithm to the estimation of power spectra, which involves sectioning the record, taking modified periodograms of these sections, and averaging these modified periodograms. In many instances this method involves fewer computations than other methods. Moreover, it involves the transformation of sequences which are shorter than the whole record which is an advantage when computations are to be performed on a machine with limited core storage. Finally, it directly yields a potential resolution in the time dimension which is useful for testing and measuring nonstationarity. As will be pointed out, it is closely related to the method of complex demodulation described by Bingham, Godfrey, and Tukey.[1]

THE METHOD

Let $X(j)$, $j=0, \cdots, N-1$ be a sample from a stationary, second-order stochastic sequence. Assume for simplicity that $E(X) = 0$. Let $X(j)$ have spectral density $P(f)$, $|f| \leq \frac{1}{2}$. We take segments, possibly overlapping, of length L with the starting points of these segments D units apart. Let $X_1(j)$, $j=0, \cdots, L-1$ be the first such segment. Then

$$X_1(j) = X(j) \qquad j = 0, \cdots, L - 1.$$

Similarly,

$$X_2(j) = X(j + D) \qquad j = 0, \cdots, L - 1,$$

and finally

$$X_K(j) = X(j + (K - 1)D) \quad j = 0, \cdots, L - 1.$$

We suppose we have K such segments; $X_1(j), \cdots, X_K(j)$, and that they cover the entire record, i.e., that $(K-1)D+L = N$. This segmenting is illustrated in Fig. 1.

The method of estimation is as follows. For each segment of length L we calculate a modified periodogram. That is, we select a data window $W(j)$, $j=0, \cdots, L-1$, and form the sequences $X_1(j)W(j), \cdots, X_K(j)W(j)$. We then take the finite Fourier transforms $A_1(n), \cdots, A_K(n)$ of these sequences. Here

$$A_k(n) = \frac{1}{L} \sum_{j=0}^{L-1} X_k(j) W(j) e^{-2kijn/L}$$

and $i=(-1)^{1/2}$. Finally, we obtain the K modified periodograms

$$I_k(f_n) = \frac{L}{U} \, | \, A_k(n) \, |^2 \qquad k = 1, 2, \cdots, K,$$

where

$$f_n = \frac{n}{L} \qquad n = 0, \cdots, L/2$$

and

$$U = \frac{1}{L} \sum_{j=0}^{L-1} W^2(j).$$

The spectral estimate is the average of these periodograms, i.e.,

$$\hat{P}(f_n) = \frac{1}{K} \sum_{k=1}^{K} I_k(f_n).$$

Now one can show that

$$E\{ \hat{P}(f_n) \} = \int_{-1/2}^{1/2} h(f) P(f - f_n) df$$

where

$$h(f) = \frac{1}{LU} \left| \sum_{j=0}^{L-1} W(j) e^{2\pi i f j} \right|^2$$

and

$$\int_{-1/2}^{1/2} h(f) df = 1.$$

[1] C. Bingham, M. D. Godfrey, and J. W. Tukey, "Modern techniques of power spectrum estimation," this issue, p. 56–66.
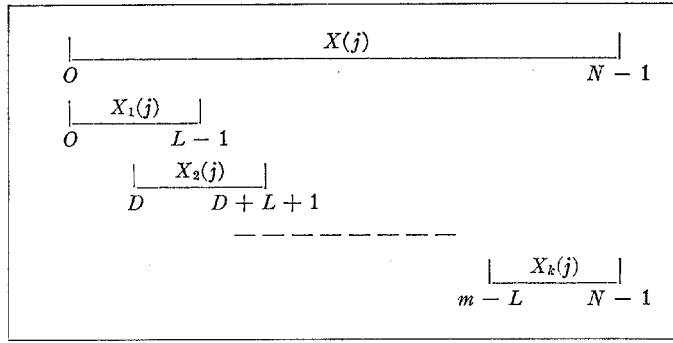
Fig. 1. Illustration of record segmentation.

