JOURNAL OF MULTIVARIATE ANALYSIS 10, 51–59 (1980)

The Inverted Complex Wishart Distribution and Its Application to Spectral Estimation

PAUL SHAMAN*

University of Pennsylvania

Communicated by the Editors

The inverted complex Wishart distribution and its use for the construction of spectral estimates are studied. The density, some marginals of the distribution, and the first- and second-order moments are given. For a vector-valued time series, estimation of the spectral density at a collection of frequencies and estimation of the increments of the spectral distribution function in each of a set of frequency bands are considered. A formal procedure applies Bayes theorem, where the complex Wishart is used to represent the distribution of an average of adjacent periodogram values. A conjugate prior distribution for each parameter is an inverted complex Wishart distribution. Use of the procedure for estimation of a 2×2 spectral density matrix is discussed.

1. INTRODUCTION

In multiple time series analysis complex multivariate distributions are commonly used to describe estimates of frequency domain parameters. A review of complex multivariate distributions and their application in time series has been given by Krishnaiah [8]. The complex Wishart distribution, in particular, was introduced and used by Goodman [4, 5] to approximate the distribution of an estimate of the spectral density matrix for a vector-valued stationary Gaussian process. In this paper the inverted complex Wishart distribution is studied and its use for the construction of spectral estimates is illustrated.

Methods of spectral estimation typically involve periodogram smoothing. The amount and type of smoothing one performs depend to a considerable extent upon prior knowledge of the spectral density to be estimated. A method

Received August 13, 1976; revised March 29, 1979.

AMS 1970 subject classifications: Primary 62H10; Secondary 62M15, 62E15.

Key words and phrases: Complex Wishart distribution, inverted complex Wishart distribution, multiple time series, spectral density, periodogram, prior distribution, posterior distribution.

* Research supported in part by the National Science Foundation under Grant No. MPS75-08286 and in part by the Office of Naval Research under Contract No. N00014-76-C-0930.

0047-259X/80/010051–09$02.00/0

of incorporating prior information about the shape and smoothness of a spectral density into the formation of a spectral estimate has been given by Shaman [9] for a univariate time series. Two types of finite-dimensional parameters are considered, the spectral density ordinates at a specified collection of frequencies and the amount of power in each of a set of frequency bands. The method is conditional upon the asymptotic distribution of nonoverlapping periodogram averages. A formal procedure applies Bayes theorem, with a conjugate prior distribution for a single parameter being an inverted gamma distribution. The mean of the posterior distribution involves a simple linear adjustment of the periodogram average, with coefficients depending upon prior distribution parameters. The prior distributions corresponding to different parameters, as well as the posteriors, are independent. Although the method is not genuinely Bayesian, it does permit one to incorporate prior information about the height and shape of the spectral density into the construction of an estimate in a formal manner.

The spectral density estimation methodology just discussed is extended to a vector time series model in the present paper. The asymptotic distribution of a periodogram average is a complex Wishart distribution. A conjugate prior distribution is an inverted complex Wishart distribution.

In Section 2 the density of the inverted complex Wishart distribution will be derived, as well as some marginals of the distribution and its first- and second-order moments. Details of the proposed use of the inverted complex Wishart distribution in spectral estimation are given in Section 3.

2. The Inverted Complex Wishart Distribution

Let $X_1, ..., X_n$ be independent $r \times 1$ vectors, each complex normal with mean 0 and Hermitian covariance matrix Σ (see Wooding [13] and Goodman [5]). Then the $r \times r$ matrix

$$W = \sum_{j=1}^{n} X_j X_j^*,$$

where the asterisk designates conjugate transpose, has a complex Wishart distribution with n degrees of freedom and covariance matrix Σ, denoted $W_C(r, n, \Sigma)$. The density is (Goodman [5])

$$\frac{1}{\tilde{\Gamma}_r(n) \mid \Sigma \mid^n} \mid W \mid^{n-r} etr(-\Sigma^{-1} W), \qquad n \geqslant r, \quad W \geqslant 0, \tag{1}$$

where

$$\tilde{\Gamma}_r(n) = \pi^{(1/2)r(r-1)} \prod_{j=1}^{r} \Gamma(n - j + 1)$$

is the complex multivariate gamma function. The mean is $E(W) = n\Sigma$ and the covariance matrix may be expressed as $\mathrm{Cov}(\mathrm{vec}\, W) = \Sigma' \otimes \Sigma$. The convention used for the covariance between complex random variables y and z is $E(y - E(y))\overline{(z - E(z))}$. The covariance result cited is known (e.g., see Brillinger [2, Problem 8.16.46], who uses the notation $B \otimes A$ for the Kronecker product written here as $A \otimes B$).

