The Multivariate Complex Normal Distribution—A Generalization

A. van den Bos

***Abstract*—The multivariate complex normal distribution usually employed in the literature is a special case since certain restrictions have been imposed on the covariances of the real and imaginary parts of its variables. A more general distribution is proposed of which the usual distribution is shown to be a special case.**

***Index Terms*— Normal distribution, complex distributions, complex stochastic variables.**

I. Introduction

Since its introduction [1], the multivariate complex normal distribution employed in the literature has been a special case: the covariance matrix associated with it satisfies a number of restrictions in addition to those required to guarantee positive semidefiniteness. The reason given in [1] for these restrictions is closely connected with the particular application studied. It concerns a complex valued stochastic process with an in-phase real part and a quadrature imaginary part. Other authors adopted this distribution in subsequent papers. For a review, see Miller [2]. In the same paper, Miller states that the restrictions imposed on the covariance matrix are an automatic consequence of defining the complex distribution in a manner analogous to the real case.

These developments have probably convinced later authors that this specialized complex normal distribution is the most general one. In any event, this is the only form found by the author in the recent literature. For an example, see [3, p. 110]. Thus the specialized distribution has become the generally accepted complex normal distribution. The purpose of this correspondence is to show that there is a more general alternative.

In the next section, the usual, specialized, form of the complex normal distribution is reviewed. The more general alternative is proposed in Section III. In Section IV, it is shown that the usual description is a specialized case of this alternative distribution. In addition, an important special case, the univariate complex normal distribution, is discussed.

II. The Usual Description of the Complex Normal Distribution

The usual definition of the complex normal distribution is [1], [3, p. 110], [4, p. 77], [5]

$$\frac{1}{\pi^N \det Z} \exp(-z^H Z^{-1} z) \tag{1}$$

where $z \in C^{N\times 1}$ is a vector of complex stochastic variables defined as

$$z = (z_1 \cdots z_N)^T \tag{2}$$

with

$$z_n = x_n + jy_n \tag{3}$$

where $x_n, y_n \in R^1$. To simplify the notation, it will be assumed that

$$E[z_n] = E[x_n] + jE[y_n] = 0 \tag{4}$$

Manuscript received May 31, 1994; revised September 9, 1994.
The author is with the Department of Applied Physics, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands.
IEEE Log Number 9408289.

for all n, where $E[\ ]$ is the expectation operator. In these expressions, the superscript T denotes transposition, the superscript H denotes complex conjugate transposition, and $j = \sqrt{-1}$. Furthermore, the matrix $Z \in C^{N\times N}$ is the complex covariance matrix of z defined as

$$Z = E[(z - E[z])(z - E[z])^H]. \tag{5}$$

Since it has been assumed that $E[z]$ is equal to zero, this simplifies to $Z = E[zz^H]$. By definition, Z is positive semidefinite and Hermitian symmetric. By assumption, it will be positive definite, hence, its inverse exists.

Substitution of (3) for the z_n in (1) produces the real multivariate normal distribution of the elements x_n and y_n of the vector $w \in R^{2N\times 1}$ defined as

$$w = (x_1\, y_1 \cdots x_N\, y_N)^T. \tag{6}$$

The expression for the covariance matrix associated with this distribution shows that [1], [2], [4, p. 79], [5]

$$E[x_n x_m] = E[y_n y_m] \tag{7}$$

and

$$E[x_n y_m] = -E[x_m y_n] \tag{8}$$

for $n, m = 1, \cdots, N$. By (7) the covariance of x_n and x_m is equal to the covariance of y_n and y_m for all n and m. This implies that the variance of x_n is equal to the variance of y_n. By (8) the covariance of x_n and y_m is the additive inverse of the covariance of x_m and y_n. This implies that the covariance of x_n and y_n is equal to zero. For what follows it is useful to notice that (7) and (8) also imply that

$$E[z_n z_m] = 0. \tag{9}$$

From these considerations it is clear that the chosen definition (1) for the multivariate distribution of the complex normal variables z_n is very restrictive with respect to the allowable covariance matrices of the real parts x_n and the imaginary parts y_n.