Hence, we have a spectral estimator $\hat{P}(f)$ with a resultant spectral window whose area is unity and whose width is of the order of $1/L$.

CHOICE OF DATA WINDOWS

We suggest two reasonable choices for the data window $W(j)$; one of them has the *shape* $1 - t^2$: $-1 \leq t \leq 1$ and gives a spectral window which, when the two are normalized to have the same half-power width, is very close in shape to the hanning or cosine arch spectral window; the other data window has the *shape* $1 - |t|$: $-1 \leq t \leq 1$ and gives the Parzen spectral window. The actual functions for a particular segment length L are

$$W_1(j) = 1 - \left[\frac{j - \frac{L-1}{2}}{\frac{L+1}{2}} \right]^2$$

and

$$W_2(j) = 1 - \left| \frac{j - \frac{L-1}{2}}{\frac{L+1}{2}} \right|$$

$$j = 0, 1, \cdots, L - 1.$$

The resultant spectral windows corresponding to these data windows are given approximately by

$$h_1(f) \approx \frac{1}{LU} \left\{ \frac{2}{\pi^2 (L+1) f^2} \left[\frac{\sin \{(L+1)\pi f\}}{(L+1)\pi f} - \cos\{(L+1)\pi f\} \right] \right\}^2$$

$$h_2(f) \approx \frac{1}{LU} \left\{ \frac{(L+1)}{2} \frac{\sin^2 \{(L+1)\pi f/2\}}{\{(L+1)\pi f/2\}^2} \right\}^2.$$

In the preceding approximations L is a scale parameter. In changing L we change the shape of $h_1(f)$ and $h_2(f)$ only in stretching or shrinking the horizontal dimension. For $h_1(f)$ the half-power width is

$$\Delta_1 f \approx \frac{(1.16)}{L+1}.$$

For $h_2(f)$ the half-power width is

$$\Delta_2 f \approx \frac{(1.28)}{L+1}.$$

THE VARIANCES OF THE ESTIMATES

As developed above our estimator is given by

$$\hat{P}(f_n) = \frac{1}{K} \sum_{k=1}^{K} I_k(f_n), \qquad (n = 0, 1, \cdots, L/2).$$

Now, if we let

$$d(j) = \text{Covariance} \{I_k(f_n), I_{k+j}(f_n)\}$$

then it is easily shown that

$$\text{Var} \{\hat{P}(f_n)\} = \frac{1}{K} \left\{ d(0) + 2 \sum_{j=1}^{K-1} \frac{K-j}{K} d(j) \right\}.$$

Further, if

$$\rho(j) = \text{Correlation} \{I_k(f_n), I_{k+j}(f_n)\} = \frac{d(j)}{d(0)}$$

then,

$$\text{Var} \{\hat{P}(f_n)\} = \frac{d(0)}{K} \left\{ 1 + 2 \sum_{j=1}^{K-1} \frac{K-j}{K} \rho(j) \right\}$$

$$= \frac{\text{Var} \{I_k(f_n)\}}{K} \left\{ 1 + 2 \sum_{j=1}^{K-1} \frac{K-j}{K} \rho(j) \right\}.$$

Assume now that $X(j)$ is a sample from a Gaussian process and assume that $P(f)$ is flat over the passband of our estimator. Then we can show[2] that

$$\text{Var} \{I_k(f_n)\} = P^2(f_n).$$

Further, under the above assumptions and assuming that $h(f - f_n) = 0$ for $f < 0$ and $f > \frac{1}{2}$ we can show[3] that

$$\rho(j) = \left[\sum_{k=0}^{L-1} W(k) W(k+jD) \right]^2 \left[\sum_{k=0}^{L-1} W^2(k) \right]^2.$$

Hence, we have the following result which enables us to estimate the variances of $\hat{P}(f_n)$ when f_n is not close to 0 or $\frac{1}{2}$.