The Jacobian of the transformation $Y = X^{-1}$ for an $r \times r$ Hermitian nonsingular matrix X is $| Y |^{-2r}$. This may be deduced from (2.8) of Khatri [7]. Then the density of $V = W^{-1}$ is

$$\frac{| \Psi |^n}{\tilde{\Gamma}_r(n)} \frac{etr(-V^{-1}\Psi)}{| V |^{n+r}}, \qquad n \geqslant r, \quad V > 0, \tag{2}$$

where $\Psi = \Sigma^{-1}$. Denote the distribution of V by $W_C^{-1}(r, n, \Psi)$. The marginal distribution of V_{11}, the $q \times q$ upper left-hand corner of V, is $W_C^{-1}(q, n - r + q, \Psi_{11})$, where Ψ_{11} is the corresponding submatrix of Ψ. This may be derived from the method used by Tiao and Zellner [10] to obtain marginals of the inverted Wishart distribution, modified for the present complex case. When $q = 1$, V_{11} has an inverted gamma distribution.

Let the $r \times r$ Hermitian matrix W be written as TT^*, where T is lower triangular and has real diagonal entries. Then the density of T when $\Sigma = I$ is (Goodman [5], p. 165)

$$\prod_{j=1}^{r} \frac{2}{\Gamma(n - j + 1)} \exp(-t_{jj}^2)\, t_{jj}^{2n-2j+1} \prod_{j=2}^{r} \prod_{k=1}^{j-1} \frac{1}{\pi} \exp(- | t_{jk} |^2), \tag{3}$$

which is the density of $\frac{1}{2}r(r + 1)$ independent random variables.

The inverted complex Wishart matrix $V = W^{-1}$ is S^*S, where $S = (s_{jk}) = T^{-1}$ is lower triangular. In terms of elements of T,

$$s_{jj} = \frac{1}{t_{jj}}, \qquad j = 1,..., r, \tag{4}$$

$$\begin{aligned} s_{jk} = \frac{1}{t_{kk}} \Bigg(&-u_{jk} + \sum_{l_1=k+1}^{j-1} u_{jl_1} u_{l_1 k} - \sum_{l_2=k+1}^{j-2} \sum_{l_1=l_2+1}^{j-1} u_{jl_1} u_{l_1 l_2} u_{l_2 k} \\ &+ \cdots + (-1)^{j-k} u_{j,j-1} u_{j-1,j-2} \cdots u_{k+1,k} \Bigg), \qquad j > k, \end{aligned} \tag{5}$$

where

$$u_{jk} = t_{jk}/t_{jj}, \qquad j > k. \tag{6}$$

The elements of V are

$$v_{jk} = \sum_{g=\max(j,k)}^{r} \bar{s}_{gj} s_{gk}, \qquad j \geqslant k. \tag{7}$$

This decomposition of W as TT^* is similar to the one indicated by Brillinger [2, problem 4.8.27; a factor $\frac{1}{2}$ is missing]. For W a real Wishart matrix Kaufman [6] uses a corresponding decomposition to deduce first and second moments of W^{-1}.

Now we consider first- and second-order moments of $V = W^{-1}$. Let $E_\Sigma(\cdot)$ and $\mathrm{Cov}_\Sigma(\cdot)$ denote expectation and covariance when W is $W_C(r, n, \Sigma)$. Assume that Σ is nonsingular and let $\Sigma^{1/2}$ designate the Hermitian square root of Σ. We shall use the result that $\Sigma^{-1/2}W\Sigma^{-1/2}$ is distributed as $W_C(r, n, I)$. Moreover, if W is $W_C(r, n, I)$, then W and UWU^* have the same distribution for any $r \times r$ unitary matrix U. The latter implies $UE_I(W^{-1}) = E_I(W^{-1})\,U$ for every $r \times r$ unitary matrix U. Thus $E_I(W^{-1})$ must be a scalar multiple of the identity matrix and it suffices to calculate, for $\Sigma = I$,

