TIME-FREQUENCY SPECTRUM ESTIMATION: AN ADAPTIVE MULTITAPER METHOD

James W. Pitton

MathSoft
1700 Westlake Ave N, Suite 500
Seattle, WA USA 98109
yojimbo@statsci.com

ABSTRACT

This paper extends Thomson's adaptive multitaper spectrum estimation method [17] to the nonstationary case. The general approach and the nonadaptive estimation procedure were first presented in [12]. The method uses time-frequency concentrated basis functions which generalize the properties of the prolate spheroidal waveforms [15]. Individual spectrograms computed with these eigenfunctions form direct time-frequency spectrum estimates, and are combined to form the multitaper time-frequency spectrum estimate. We then develop a new adaptive procedure which reduces the bias of the individual eigenestimates using an estimate of their leakage characteristics. The revised multitaper estimator then has correspondingly improved bias properties. An expression for the variance of the adaptive estimator is also derived, providing a complete characterization of the statistical time-frequency estimator.

1. INTRODUCTION

Time-frequency analysis is the process of estimating the time-varying spectral content of nonstationary signals, which would not be completely described by stationary spectral analysis. A general approach to time-frequency analysis is given by Cohen's bilinear class [3], which provides a general formulation for time-frequency representations. This general class includes well-known analysis methods such as the spectrogram and Wigner distribution. Other methods of time-frequency analysis are based on an evolutionary spectrum approach (*e.g.*, [14, 6, 7]), which models the spectrum as a slowly varying envelope of a complex sinusoid. This assumption allows the averaging of short-time spectral estimates to stabilize the variance. Computationally, the evolutionary spectrum methods fall within Cohen's class; however, they have a distinctly different theoretical foundation, being statistical estimation methods.

Most methods of time-frequency analysis are fixed processing techniques — the same analysis is used for all signals. In other words, the time-frequency representation is *signal-independent*. Such representations either have trade-offs in time and frequency resolution, like the spectrogram, or are not generally nonnegative, like the Wigner distribution. Nonnegativity is particularly important for spectral estimation, whether stationary or not. The shortcomings of these fixed time-frequency analysis methods may be avoided by restricting oneself to positive time-frequency distributions (TFDs) [4]. Positive TFDs are, of course, everywhere nonnegative, and do not suffer a trade-off in time and frequency resolution. Unlike any fixed methods of time-frequency analysis, they yield the correct univariate marginal distributions in time and frequency (the instantaneous energy and the energy spectral density):

$$P(t,\omega) \geq 0, \tag{1}$$

$$\int P(t,\omega)d\omega = |s(t)|^2, \tag{2}$$

$$\int P(t,\omega)dt = |S(\omega)|^2, \tag{3}$$

where $S(\omega)$ denotes the Fourier transform of the finite energy signal $s(t)$, and all integrals are from $-\infty$ to ∞.

The first method for generating positive TFDs used constrained optimization, minimizing the cross-entropy to a prior distribution subject to a set of linear constraints [8]. Positive TFDs have been linked to the evolutionary spectrum and estimated via deconvolution [13]. Least-squares estimation has also been used to compute positive TFDs [10]. Approximate solutions for positive TFDs have been obtained through a nonlinear combination of spectrograms [9].

Another approach to computing time-frequency spectra has been to extend Thomson's multitaper spectral estimation method [17] to the nonstationary case through a sliding-window framework [16]. [1] developed a multitaper time-frequency spectrum, including a significance test for nonstationary tones, using Hermite windows, which have previously been shown to maximize a time-frequency concentration measure [5]. [2] extended the Hermite multiwindow method to include a means of reducing artifacts using a time-frequency mask.