Result: If $X(j)$ is a sample from a Gaussian process, and $P(f)$ is flat over the passband of the estimator, and $h(f - f_n) = 0$ for $f < 0$ and $f > \frac{1}{2}$, then

$$\text{Var} \{\hat{P}(f_n)\} = \frac{P^2(f_n)}{K} \left\{ 1 + 2 \sum_{j=1}^{K-1} \frac{K-j}{K} \rho(j) \right\}$$

[2] P. D. Welch, "A direct digital method of power spectrum estimation," *IBM J. Res. and Dev.*, vol. 5, pp. 141–156, April 1961.
[3] In Welch[2] we obtained the variance spectrum of $I_k(f_n)$ considered as a function of time. The above result is obtained by taking the Fourier transform of this spectrum.

where

$$\rho(j) = \left[\sum_{k=0}^{L-1} W(k)W(k+jD) \right]^2 \Big/ \left[\sum_{k=0}^{L-1} W^2(k) \right]^2 .$$

For estimating the spectrum of $P(f_n)$ at 0 and $\frac{1}{2}$ the variance is twice as great, as given by the following result:

Result: If $X(j)$ is a sample from a Gaussian process and $P(f)$ is flat over the passband of the estimator, then

$$\text{Var}\left\{ \hat{P}(0 \text{ or } 1/2) \right\}$$
$$= \frac{2P^2(0 \text{ or } 1/2)}{K} \left\{ 1 + 2 \sum_{=1}^{K-1} \frac{K-j}{K} \rho(j) \right\}$$

where $\rho(j)$ is as defined above.

In the above results note that $\rho(j) \geq 0$ and that $\rho(j) = 0$ if $D \geq L$. Hence, if we average over K segments the best we can do is obtain a reduction of the variance by a factor $1/K$. Further, this $1/K$ reduction can be achieved (under these conditions) if we have nonoverlapping segments. Hence, if the total number of points N can be made sufficiently large the computationally most efficient procedure for achieving any desired variance is to have nonoverlapping segments, i.e., to let $D = L$. In this case we have

$$\text{Var}\left\{ \hat{P}(f_n) \right\} = \frac{P^2(f_n)}{K} = \frac{P^2(f_n)L}{N} .$$

Further, under these conditions $E\left\{ \hat{P}(f_n) \right\} = E\left\{ I_K(f_n) \right\} = P(f_n)$ and, hence,

$$\frac{E^2\{ P(f_n) \}}{\text{Var}\left\{ P(f_n) \right\}} = K$$

and the equivalent degrees of freedom of the approximating chi-square distribution is given by

$$\text{E.D.F.} \left\{ \hat{P}(f_n) \right\} = 2K.$$

If the total number of points N cannot be made arbitrarily large, and we wish to get a near maximum reduction in the variance out of a fixed number of points then a reasonable procedure is to overlap the segments by one half their length, i.e., to let $D = L/2$. In this case, if we use $W_1(j)$ as the data window we get $\rho(1) \approx 1/9$ and $\rho(j) = 0$ for $j > 1$. Letting $\hat{P}_1(f_n)$ be the estimate, we have

$$\text{Var}\left\{ \hat{P}_1(f_n) \right\} = \frac{P^2(f_n)}{K} \left\{ 1 + \frac{2}{9} - \frac{2}{9K} \right\}$$
$$\approx \frac{11 P^2(f_n)}{9K} .$$

The factor 11/9, compared with the factor 1.0 for non-overlapped segments, inflates the variance. However, an overall reduction in variance for fixed record length is achieved because of the difference in the value of K. For nonoverlapped segments we have $K = N/L$; for the overlapping discussed here

$$K = \frac{N}{L/2} - 1 = \frac{2N}{L} - 1 \approx \frac{2N}{L} .$$

Therefore, for fixed N and L the overall reduction in variance achieved by this overlapping is by a factor of 11/18. Now again $E\left\{ \hat{P}_1(f_n) \right\} = P(f_n)$ and, hence,

$$\frac{E^2\{ \hat{P}_1(f_n) \}}{\text{Var}\left\{ \hat{P}_1(f_n) \right\}} \approx \frac{9K}{11} \approx \frac{18N}{11L} .$$