$$E(v_{rr}) = \frac{1}{n-r}, \qquad n > r,$$

by (3), (4), and (7). Then

$$\begin{aligned} E_\Sigma(W^{-1}) &= \Sigma^{-1/2}E_\Sigma(\Sigma^{1/2}W^{-1}\Sigma^{1/2})\,\Sigma^{-1/2} = \Sigma^{-1/2}E_I(W^{-1})\,\Sigma^{-1/2} \\ &= \frac{1}{n-r}\,\Sigma^{-1}, \qquad n > r. \end{aligned} \tag{8}$$

This result was given by Wahba [11].

To deduce the covariance matrix of $\mathrm{vec}(W^{-1})$, note that for every $r \times r$ unitary matrix U

$$\begin{aligned} \mathrm{Cov}_I[\mathrm{vec}(W^{-1})] &= \mathrm{Cov}_I[(\bar{U} \otimes U)\,\mathrm{vec}(W^{-1})] \\ &= (\bar{U} \otimes U)\,\mathrm{Cov}_I[\mathrm{vec}(W^{-1})](U' \otimes U^*), \end{aligned}$$

or

$$(\bar{U} \otimes U)\,\mathrm{Cov}_I[\mathrm{vec}(W^{-1})] = \mathrm{Cov}_I[\mathrm{vec}(W^{-1})](\bar{U} \otimes U).$$

This implies

$$\mathrm{Cov}_I[\mathrm{vec}(W^{-1})] = b(I_r \otimes I_r) + c\,\mathrm{vec}\,I_r(\mathrm{vec}\,I_r)' + (a - b - c)\,J,$$

where a, b, and c are real constants, I_r is the $r \times r$ identity matrix, and J is an $r^2 \times r^2$ diagonal matrix with ones in diagonal positions $1 + (r+1)j$, $j = 0, 1, \ldots, r-1$, and zeros elsewhere. It is therefore sufficient to determine for $\Sigma = I$, from (3)–(8),

$$\mathrm{Var}(v_{rr}) = \frac{1}{(n-r)^2(n-r-1)}, \qquad n > r+1,$$

$$\mathrm{Var}(v_{r-1,r}) = \frac{1}{(n-r+1)(n-r)(n-r-1)}, \qquad n > r+1,$$

and

$$\mathrm{Cov}(v_{r-1,r-1}, v_{rr}) = \frac{1}{(n-r+1)(n-r)^2(n-r-1)}, \qquad n > r+1.$$

Then

$$\begin{aligned}
\mathrm{Cov}_\Sigma[\mathrm{vec}(W^{-1})] &= \mathrm{Cov}_\Sigma[\mathrm{vec}(\Sigma^{-1/2}(\Sigma^{1/2}W^{-1}\Sigma^{1/2})\,\Sigma^{-1/2})] \\
&= \mathrm{Cov}_\Sigma[((\Sigma^{-1/2})' \otimes \Sigma^{-1/2})\,\mathrm{vec}(\Sigma^{1/2}W^{-1}\Sigma^{1/2})] \\
&= [(\Sigma^{-1/2})' \otimes \Sigma^{-1/2}]\,\mathrm{Cov}_I[\mathrm{vec}(W^{-1})][(\Sigma^{-1/2})' \otimes \Sigma^{-1/2}] \\
&= \frac{(\Sigma^{-1})' \otimes \Sigma^{-1} + 1/(n-r)\,\mathrm{vec}(\Sigma^{-1})(\mathrm{vec}(\Sigma^{-1})')'}{(n-r+1)(n-r)(n-r-1)}, \\
&\qquad\qquad n > r+1. \qquad (9)
\end{aligned}$$

The inverse of (9), or the precision matrix, is

$$(n-r+1)(n-r)(n-r-1)\left[\Sigma' \otimes \Sigma - \frac{1}{n}\,\mathrm{vec}\,\Sigma(\mathrm{vec}(\Sigma'))'\right].$$

3. Application of the Inverted Complex Wishart Distribution to Spectral Estimation

The methodology described below is based upon distributional approximations and we make assumptions which allow these to hold. The conditions in Brillinger [1] and [2], Chapters 5 and 7, in particular, are used.