While these methods all provide some representation of the time-varying frequency content of a signal, they do not relate the computed distribution to an underlying time-frequency spectrum (*e.g.*, [1] minimizes the bias between the multitaper TFD and the Wigner distribution; however, the Wigner distribution is not nonnegative for arbitrary signals, and as such is not a valid time-frequency spectrum). As a result, there is no quantitative measurement of the accuracy of the representation. Many real-world applications require bias and variance expressions for measured quantities,

0-7803-5073-1/98/$10.00 ©1998 IEEE

so that confidence intervals may be assigned or hypothesis testing performed. A time–frequency analysis technique meeting these requirements was presented in [12]. Based on a time–varying filter formulation for positive TFDs [11], the method constructed a multitaper time–frequency spectral estimate using eigenvectors which diagonalize the nonstationary spectral generating function. These eigenvectors are maximally concentrated (and doubly–orthogonal) in time–frequency. We review this multitaper estimation procedure, and then develop a new adaptive procedure which reduces the bias of the individual eigenestimates using an estimate of their leakage characteristics. We also derive an expression for the variance of the adaptive estimator, providing a complete characterization of the statistical time–frequency estimation procedure.

2. EIGENVECTORS FOR TIME–FREQUENCY SPECTRAL ESTIMATION

A signal $s(t)$ may frequently be modeled as the output of a linear time–varying filter driven by white noise:

$$s(t) = \int h(t,\tau)e(\tau)d\tau. \tag{4}$$

$e(t)$ is bandlimited Gaussian white noise with bandwidth much greater than that of the filter $h(t,\tau)$:

$$e(t) = \int e^{j\omega t}dZ(\omega). \tag{5}$$

$dZ(\omega)$ is an orthogonal process with unit variance. The signal can then be written as:

$$s(t) = \int H(t,\omega)e^{j\omega t}dZ(\omega), \tag{6}$$

where $H(t,\omega)$ is defined as the Fourier transform of $h(t, t-\tau)$ [11]. This formulation is a straightforward extension of the spectral representation theorem for stationary processes [14]. The time–frequency spectrum is defined by:

$$P(t,\omega) = |H(t,\omega)|^2. \tag{7}$$

This formulation for a time–frequency spectrum is of the same general form as Priestley's evolutionary spectrum [14]. However, we have not constrained $H(t,\omega)$ to be slowly–varying. This form for $P(t,\omega)$ also satisfies the stochastic equivalent of the time and frequency marginals (equations (2)–(3)); the relationship between the above time–varying spectrum and positive TFDs is discussed in [11].

Given a signal $s(t)$, we want to estimate $P(t,\omega)$; however, direct inversion of equation (6) is impossible. A rough estimate of the time–varying frequency content of $s(t)$ may be obtained by computing its short-time Fourier transform (STFT):

$$S_s(t,\omega) = \int s(\tau)g(t-\tau)e^{-j\omega\tau}d\tau, \tag{8}$$

where $g(t)$ is a rectangular window of length T. A relationship between the STFT and $H(t,\omega)$ is obtained by replacing $s(t)$ by its time–frequency spectral formulation:

$$S_s(t,\omega) = \int\int H(\tau,\theta)g(t-\tau)e^{-j(\omega-\theta)\tau}dZ(\theta)d\tau. \tag{9}$$

To solve for the time-varying spectrum $H(\tau,\theta)$, we need to invert the STFT operator $g(t-\tau)e^{-j\omega\tau}$. This inversion is an inherently ill-posed problem. Instead, we approximate the inverse solution by regularizing it to some region $R(t,\omega)$ in the time-frequency plane, much as Thomson regularized the spectral inversion to a bandwidth W in his multitaper approach [17]. For simplicity throughout, we will define $R(t,\omega)$ to be a square region of time–frequency of dimension $\Delta T \times \Delta W$; however, the results readily generalize to arbitrary regions $R(t,\omega)$.