Finally,

$$\Delta_1(f) \approx \frac{(1.16)}{L+1} \approx \frac{7}{6L} .$$

Thus,

$$\frac{E^2\{ \hat{P}_1(f_n) \}}{\text{Var}\left\{ \hat{P}_1(f_n) \right\}} \approx 1.4 N \Delta f$$

and the equivalent degrees of freedom of the approximating chi-square distribution is

$$\text{E.D.F.} \left\{ \hat{P}_1(f_n) \right\} \approx 2.8 N \Delta f.$$

Similarly, if we use $W_2(j)$ as our data window we get $\rho(1) = 1/16$ and $\rho(j) = 0$, $j > 1$. Letting $\hat{P}_2(f_n)$ be the estimate in this case we get, by following the above steps and using the result $\Delta_2 f = (1.28)/(L+1)$, that the equivalent degrees of freedom is again approximately

$$\text{E.D.F.} \left\{ \hat{P}_2(f_n) \right\} \approx 2.8 N \Delta f.$$

Thus, both $W_1(j)$ and $W_2(j)$ yield roughly the same variance when adjusted to have windows of equal half power width. Finally, we should point out that the above variances need to be doubled and the equivalent degrees of freedom halved for the points $f_n = 0$ and $\frac{1}{2}$.

DETAILS IN THE APPLICATION OF THE FAST FOURIER TRANSFORM ALGORITHM

Our estimator $\hat{P}(f_n)$ is given by

$$\hat{P}(f_n) = \frac{1}{K} \sum_{k=1}^{K} I_k(f_n) = \frac{L}{UK} \sum_{k=1}^{K} |A_k(n)|^2,$$

where L is the length of the segments, and K is the number of segments into which the record is broken, and

$$U = \frac{1}{L} \sum_{j=0}^{L-1} W^2(j).$$

We will first discuss how the complex algorithm can be used to obtain the summation $\sum_{k=1}^{K} |A_k(n)|^2$ two terms at a time with $K/2$ [or $(K+1)/2$, if K is odd] rather than K transforms. Suppose K is even and let

$$\left. \begin{aligned} Y_1(j) &= X_1(j)W(j) + iX_2(j)W(j) \\ &\;\;\vdots \\ Y_{K/2}(j) &= X_{K-1}(j)W(j) + iX_K(j)W(j) \end{aligned} \right\} j = 0, \cdots, L-1.$$

Let $B_k(n)$ be the transform of $Y_k(j)$. Then, by the linearity property of the finite Fourier transform

$$B_k(n) = A_{2k-1}(n) + iA_{2k}(n).$$

Further,

$$B_k(N - n) = A_{2k-1}(N - n) + iA_{2k}(N - n)$$
$$= \widetilde{A_{2k-1}(n)} + i\widetilde{A_{2k}(n)}.$$

Now,

$$|B_k(n)|^2 = (A_{2k-1}(n) + iA_{2k}(n))(\widetilde{A_{2k-1}(n)} - i\widetilde{A_{2k}(n)})$$
$$|B_k(N - n)|^2 = (A_{2k-1}(n) - iA_{2k}(n))(\widetilde{A_{2k-1}(n)} + i\widetilde{A_{2k}(n)}).$$

These equations yield, with some algebra,

$$|B_k(n)|^2 + |B_k(N - n)|^2$$
$$= 2(|A_{2k-1}(n)|^2 + |A_k(n)|^2).$$

Hence, finally,

$$\hat{P}(f_n) = \frac{L}{2UK} \sum_{k=1}^{K/2} (|B_k(n)|^2 + |B_k(N - n)|^2).$$

If K is odd this procedure can be extended in an obvious fashion by defining $Y_{(K+1)/2}(j) = X_K(j)$ and summing from 1 to $(K+1)/2$.

