On the Maximum Entropy Theorem for Complex Random Vectors

Georg Tauböck
Telecommunications Research Center Vienna (ftw.)
Tech Gate Vienna, Donau-City-Strasse 1/3, A-1220 Vienna, Austria
e-mail: tauboeck@ftw.at

Abstract — **We consider complex random vectors and study some important properties. We develop a theory which is based on the concept of covariance and *pseudo-covariance* matrix in order to prove a stronger version of the *Maximum Entropy Theorem* for the complex multivariate case [1].**

I. Introduction

It is well known, c.f., [1] or [2], that a complex random vector $\mathbf{x} \in \mathbb{C}^n$ with a given non-singular covariance matrix $\mathbf{C_x} = \mathcal{E}\{(\mathbf{x}-\mu_\mathbf{x})(\mathbf{x}-\mu_\mathbf{x})^\mathbf{H}\}$, where $\mu_\mathbf{x} = \mathcal{E}\{\mathbf{x}\}$ denotes the mean / expectation vector[1], has a *differential entropy* $h(\mathbf{x})$ which is upper bounded by an expression that depends only on the covariance matrix, i.e.,

$$h(\mathbf{x}) \le \log\det(\pi e \mathbf{C_x})$$

with equality if and only if $\mathbf{x}$ is *rotationally invariant* and Gaussian. A complex random vector is called *rotationally invariant* (c.f., [4]) or *proper* (c.f., [1]) or *circularly symmetric* (c.f., [2]) if its *pseudo-covariance* matrix $\mathbf{P_x} = \mathcal{E}\{(\mathbf{x}-\mu_\mathbf{x})(\mathbf{x}-\mu_\mathbf{x})^\mathbf{T}\}$ vanishes.

Recently [4, 5, 6] it was found that there are (practical) situations where complex random vectors have a non-vanishing pseudo-covariance matrix, which can then be utilized to achieve (significant) performance gains. Unfortunately the upper bound above is not tight in this case and therefore the benefit of this *Maximum Entropy Theorem* is limited. We will generalize this theorem such that it takes into account covariance **and** pseudo-covariance matrix, in order to achieve a tighter upper bound.

II. Preliminaries

Definition 1 *A matrix* $\mathbf{B} \in \mathbb{C}^{n\times n}$ *is called* generalized Cholesky factor *of a positive definite Hermitian matrix* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *if it satisfies the equation*

$$\mathbf{A} = \mathbf{B}\mathbf{B}^\mathbf{H}.$$

Since $\det\mathbf{A} = |\det\mathbf{B}|^2$, a generalized Cholesky factor is always a non-singular matrix. Note that the conventional Cholesky decomposition (c.f., [3]), $\mathbf{A} = \mathbf{L}\mathbf{L}^\mathbf{H}$, where $\mathbf{L}$ is lower-triangular, yields a generalized Cholesky factor $\mathbf{L}$. Another possibility to construct a generalized Cholesky factor is based on the eigenvalue decomposition of $\mathbf{A}$. The next theorem tells us that if we know one generalized Cholesky factor, we know them all:

[1] Usually the random vector is assumed to be zero-mean.

Theorem 1 *Suppose* $\mathbf{B}$ *is a generalized Cholesky factor of* $\mathbf{A}$. *Then, for any unitary matrix* $\mathbf{U}$, $\mathbf{C} = \mathbf{B}\mathbf{U}$ *is also a generalized Cholesky factor of* $\mathbf{A}$. *Conversely, if* $\mathbf{B}$ *and* $\mathbf{C}$ *are generalized Cholesky factors, there exists a unitary matrix* $\mathbf{U}$, *such that* $\mathbf{C} = \mathbf{B}\mathbf{U}$.

The concept of generalized Cholesky factors enables us to formulate a criterion for a matrix to be a pseudo-covariance matrix:

Theorem 2 *Let* $\mathbf{C} \in \mathbb{C}^{n\times n}$ *be an Hermitian positive definite matrix and* $\mathbf{B}$ *a generalized Cholesky factor of* $\mathbf{C}$. $\mathbf{C}$ *and a matrix* $\mathbf{P} \in \mathbb{C}^{n\times n}$ *are covariance matrix and pseudo-covariance matrix of a complex random vector, respectively, if and only if* $\mathbf{P}$ *is symmetric and the singular values(c.f., [3]) of* $\mathbf{B}^{-1}\mathbf{P}\mathbf{B}^{-\mathbf{T}}$ *are smaller or equal to* 1.

III. Generalized Maximum Entropy Theorem

Using the theory of generalized Cholesky factors we can prove the following main result:

Theorem 3 *Suppose the complex random vector* $\mathbf{x} \in \mathbb{C}^n$ *has a non-singular covariance matrix* $\mathbf{C_x}$ *and a pseudo-covariance matrix* $\mathbf{P_x}$. *Let* $\mathbf{B_x}$ *be a generalized Cholesky factor of* $\mathbf{C_x}$ *and let* λ_i *denote the singular values of* $\mathbf{B_x}^{-1}\mathbf{P_x}\mathbf{B_x}^{-\mathbf{T}}$, *which must be smaller than* 1. *Then the entropy of* $\mathbf{x}$ *satisfies*

$$h(\mathbf{x}) \le \log\det(\pi e \mathbf{C_x}) + \frac{1}{2}\sum_{\lambda_i}\log\left(1-\lambda_i^2\right),$$

with equality if and only if $\mathbf{x}$ *is Gaussian.*

Acknowledgments

The author would like to thank Jossy Sayir for helpful hints and discussions.

References

[1] F.D. Neeser, J.L. Massey, “Proper Complex Random Processes with Applications to Information Theory”, *IEEE Transactions on Information Theory*, Vol. 39, No. 4, pp. 1293-1302, Jul 1993.

[2] I.E. Telatar, “Capacity of Multi-antenna Gaussian Channels”, *European Transactions on Telecommunications*, Vol. 10, No. 6, pp. 585-595, Nov/Dec 1999.

[3] G.H. Golub, C.F. Van Loan, “Matrix Computations”, *North Oxford Academic Publishers Ltd, a subsidiary of Kogan Page Ltd*, p. 16, 1986.

[4] A. Lampe, R. Schober, W.H. Gerstacker, “Iterative Multiuser Detection for Complex Modulation Schemes”, *Proc. of IEEE International Symposium on Information Theory (ISIT)*, (Washington D.C.), p. 33, 2001.

[5] G. Tauböck, “Noise Analysis of DMT”, *Proc. of Globecom 2003*, San Francisco, USA, Dec 1-5, 2003.

[6] G. Tauböck, “Rotationally Variant Complex Channels”, *Proc. of the 23rd Symposium on Information Theory in the Benelux*, Louvain-la-Neuve, Belgium, May 29-31, 2002.