III. A More General Alternative

Let again $z_n = x_n + jy_n$, $n = 1, \cdots, N$ be complex stochastic variables and suppose that the real quantities x_n and y_n, $n = 1, \cdots, N$ are normally distributed. Consider the vector $w \in R^{2N\times 1}$

$$w = (x_1\, y_1 \cdots x_N\, y_N)^T \tag{10}$$

and suppose that $E[x_n] = E[y_n] = 0$ for all n. Therefore, the matrix $W \in R^{2N\times 2N}$

$$W = E[ww^T] \tag{11}$$

is the covariance matrix of w. Then the distribution of the x_n and y_n is described by

$$f(w) = \frac{1}{(2\pi)^N (\det W)^{1/2}} \exp(-1/2 w^T W^{-1} w). \tag{12}$$

If

$$z_n^* = x_n - jy_n \tag{13}$$

is defined as the complex conjugate of z_n it follows from (3) and (13) that

$$\begin{pmatrix} z_n \\ z_n^* \end{pmatrix} = J \begin{pmatrix} x_n \\ y_n \end{pmatrix} \tag{14}$$

where the matrix $J \in C^{2\times 2}$ is defined as

$$J = \begin{pmatrix} 1 & j \\ 1 & -j \end{pmatrix}. \tag{15}$$

For what follows it is important to note that

$$J^{-1} = 1/2J^H. \tag{16}$$

Next define the vector $v \in C^{2N\times 1}$ as

$$v = (z_1\, z_1^* \cdots z_N\, z_N^*)^T. \tag{17}$$

Then, by (14)

$$v = Aw \tag{18}$$

where the block-diagonal matrix $A \in C^{2N\times 2N}$ is defined as

$$A = \operatorname{diag}(J \cdots J). \tag{19}$$

Hence, by (16)

$$w = A^{-1}v = 1/2A^H v \tag{20}$$

and since w is real

$$w = 1/2A^T v^*. \tag{21}$$

The covariance matrix $V \in C^{2N\times 2N}$ of v is defined as

$$V = E[vv^H]. \tag{22}$$

Hence, by (18)

$$V = AWA^H. \tag{23}$$

Next consider the quadratic form in (12):

$$w^T W^{-1} w. \tag{24}$$

By (20), (21), and (23) this form may be written

$$\begin{aligned} v^H(1/2A)W^{-1}(1/2A^H)v &= v^H(AWA^H)^{-1}v \\ &= v^H V^{-1} v. \end{aligned} \tag{25}$$

Furthermore, the determinant in (12) may be written

$$\begin{aligned} \det W &= \det(1/2A^H V 1/2A) \\ &= (1/2)^{2N}(\det J^H)^N \det V (1/2)^{2N}(\det J)^N \\ &= (1/2)^{4N}(2j)^N(-2j)^N \det V = (1/2)^{2N}\det V. \end{aligned} \tag{26}$$

Substitution of (25) and (26) in (12) yields

$$f(v) = \frac{1}{\pi^N(\det V)^{1/2}}\exp(-1/2v^H V^{-1} v) \tag{27}$$

with v defined by (17). This is the main result of this correspondence. It has been derived without any assumption with respect to the covariance matrix W of the x_n and y_n with the exception of nonsingularity.

IV. Special Cases

Under the restrictions (7) and (8) the distribution (27) should produce the distribution (1) as a special case. To show this the elements of v, defined by (17), are first rearranged as follows

$$u = Pv = (z_1 \cdots z_N\; z_1^* \cdots z_N^*)^T \tag{28}$$

where $P \in R^{2n\times 2N}$ is the appropriate permutation matrix. The elements of a permutation matrix are either equal to one or to zero. In addition, exactly one element of each row and column is equal to one. A permutation matrix is orthogonal [6, p. 360] and the absolute value of its determinant is equal to one [6, p. 25]. That is

$$P^T = P^{-1} \tag{29}$$

and

$$|\det P| = 1. \tag{30}$$

Using (29) the following transformation of the quadratic form in (27) may be carried out:

$$\begin{aligned} v^H V^{-1} v &= u^H P V^{-1} P^T u \\ &= u^H (PVP^T)^{-1} u = u^H U^{-1} u. \end{aligned} \tag{31}$$

In this expression, the matrix $U \in C^{2N\times 2N}$ defined as

$$U = PVP^T \tag{32}$$

is, by (28), equal to the covariance matrix of u. But since, by (9), $E[z_m z_n] = 0$

$$U = \operatorname{diag}(ZZ^*) \tag{33}$$

where the matrix $Z \in C^{N\times N}$ is the covariance matrix Z of $z = (z_1 \cdots z_N)^T$. From (28), (31), and (33) it is concluded that

$$\begin{aligned} v^H V^{-1} v &= z^H Z^{-1} z + z^T (Z^*)^{-1} z^* \\ &= 2z^H Z^{-1} z. \end{aligned} \tag{34}$$

Furthermore, from (29), (30), (32), and (33)

$$\begin{aligned} \det V &= \det P^T \det U \det P \\ &= \det U = \det Z \det Z^* = (\det Z)^2. \end{aligned} \tag{35}$$

Substituting (34) and (35) in (27) yields (1). Therefore, it is concluded that the usual distribution (1) is a special case of the proposed general distribution (27).