A second observation on the actual application of the algorithm concerns the bit-inverting. If the algorithm is applied as described here, and one is especially concerned with computation time, then the bit-inverting could be postponed until after the summation. Thus, instead of bit-inverting $K/2$ times, one would only have to bit-invert once.

COMPUTATION TIME

The time required to perform a finite Fourier transform on a sequence of length L is approximately $k'L \log_2 L$ where k' is a constant which depends upon the program and type of computer. Hence, if we overlap segments by an amount $L/2$ we require an amount of computing time (performing two transforms simultaneously) approximately equal to

$$\left(\frac{1}{2}\right)\left(\frac{N}{L/2}\right)k'L \log_2 L = k'N \log_2 L,$$

plus the amount of time required to premultiply by the data window and average. If we only consider the time required for the Fourier transformation this compares with approximately $k'N(\log_2 N)/2$ for the smoothing of the periodogram. Hence, if $L < (N)^{1/2}$ it requires less computing time than the smoothing of the periodogram.

RELATION OF THIS METHOD TO COMPLEX DEMODULATION

It is appropriate to mention here the process of complex demodulation and its relation to this method of spectral estimation. Complex demodulation is discussed in Tukey,[4] Godfrey,[5] and Bingham, Godfrey, and Tukey.[1] The functions $A_k(n)e^{-2\pi ikD/L}$ considered as functions of k are complex demodulates sampled at the sampling period D. In this case the demodulating function is $e^{-2\pi ifnj}$. A phase coherency from sample to sample is retained in the complex demodulates. This phase is lost in estimating the spectrum and, hence, as a method of estimating spectra, complex demodulation is identical to the method of this section. However, additional information can be obtained from the time variation of the phase of the demodulates.

THE SPACING OF THE SPECTRAL ESTIMATES

This method yields estimates spaced $1/L$ units apart. If more finely spaced estimates are desired zeros can be added to the sequences $X_k(j)W(j)$ before taking the transforms. If L' zeros are added giving time sequences $L+L' = M$ long and we let $A_k'(n)$ be the finite Fourier transforms of these extended sequences, i.e.,

$$A_k'(n) = \frac{1}{M} \sum_{j=0}^{L-1} X_k(j)W(j)e^{-2\pi ijn/M}$$

then the modified periodogram is given by

$$I_k(f_n) = \frac{M^2}{LU} |A_k'(n)|^2$$

where

$$f_n = \frac{n}{M} \qquad n = 0, 1, \cdots, M/2.$$

Everything proceeds exactly as earlier except that we have estimates spaced at intervals of $1/M$ rather than $1/L$.

ESTIMATION OF CROSS SPECTRA

Let $X(j)$, $j=0, \cdots, N-1$, and $Y(j)$, $j=0, \cdots, N-1$, be samples from two second-order stochastic sequences. This method can be extended in a straightforward manner to the estimation of the cross spectrum, $P_{xy}(f)$. In exactly the same fashion each sample is divided in K segments of length L. Call these segments $X_1(j), \cdots, X_K(j)$ and $Y_1(j), \cdots, Y_K(j)$. Modified cross periodograms are calculated for each pair of segments $X_k(j), Y_k(j)$, and the average of these modified cross periodograms constitutes the estimate $\hat{P}_{xy}(f_n)$. The spectral window is the same as is obtained using this method for the estimation of the spectrum.

[4] J. W. Tukey, "Discussion, emphasizing the connection between analysis of variance and spectrum analysis," *Technometrics*, vol. 3, pp. 191–219, May 1961.

[5] M. D. Godfrey, "An exploratory study of the bispectrum of economic time series," *Applied Statistics*, vol. 14, pp. 48–69, January 1965.