In the case of spectral estimation, the operator is square and Toeplitz; its regularized inverse is found through an eigenvector decomposition. Such is not the case in the time–frequency problem; the STFT operator is neither full rank nor square. To diagonalize it, we apply a Singular Value Decomposition, finding the left and right eigenvectors $u(\tau)$ and $V(t,\omega)$ and the associated eigen (singular) values λ:

$$g(t-\tau)e^{-j\omega\tau} = \sum_k \lambda_k u_k(\tau)V_k^*(t,\omega). \tag{10}$$

The eigenvectors $u(\tau)$ and $V(t,\omega)$ form an STFT pair:

$$V(t,\omega) = \int u(\tau)g(t-\tau)e^{-j\omega\tau}d\tau. \tag{11}$$

The SVD relationship between $u(\tau)$ and $V(t,\omega)$ is obtained by applying the STFT operator to $V(t,\omega)$, computing the integrals only over $\Delta T \times \Delta W$:

$$\lambda u(\tau) = \int_{\Delta T}\int_{\Delta W} V(t,\omega)g(t-\tau)e^{j\omega\tau}d\omega dt. \tag{12}$$

The inverse STFT computed over all (t,ω) also holds. This equation can be reduced to a standard eigenvector equation by substituting for $V(t,\omega)$:

$$\lambda u(\tau) = \int_{\Delta T}\int_{\Delta W}\int u(s)g(t-s)g(t-\tau)e^{j\omega(\tau-s)}ds d\omega dt. \tag{13}$$

The integral in ω reduces to a sinc function, or for the discrete case, Dirichlet's kernel. The integral in t can be easily computed over the two rectangular windows, the result of which for convenience we define by:

$$f(\tau,s) = \int_{\Delta T} g(t-s)g(t-\tau)dt. \tag{14}$$

The eigenvalue equation for $u(\tau)$ is then:

$$\lambda u(\tau) = \int 2\Delta W \operatorname{sinc}(\Delta W(\tau - s))f(\tau,s)u(s)ds. \tag{15}$$

We can then solve for $u(\tau)$ using standard eigenvector solution methods.

There are two extremely important properties of these eigenvectors for the time–frequency spectral estimation problem. These properties are obvious results of diagonalizing the STFT operator over a finite region of the time–frequency plane. First, the first left eigenvector u_1 maximizes a time–frequency energy concentration measure

$$u_1(\tau) = \arg_u \max \int_{\Delta T}\int_{\Delta W} |S_u(t,\omega)|^2 \, d\omega dt \tag{16}$$

over the space of unit–energy functions. In other words, the first right eigenvector $V_1(t,\omega)$, the STFT of the first left eigenvector $u_1(\tau)$, has the greatest proportion of energy in $\Delta T \times \Delta W$ (or more generally in $R(t,\omega)$) of any STFT. The second eigenvector maximizes the energy in a subspace orthogonal to the first, and so on. It is straightforward to show that maximization of this integral reduces to the eigenequation for $u(\tau)$. This time–frequency concentration property is akin to the frequency concentration property of the prolate spheroidal waveforms (or the prolate spheroidal sequences, in the discrete case) [15]. The second important property, again similar to the Slepian functions, is a double–orthogonality property. Since the $u_k(\tau)$ are orthonormal functions on the real line, the $V_k(t,\omega)$ are orthonormal on the entire time–frequency plane. In addition, the $V_k(t,\omega)$ are also orthogonal over $\Delta T \times \Delta W$. This orthogonality is to be expected, given the maximization of the time–frequency concentration measure. The Slepian functions, by contrast, are orthonormal on the real line and orthogonal on the interval W. This orthogonality on $\Delta T \times \Delta W$ lets us construct a regularized estimate of the time–frequency spectrum, just as the orthogonality of the Slepian functions on W provides the basis for Thomson's multitaper spectral estimation method.