As a second special case consider (27) for $N = 1$. Denote $z_1 = z$, $x_1 = x$, and $y_1 = y$. Then

$$V = \begin{pmatrix} C_{zz^*} & C_{zz} \\ C_{z^*z^*} & C_{z^*z} \end{pmatrix} \tag{36}$$

where $C_{zz^*} = E[zz^*]$ and the other elements are defined accordingly. Hence

$$V^{-1} = \frac{1}{\det V}\begin{pmatrix} C_{z^*z} & -C_{zz} \\ -C_{z^*z^*} & C_{zz^*} \end{pmatrix} \tag{37}$$

with

$$\det V = C_{zz^*}C_{z^*z} - C_{zz}C_{z^*z^*}. \tag{38}$$

It is observed that $C_{zz^*} = C_{z^*z} = E[|z|^2] = \sigma_z^2 = E[x^2] + E[y^2] = \sigma_x^2 + \sigma_y^2$ and $C_{zz} = (C_{z^*z^*})^* = E[z^2] = \sigma_x^2 - \sigma_y^2 + 2j\rho_{xy}\sigma_x\sigma_y$ where $\rho_{xy} = E[xy]/\sigma_x\sigma_y$. Therefore

$$V^{-1} = \frac{1}{\det V}\begin{pmatrix} \sigma_x^2 + \sigma_y^2 & -\sigma_x^2 + \sigma_y^2 - 2j\rho_{xy}\sigma_x\sigma_y \\ -\sigma_x^2 + \sigma_y^2 + 2j\rho_{xy}\sigma_x\sigma_y & \sigma_x^2 + \sigma_y^2 \end{pmatrix} \tag{39}$$

with

$$\det V = 4\sigma_x^2\sigma_y^2(1 - \rho_{xy}^2). \tag{40}$$

Substitution of (39) and (40) in (27) yields $f(v)$ with $v = (x + jy\;\, x - jy)^T$. Notice that this expression applies to any $|\rho_{xy}| < 1$ and any σ_x and σ_y. If ρ_{xy} is equal to zero and σ_x^2 is equal to σ_y^2 the following expression is obtained:

$$f(z) = \frac{1}{\pi\sigma_z^2}\exp\left(-\frac{|z|^2}{\sigma_z^2}\right). \tag{41}$$

This is the usual expression found in the statistical literature [5], and in the signal processing literature [3, p. 110] for the univariate complex normal distribution.

REFERENCES

[1] R. A. Wooding, "The multivariate distribution of complex normal variables," *Biometrika*, vol. 43, pp. 212–215, 1956.
[2] K. S. Miller, "Complex Gaussian processes," *SIAM Rev.*, vol. 11, pp. 544–567, 1969.
[3] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.
[4] K. S. Miller, *Complex Stochastic Processes*. Reading, MA: Addison-Wesley, 1974.
[5] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution (An introduction)," *Annals Math. Statist.*, vol. 34, pp. 152–177, 1963.
[6] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge Univ. Press, 1992.

On Lower Bounds for the Smallest Eigenvalue of a Hermitian Positive-Definite Matrix

Evan M. Ma and Christopher J. Zarowski, *Senior Member, IEEE*

***Abstract*—This correspondence presents an improvement to Dembo's lower bound on the smallest eigenvalue of a Hermitian positive-definite matrix. Unlike Dembo's bound the improved bound is always positive.**

***Index Terms*—Hermitian positive-definite matrices, eigenvalue bounds.**

I. INTRODUCTION

Dembo [1] presents a set of upper and lower bounds on the largest and smallest eigenvalues of Hermitian positive-definite matrices. When employed with certain parameters generated by the well-known Levinson–Durbin algorithm [2], [3] these bounds may be applied to the important special case of Hermitian positive-definite Toeplitz matrices.

However, the lower bound on the smallest eigenvalue of a Hermitian positive-definite matrix as given in [1, Theorem 1] can at times be negative. We present here, based on [4], an improved lower bound that is never negative.

II. SUMMARY OF DEMBO'S BOUND

Consider the $n \times n$ Hermitian positive-definite matrix R_{n-1} in partitioned form, i.e.,

$$R_{n-1} = \begin{bmatrix} R_{n-2} & b_{n-1} \\ b_{n-1}^H & c \end{bmatrix} \tag{1}$$

where $c > 0$, and $b_{n-1}^H = [r_{n-1,0} \; r_{n-1,1} \; \cdots \; r_{n-1,n-2}]$. Because R_{n-1} is Hermitian, $R_{n-1} = R_{n-1}^H$, where the superscript H denotes Hermitian transposition.