Let $X(t)$, $t = 0, \pm 1, \ldots$, be a vector-valued strictly stationary stochastic process for which all moments exist. Denote the components of $X(t)$ by $X_j(t)$, $j = 1, \ldots, r$, the mean by $E\{X(t)\} = m$, and the spectral density by $f(\lambda)$, $-\pi \leqslant \lambda \leqslant \pi$. The cumulant functions of the process are

$$C_{j_1 \cdots j_k}(t_1, \ldots, t_{k-1}) = cum\{X_{j_1}(t_1 + t), \ldots, X_{j_{k-1}}(t_{k-1} + t), X_{j_k}(t)\},$$

$$j_i = 1, \ldots, r,\ i = 1, \ldots, k, \qquad t_1 + t, \ldots, t_{k-1} + t,\ t = 0, \pm 1, \ldots,\ k = 2, 3, \ldots.$$

Assume for $j_i = 1, \ldots, r$, $i = 1, \ldots, k$, that

$$\sum_{t_1, \ldots, t_{k-1} = -\infty}^{\infty} | C_{j_1, \ldots, j_k}(t_1, \ldots, t_{k-1})| < \infty, \qquad k = 2, 3, \ldots. \qquad (10)$$

This ensures the existence and uniform continuity of cumulant spectra of all orders.

Assume a time series $X(t)$, $t = 0, 1, \ldots, T-1$, is available. The periodogram is

$$I(\lambda) = \frac{1}{2\pi T} Z(\lambda)Z(\lambda)^*, \qquad -\pi \leqslant \lambda \leqslant \pi, \qquad (11)$$

where $Z(\lambda) = \sum_{t=0}^{T-1} e^{i\lambda t}X(t)$, $-\pi \leqslant \lambda \leqslant \pi$. Let $p = [\frac{1}{2}(T-1)]$. Then $I(2\pi j/T)$ are asymptotically independent variables distributed as $W_C\{r, 1, f(2\pi j/T)\}$, $j = 1,..., p$ (see Brillinger [2], Theorem 7.2.4). Also $I(\pi)$ is asymptotically an $r \times r$ Wishart variable with one degree of freedom and covariance matrix $f(\pi)$, and is independent of the other variables. If $m \neq 0$, $I(0)$ is approximately an $r \times r$ noncentral Wishart variable with one degree of freedom.

Restrict attention to frequencies $0 \leqslant \lambda \leqslant \pi$ and let $j(T)$ be a sequence of integers such that $2\pi j(T)/T$ is near $\lambda(\neq 0, \pi)$ and converges to λ as $T \to \infty$. Then

$$z = \frac{1}{2n+1} \sum_{h=-n}^{n} I[2\pi\{j(T)+h\}/T] \tag{12}$$

is an estimate of $f(\lambda)$ and is asymptotically distributed as $(2n+1)^{-1} \times W_C\{r, 2n+1, f(\lambda)\}$. If $\lambda = 0$ (12) is replaced by

$$z = \frac{1}{n} \sum_{h=1}^{n} I(2\pi h/T), \tag{13}$$

and if $\lambda = \pi$,

$$\begin{aligned} z &= \frac{1}{n} \sum_{h=1}^{n} I(\pi - 2\pi h/T) && (T \text{ even}) \\ &= \frac{1}{n} \sum_{h=1}^{n} I(\pi - \pi/T - 2\pi h/T) && (T \text{ odd}) \end{aligned} \tag{14}$$

is used. The complex Wishart was established as a limiting distribution for (11) and (12) with fixed n by Brillinger [1] for the case $m = 0$. For $\lambda \neq 0$, the asymptotic distributions of $I(\lambda)$ and z do not depend on m and in [2] Brillinger treats an arbitrary m. Wahba [12] and Gleser and Pagano [3] allow $n \to \infty$ under the assumption $X(t)$ is Gaussian. Under appropriate conditions, M nonoverlapping sums of the form (12) are asymptotically independent complex Wishart matrices as n, M, $T \to \infty$. The covariance matrices of the asymptotic complex Wishart distributions are $f[2\pi j(T)/T]$, where $2\pi j(T)/T$ converges to some λ as $T \to \infty$.