3. MULTITAPER TIME–FREQUENCY SPECTRUM ESTIMATION

We now want to estimate $H(t,\omega)$ regularized to $\Delta T \times \Delta W$. We project $H(t,\omega)$ onto the rectangular region $\Delta T \times \Delta W$ around (t,ω) using the k^{th} left eigenvector $u_k(t)$:

$$H_k(t,\omega) \doteq \frac{1}{\sqrt{\lambda_k}} \int_{\Delta T}\int_{\Delta W} H(\tau,\theta) u_k(t-\tau) e^{-j(\omega-\theta)\tau} dZ(\theta) d\tau. \tag{17}$$

H_k is thus a direct, but unobservable, projection of $H(t,\omega)$ onto $\Delta T \times \Delta W$.

We next form an estimate of these expansion coefficients using the STFT. We compute the STFT of $s(t)$ using $u_k(t)$:

$$S_k(t,\omega) = \int\int H(\tau,\theta) u_k(t-\tau) e^{-j(\omega-\theta)\tau} dZ(\theta) d\tau, \tag{18}$$

i.e., the k^{th} eigenspectrum $S_k(t,\omega)$ is a projection of $H(t,\omega)$ onto the k^{th} left eigenvector $u_k(t)$, estimating $H_k(t,\omega)$ over $\Delta T \times \Delta W$. When $s(t)$ is a stationary white noise process, it follows that

$$E[|S_k(t,\omega)|^2] = |H(t,\omega)|^2 = P(t,\omega). \tag{19}$$

Thus, the individual eigenspectra are direct estimates of $P(t,\omega)$, and are unbiased when the spectrum is white.

Next, we form an estimate of $H(t,\omega)$ over $\Delta T \times \Delta W$ using the right eigenvectors $V_k(t,\omega)$ weighted by the projections of $H(t,\omega)$ onto $u_k(t)$:

$$\hat{H}(\bar{t},\bar{\omega};t,\omega) = \sum_{k=1}^{K} V_k(\bar{t}-t,\bar{\omega}-\omega) \frac{S_k(t,\omega)}{\lambda_k}, \tag{20}$$

where $K \approx \Delta T \Delta W$. Choosing $\Delta T \Delta W$ too small will result in estimates with poor bias and variance properties. The magnitude–square of $\hat{H}(\bar{t},\bar{\omega};t,\omega)$ yields our estimate of $P(t,\omega)$ over $\Delta T \times \Delta W$. This estimate is a χ^2 random variable with two degrees of freedom (except for DC and Nyquist) with variance $P^2(t,\omega)$. To reduce the variance, we average the estimate over $\Delta T \times \Delta W$, making use of the orthogonality of $V_k(t,\omega)$ on this region:

$$\begin{aligned}\hat{P}(t,\omega) &= \frac{1}{\Delta T \Delta W} \int_{\Delta T}\int_{\Delta W} \left|\hat{H}(\bar{t},\bar{\omega};t,\omega)\right|^2 d\bar{t}\, d\bar{\omega} \\ &= \frac{1}{\Delta T \Delta W} \sum_{k=1}^{K} \frac{|S_k(t,\omega)|^2}{\lambda_k}.\end{aligned} \tag{21}$$

The average of K direct estimates is a χ^2 random variable with $2K$ degrees of freedom; hence, the variance of this estimate is $P^2(t,\omega)/K$. If we choose ΔT to be a fixed proportion of the window length T, then this estimator is consistent for fixed ΔW.

4. ADAPTIVE TIME–FREQUENCY ESTIMATOR

When the time–frequency spectrum is not white, the individual eigenestimates will no longer be unbiased. Since the eigenvalues decrease with increasing order, the bias properties of the corresponding spectral estimates will likewise degrade. Regions where the time–frequency spectrum is low will be corrupted by leakage from high–energy regions. To reduce the bias in these estimates, we introduce a series of time–frequency weighting functions $d_k(t,\omega)$, and minimize the mean squared error between the unknown $H_k(t,\omega)$ and the k^{th} eigenspectrum $S_k(t,\omega)$. The mse is given by:

$$\text{mse} = E[|H_k(t,\omega) - d_k(t,\omega) S_k(t,\omega)|^2]. \tag{22}$$

Minimizing this quantity with respect to $d_k(t,\omega)$ gives:

$$d_k(t,\omega) = \frac{E[H_k(t,\omega) S_k^*(t,\omega)]}{E[|S_k(t,\omega)|^2]}. \tag{23}$$