According to Dembo [1, Theorem 1] we have

Theorem 2.1: Let λ_1 denote the smallest eigenvalue of R_{n-1}, then

$$\underline{\lambda}_1 = \frac{c+\eta_1}{2} - \sqrt{\frac{(c-\eta_1)^2}{4} + b_{n-1}^H b_{n-1}} \le \lambda_1 \tag{2}$$

where η_1 is the lower bound on the minimum eigenvalue of R_{n-2}.

Manuscript received May 1, 1994; revised June 22, 1994. This work was funded by the Canadian Institute for Telecommunications Research.

The authors are with the Department of Electrical and Computer Engineering, Queen's University, Kingston, Ont. K7L 3N6, Canada .

IEEE Log Number 9407922.

To see that the bound in (2) may be negative, consider the Toeplitz matrix of dimension three, and all the entries of which are unity. In this case $\underline{\lambda}_1 = -1$.

III. AN IMPROVED BOUND

Recall first of all that the eigenvalues of Hermitian matrix R_{n-1} are the roots of the secular equation [1], [5]

$$f(\lambda) = \lambda - c + \sum_{j=1}^{t} \frac{k_j^H k_j}{\tau_j - \lambda} \tag{3}$$

where t is the number of distinct eigenvalues of R_{n-1}, which is partitioned as in (1), and τ_j is an eigenvalue of R_{n-2} such that $\tau_1 < \tau_2 < \cdots < \tau_t$, and $k_j = (u_j^H b_{n-1})u_j$, with u_j being the eigenvector associated with τ_j.

With the aid of the secular equation (3), we have

Theorem 3.1:

$$\underline{\lambda}_1 = \frac{c+\eta_1}{2} - \sqrt{\frac{(c+\eta_1)^2}{4} - (c - b_{n-1}^H R_{n-2}^{-1} b_{n-1})\eta_1} \le \lambda_1 \tag{4}$$

where η_1 is the lower bound on the minimum eigenvalue of R_{n-2}.

Proof: Over the interval $\lambda \in [0, \tau_1)$, since $\tau_1 < \tau_2 < \cdots < \tau_t$, we have

$$(1-\lambda/\tau_1)^{-1} \ge (1-\lambda/\tau_2)^{-1} \ge \cdots \ge (1-\lambda/\tau_t)^{-1} \tag{5}$$

and, therefore, we have established

$$\sum_{j=1}^{t} \frac{k_j^H k_j}{\tau_j - \lambda} \le \frac{1}{1-\lambda/\tau_1} \sum_{j=1}^{t} \frac{k_j^H k_j}{\tau_j} = \frac{1}{1-\lambda/\tau_1} b_{n-1}^H R_{n-2}^{-1} b_{n-1} \tag{6}$$

where the equality follows from a consideration of the singular value decomposition of R_{n-2}^{-1}. Now define

$$g(\lambda) = \lambda - c + \frac{1}{1-\lambda/\tau_1} \sum_{j=1}^{t} \frac{k_j^H k_j}{\tau_j}. \tag{7}$$

Via (6) and the fact that $f(\lambda)$ is a monotonically increasing function over the interval $[0, \tau_1)$, $g(\lambda)$ is therefore an overbounding function of $f(\lambda)$ over the same interval. Since the eigenvalues of R_{n-1} satisfy the well-known interleaving property of eigenvalues of leading principal submatrices (i.e., $\lambda_1 < \tau_1 < \lambda_2 < \tau_2 < \cdots < \tau_t < \lambda_{t+1}$) for Hermitian matrices [5], the minimal root of $g(\lambda)$ can be used as a lower bound for λ_1. Letting $\underline{\lambda}_1$ be the minimal root of $g(\lambda)$ and setting the left side of (7) to zero, one obtains

$$\underline{\lambda}_1 = \frac{(\tau_1 + c) \pm \sqrt{(\tau_1+c)^2 - 4(c - b_{n-1}^H R_{n-2}^{-1} b_{n-1})\tau_1}}{2}. \tag{8}$$

By completing the square inside the square root of (8) we may see that

$$\tau_1 < \frac{(\tau_1 + c) + \sqrt{(\tau_1+c)^2 - 4(c - b_{n-1}^H R_{n-2}^{-1} b_{n-1})\tau_1}}{2}. \tag{9}$$

Thus the lower bound on λ_1 is as stated in (4). ■

There is an alternative proof based on consideration of the matrices

$$A = \begin{bmatrix} \sqrt{\eta_1} R_{n-2}^{-1/2} & 0 \\ 0 & 1 \end{bmatrix}$$
$$ABA = \begin{bmatrix} \eta_1 I_{n-1} & d \\ d^H & c \end{bmatrix} \tag{10}$$

where $B = R_{n-1}$, and $d^H d = \eta_1 b_{n-1}^H R_{n-2}^{-1} b_{n-1}$. If A and $I - A$ are nonnegative-definite then the eigenvalues of ABA lower bound