Consider estimation of the spectral density at a fixed, preassigned set of frequencies, $0 \leqslant \lambda_1 < \cdots < \lambda_M \leqslant \pi$. The choice of M and the frequencies may involve use of prior information. For example, if the spectral density is considered a priori to be approximately constant in certain bands, the frequencies may be interior points of the bands.

To avoid anomalous cases assume $\lambda_1 > 0$, $\lambda_M < \pi$. Since the spectral estimates are asymptotically independent and the priors will also be independent, it suffices to discuss a single frequency, labelled λ for simplicity, in the preassigned set. We use the asymptotic distribution described above of z in (12). Then z has density

$$h(z \mid f(\lambda)) = \frac{(2n+1)^{(2n+1)r}}{\tilde{\Gamma}_r(2n+1)} \frac{|z|^{2n+1-r}}{|f(\lambda)|^{2n+1}} \, etr\{-f(\lambda)^{-1}(2n+1)\, z\}, \qquad 2n+1 \geqslant r. \tag{15}$$

A conjugate prior density is an inverted complex Wishart,

$$h(f(\lambda)) = \frac{|B|^{\alpha} \, etr\{-f(\lambda)^{-1} B\}}{\tilde{\Gamma}_r(\alpha) \, |f(\lambda)|^{\alpha+r}}, \qquad \alpha \geqslant r, \quad B \geqslant 0, \tag{16}$$

where B is Hermitian. The posterior density from (15) and (16) is

$$h(f(\lambda) \mid z) = \frac{|(2n+1)\, z + B|^{2n+1+\alpha} \, etr[-f(\lambda)^{-1}\{(2n+1)\, z + B\}]}{\tilde{\Gamma}_r(2n+1+\alpha) \, |f(\lambda)|^{2n+r+1+\alpha}}. \tag{17}$$

The mean of the posterior occurs at $f(\lambda) = \{(2n+1)\, z + B\}/(2n+1+\alpha-r)$.

Knowledge of the spectral density ordinate at each of a number of specified frequencies can convey an accurate picture of the shape of the curve. However, more basic interest may concern the increments of the spectral distribution function in certain frequency bands. Consider a partition $0 = \lambda_0 < \lambda_1 < \cdots < \lambda_M < \lambda_{M+1} = \pi$ and let $F(\lambda)$ denote the spectral distribution function. Consider the parameter $p = (p_1, ..., p_{M+1})$, where $p_l = F(\lambda_l) - F(\lambda_{l-1})$, $l = 1, ..., M+1$. The partition is fixed for all sample sizes. Let $k_l(T)$, $l = 0, ..., M+1$, $k_0(T) = 1$, $k_{M+1}(T) = [\frac{1}{2}(T-1)]$, be integers such that the frequencies $2\pi j/T$ are in the interval $(\lambda_{l-1}, \lambda_l)$ for $k_{l-1}(T) \leqslant j \leqslant k_l(T) - 1$ and define $m_l(T) = k_l(T) - k_{l-1}(T)$. Then under the assumption (10) and the conditions in Wahba [12] or Gleser and Pagano [3] the sums

$$y_l = \sum_{j=k_{l-1}(T)}^{k_l(T)-1} I(2\pi j/T), \qquad l = 1, ..., M+1,$$

are approximately distributed as independent $r \times r$ complex Wishart variables with $m_l(T)$ degrees of freedom and covariance matrices

$$\frac{1}{m_l(T)} \sum_{j=k_{l-1}(T)}^{k_l(T)-1} f(2\pi j/T), \qquad l = 1, ..., M+1.$$

We further approximate the distribution of y_l as that of $W_C\{r, m_l(T), Tp_l/(2\pi m_l(T))\}$, $l = 1, ..., M+1$. Details of the transition to a posterior distribution for the parameter p are similar to those given at (15)–(17). One can restrict attention to a set of frequency bands whose union forms a subset of $[0, \pi]$.