The value of the denominator has previously been shown [13] to be:

$$E[|S_k(t,\omega)|^2] = \int\int P(t-\tau,\omega-\theta) W_k(\tau,\theta) d\theta d\tau, \tag{24}$$

where $W_k(t,\omega)$ is the Wigner distribution of $u_k(t)$.

In a similar fashion, we can obtain an estimate for the numerator of equation (23). Inserting equations (17)- (18) for $H_k(t,\omega)$ and $S_k(t,\omega)$ in the numerator and doing a simple change of variables gives:

$$E[H_k(t,\omega)S_k^*(t,\omega)] = \int\int_{\Delta T}\int_{\Delta W} H_k(t-u,\omega-\theta) H_k^*(t-v,\omega-\theta) u_k(u) u_k^*(v) e^{-j\theta(u-v)} d\theta du dv. \tag{25}$$

After again making use of the assumptions from [13], performing a change of variables, and changing the order of integration, this equation simplifies to:

$$E[H_k(t,\omega)S_k^*(t,\omega)] = \int_{\Delta W}\int\int_{u-\Delta T/2}^{u+\Delta T/2} P(t-u,\omega-\theta) u_k(u+v/2) u_k^*(u-v/2) e^{-j\theta v} dv du d\theta. \tag{26}$$

The inner integral (in v) may be evaluated separately, and is equal to:

$$\int_{u-\Delta T}^{u+\Delta T} u_k(u+\frac{v}{2})u_k^*(u-\frac{v}{2})e^{-j\theta v}dv = W_k(u,\theta) * \mathrm{sinc}(\theta\Delta T)e^{-j\theta u}. \tag{27}$$

Denoting the value of this integral by $W_k^{(*)}(u,\theta)$, the numerator of equation (23) is given by:

$$E[H_k(t,\omega)S_k^*(t,\omega)] = \int\int_{\Delta W} P(t-\tau,\omega-\theta)W_k^{(*)}(\tau,\theta)d\theta d\tau. \tag{28}$$

These results give the time–frequency weights in equation (23), and the corresponding estimate of $P(t,\omega)$ is:

$$\hat{P}(t,\omega) = \sum_{k=1}^{K} \frac{|d_k(t,\omega)|^2\,|S_k(t,\omega)|^2}{\sum_{k=1}^{K}|d_k(t,\omega)|^2}. \tag{29}$$

Note that these equations define an adaptive procedure for estimating $P(t,\omega)$. The K time–frequency weighting functions $d_k(t,\omega)$ are computed using an initial multitaper estimate of $P(t,\omega)$. This estimate is then updated using the weights, which in turn are recomputed using the new $\hat{P}(t,\omega)$. This procedure is repeated until the weights have converged to a steady state.

As mentioned above, the individual direct estimates $|S_k(t,\omega)|^2$ are unbiased estimators of $P(t,\omega)$ when $P(t,\omega)$ is white. As a result, the normalization in the denominator of equation (29) makes $\hat{P}(t,\omega)$ a similarly unbiased estimate of $P(t,\omega)$.

The variance properties of $\hat{P}(t,\omega)$ are also easily obtained. The individual $S_k(t,\omega)$ are approximately uncorrelated, given the assumption that $P(t,\omega)$ is white over $\Delta T \times \Delta W$. It immediately follows that the variance of $\hat{P}(t,\omega)$ is:

$$\mathrm{var}\left\{\hat{P}(t,\omega)\right\} = \frac{\sum_{k=1}^{K}|d_k(t,\omega)|^4}{\left(\sum_{k=1}^{K}|d_k(t,\omega)|^2\right)^2}P^2(t,\omega). \tag{30}$$

The independence assumption will be valid so long as the K eigenvectors are well-localized to $\Delta T \times \Delta W$. If additional eigenvectors with worse leakage properties are used, the individual direct estimators will no longer be uncorrelated, and the variance of the adaptive estimator will increase.

5. CONCLUSIONS

We have presented an adaptive multitaper estimator for time–varying spectra. This method extends the estimation method first presented in [12]. As with the non–adaptive method, the approach follows that used by Thomson [17]. A family of orthonormal windows with double–orthogonality properties in the time–frequency plane are used to form the estimate. Leakage in the estimator is controlled by the time–frequency concentration properties of the windows. The adaptive weights minimize the mean squared error of the multitaper estimator. An expression for the variance of the adaptive estimator is also presented.

6. REFERENCES

[1] M. Bayram and R. G. Baraniuk. Multiple window time-frequency analysis. In *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, pages 173–176, Paris, France, June 1996.

[2] F. Çakrak and P. Loughlin. Multiple window non–linear time–varying spectral analysis. *IEEE Transactions on Signal Processing*, 1997. submitted.

[3] L. Cohen. *Time-Frequency Analysis*. Prentice-Hall, 1995.

[4] L. Cohen and T. Posch. Positive time-frequency distributions. *IEEE Transactions on Acoustics Speech and Signal Processing*, 33(1):31–38, 1985.

[5] I. Daubechies. Time-frequency localization operators: a geometric phase space approach. *IEEE Transactions on Information Theory*, 34(4):605–612, 1992.

[6] A. Kayhan, A. El-Jaroudi, and L. Chaparro. The evolutionary periodogram for non-stationary signals. *IEEE Transactions on Signal Processing*, 42(6), 1994.

[7] A. Kayhan, A. El-Jaroudi, and L. Chaparro. Data–adaptive evolutionary spectral estimation. *IEEE Transactions on Signal Processing*, 43(1), 1995.

[8] P. Loughlin, J. Pitton, and L. Atlas. Construction of positive time-frequency distributions. *IEEE Transactions on Signal Processing*, 42(10):2697–2705, 1994.

[9] P. Loughlin, J. Pitton, and B. Hannaford. Approximating time-frequency density functions via optimal combinations of spectrograms. *IEEE Signal Processing Letters*, 1(12):199–202, 1994.

[10] J. Pitton. An algorithm for weighted least squares positive time–frequency distributions. In *SPIE Advanced Sig. Proc. Algs. Archs., Impl. VII*, volume 3162, 1997.

[11] J. Pitton. Linear and quadratic methods for positive time-frequency distributions. In *IEEE Conference on Acoustics, Speech, and Signal Processing*, volume 5, pages 3649–3652. IEEE, 1997.

[12] J. Pitton. Nonstationary spectrum estimation and time-frequency concentration. In *IEEE Conference on Acoustics, Speech, and Signal Processing*, volume IV, pages 2425–2428. IEEE, 1998.

[13] J. Pitton, L. Atlas, and P. Loughlin. Applications of positive time-frequency distributions to speech processing. *IEEE Trans. on Speech and Audio Proc.*, 2(4):554–566, 1994.

[14] M. Priestley. *Spectral Analysis of Time Series*. Academic Press - London, 1981.

[15] D. Slepian and H. Pollack. Prolate spheroidal wave functions, Fourier analysis and uncertainty. *Bell Sys. Tech. J.*, 40(1), 1961.

[16] D. Thomson and A. Chave. Jackknifed error estimates for spectra, coherences, and transfer functions. In S. Haykin, editor, *Advances in Spectrum Analysis and Array Processing, Vol. I*, pages 58–113. Prentice–Hall, 1991.

[17] D. J. Thomson. Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, 70:1055–1096, 1982.