Selection of a prior distribution for spectral estimation may be based upon first- and second-order moments. For illustration let $r = 2$ and denote the spectral and cross spectral densities by $f_{11}(\lambda)$, $f_{22}(\lambda)$, and $f_{12}(\lambda) = c(\lambda) + iq(\lambda)$. For a given frequency λ the moments of the prior (16), with $B = (\beta_{jkR} + i\beta_{jkI})$, may be written as

$$E\begin{pmatrix} f_{11}(\lambda) & c(\lambda) \\ q(\lambda) & f_{22}(\lambda) \end{pmatrix} = \frac{1}{\alpha - 2}\begin{pmatrix} \beta_{11} & \beta_{12R} \\ \beta_{12I} & \beta_{22} \end{pmatrix}$$

and

$$\operatorname{Cov}\begin{pmatrix} f_{11}(\lambda) \\ f_{22}(\lambda) \\ c(\lambda) \\ q(\lambda) \end{pmatrix} = \frac{1}{(\alpha - 3)(\alpha - 2)^2}$$

$$\times \begin{bmatrix} \beta_{11}^2 & & & \\ \dfrac{\beta_{11}\beta_{22}}{\alpha - 1} + \dfrac{\alpha - 2}{\alpha - 1}\,|\beta_{12}|^2 & \beta_{22}^2 & & \\ \beta_{11}\beta_{12R} & \beta_{22}\beta_{12R} & \dfrac{\alpha\beta_{12R}^2 + (\alpha - 2)(\beta_{11}\beta_{22} - \beta_{12I}^2)}{2(\alpha - 1)} & \\ \beta_{11}\beta_{12I} & \beta_{22}\beta_{12I} & \beta_{12R}\beta_{12I} & \dfrac{\alpha\beta_{12I}^2 + (\alpha - 2)(\beta_{11}\beta_{22} - \beta_{12R}^2)}{2(\alpha - 1)} \end{bmatrix}$$

Thus $(\alpha - 2)^{-1} B$ may be chosen to match prior opinion about the value of $f(\lambda)$, and the magnitude of α should reflect the firmness of this opionion.

Acknowledgment

The author is grateful to M. L. Eaton, who indicated the use of a unitary matrix to deduce the structure of $E_I(W^{-1})$ and of $\operatorname{Cov}_I[\operatorname{vec}(W^{-1})]$.

References

[1] Brillinger, D. R. (1969). Asymptotic properties of spectral estimates of second order. *Biometrika* **56** 375–390.

[2] Brillinger, D. R. (1975). *Time Series Data Analysis and Theory.* Holt, Rinehart & Winston, New York.

[3] Gleser, L. J., and Pagano, M. (1973). Approximating circulant quadratic forms in jointly stationary Gaussian time series. *Ann. Statist.* **1** 322–333.

[4] Goodman, N. R. (1957). *On the Joint Estimation of the Spectra, Cospectrum and Quadrature Spectrum of a Two-Dimensional Stationary Gaussian Process.* Ph. D. Thesis, Princeton University.

[5] Goodman, N. R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). *Ann. Math. Statist.* **34** 152–177.

[6] KAUFMAN, G. (1967). *Some Bayesian Moment Formulae.* Report No. 6710, Center for Operations Research and Econometrics, Catholic University of Louvain.
[7] KHATRI, C. G. (1965). Classical statistical analysis based on a certain multivariate complex Gaussian distribution. *Ann. Math. Statist.* **36** 98–114.
[8] KRISHNAIAH, P. R. (1976). Some recent developments on complex multivariate distributions. *J. Multivariate Anal.* **6** 1–30.
[9] SHAMAN, P. (1977). Some Bayesian considerations in spectral estimation. *Biometrika* **64** 79–84.
[10] TIAO, G. C., AND ZELLNER, A. (1964). On the Bayesian estimation of multivariate regression. *J. Roy. Statist. Soc. Ser. B* **26** 277–285.
[11] WAHBA, G. (1966). *Cross Spectral Distribution Theory for Mixed Spectra and Estimation of Prediction Filter Coefficients.* Ph. D. Thesis, Stanford University.
[12] WABHA, G. (1968). On the distribution of some statistics useful in the analysis of jointly stationary time series. *Ann. Math. Statist.* **39** 1849–1862.
[13] WOODING, R. A. (1956). The multivariate distribution of complex normal variables. *Biometrika* **43** 212–215.