A Note on the Wiener Filter for Vector Random Processes

Jesús Gutiérrez-Gutiérrez[1], Adam Podhorski[1], Iñaki Iglesias[2], and Javier Del Ser[3]

[1] CEIT and Tecnun (University of Navarra),
Manuel de Lardizábal 15, 20018 San Sebastián, Spain
{jgutierrez,apodhorski}@ceit.es
[2] Department of Electrical and Computer Engineering,
University of Delaware, 302 Evans Hall, Newark, DE 19716 USA
iglesias@udel.edu
[3] Tecnalia Research & Innovation
P. Tecnológico, Ed. 202, 48170 Zamudio, Spain
javier.delser@tecnalia.com

Abstract. In the present paper we compute the geometric minimum mean square error for the vector linear estimation problem. We do this by proving that the vector linear estimator that minimizes the mean square error (MSE) also minimizes the geometric MSE.

Keywords: minimum mean square error (MMSE), vector linear estimation, vector linear interpolation, vector linear prediction.

1 Introduction

The linear estimation of a random vector from another random vector, which, at first sight, may seem to be a very particular case of vector linear estimation, represents, in fact, a very general framework, which includes, as it will be shown in the appendix, vector linear prediction, vector linear interpolation, multiple-input multiple-output (MIMO) linear equalization and MIMO decision-feedback equalization (DFE).

As we are dealing with random vectors, the definition of mean square error (MSE) is not unique. Here we will consider two possible definitions, which will be called MSE and geometric MSE (GMSE), respectively. We will name the former simply MSE because it is the definition commonly used [1,2]. The latter was introduced in [3] for MIMO DFE. The interest in using one or the other error measure depends on the problem considered.

In this paper we will prove that the vector linear estimator that minimizes the MSE also minimizes the GMSE. This fact was proved in [4] for MIMO DFE, and it was assumed true, but not proved, in [5] for MIMO linear prediction.

J. Del Ser et al. (Eds.): MOBILIGHT 2011, LNICST 81, pp. 30–36, 2012.

2 General Framework

2.1 Vector Linear Estimation Problem Considered

Let $\mathbf{x}$ and $\mathbf{y}$ be two complex random vectors of dimensions N and M, respectively. We assume that the correlation matrix of $\mathbf{y}$ is invertible, and we estimate $\mathbf{x}$ from $\mathbf{y}$ in the following manner:

$$\hat{\mathbf{x}} = W\mathbf{y}, \tag{1}$$

where $W \in \mathbb{C}^{N\times M}$ and $\mathbb{C}^{N\times M}$ is the set of all $N \times M$ complex matrices. We denote by $\mathbf{e}$ the error vector $\mathbf{x} - \hat{\mathbf{x}}$, and by R its correlation matrix, that is,

$$\begin{aligned} R &= \mathrm{E}\left[\mathbf{e}\mathbf{e}^*\right] \\ &= \mathrm{E}\left[(\mathbf{x} - W\mathbf{y})\left(\mathbf{x}^* - \mathbf{y}^* W^*\right)\right] \\ &= \mathrm{E}\left[\mathbf{x}\mathbf{x}^*\right] - \mathrm{E}\left[\mathbf{x}\mathbf{y}^*\right] W^* - W\mathrm{E}\left[\mathbf{y}\mathbf{x}^*\right] + W\mathrm{E}\left[\mathbf{y}\mathbf{y}^*\right] W^*, \end{aligned} \tag{2}$$

where $*$ indicates complex conjugate transpose. MSE and GMSE are defined as the trace and the determinant of R, respectively.

2.2 Principle of Orthogonality

We now compute the matrices W and R associated with the linear estimator $\hat{\mathbf{x}}$ that satisfies the following condition that is called principle of orthogonality:

$$0_{N\times M} = \mathrm{E}\left[\mathbf{e}\mathbf{y}^*\right], \tag{3}$$

where $0_{N\times M}$ is the $N \times M$ zero matrix. We denote these two matrices by W_0 and R_0. From (3) we have

$$0_{N\times M} = \mathrm{E}\left[\mathbf{x}\mathbf{y}^*\right] - W_0\mathrm{E}\left[\mathbf{y}\mathbf{y}^*\right], \tag{4}$$

and consequently,

$$W_0 = \mathrm{E}\left[\mathbf{x}\mathbf{y}^*\right]\left(\mathrm{E}\left[\mathbf{y}\mathbf{y}^*\right]\right)^{-1}. \tag{5}$$

Combining (2) and (4) we obtain

$$R_0 = \mathrm{E}\left[\mathbf{x}\mathbf{x}^*\right] - W_0\mathrm{E}\left[\mathbf{y}\mathbf{x}^*\right], \tag{6}$$

and applying (5) yields

$$R_0 = \mathrm{E}\left[\mathbf{x}\mathbf{x}^*\right] - \mathrm{E}\left[\mathbf{x}\mathbf{y}^*\right]\left(\mathrm{E}\left[\mathbf{y}\mathbf{y}^*\right]\right)^{-1}\mathrm{E}\left[\mathbf{y}\mathbf{x}^*\right]. \tag{7}$$

Observe that (4) and (6) can be written as

$$\begin{cases} R_0 = \begin{pmatrix} I_N & -W_0 \end{pmatrix} \begin{pmatrix} \mathrm{E}\left[\mathbf{x}\mathbf{x}^*\right] \\ \mathrm{E}\left[\mathbf{y}\mathbf{x}^*\right] \end{pmatrix}, \\ 0_{N\times M} = \begin{pmatrix} I_N & -W_0 \end{pmatrix} \begin{pmatrix} \mathrm{E}\left[\mathbf{x}\mathbf{y}^*\right] \\ \mathrm{E}\left[\mathbf{y}\mathbf{y}^*\right] \end{pmatrix}, \end{cases} \tag{8}$$

where I_N is the $N \times N$ identity matrix. The system of two equations (8) is equivalent to the following single equation:

$$(R_0 \quad 0_{N\times M}) = (I_N \quad -W_0)\,\mathrm{E}[\mathbf{z}\mathbf{z}^*], \tag{9}$$

with $\mathbf{z} = \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}$. As in the theory of linear prediction [1], we can call equations (4) and (9) normal equation and augmented normal equation, respectively.

2.3 Relation between R and R_0

Since R and R_0 are Hermitian[1], $S = R - R_0$ is Hermitian. In this subsection we prove that S is also positive semidefinite. Using (2) and (7) we have

$$S = \mathrm{E}\left[\mathbf{x}\mathbf{y}^*\right]\left(\mathrm{E}\left[\mathbf{y}\mathbf{y}^*\right]\right)^{-1}\mathrm{E}\left[\mathbf{y}\mathbf{x}^*\right] - \mathrm{E}\left[\mathbf{x}\mathbf{y}^*\right]W^* - W\mathrm{E}\left[\mathbf{y}\mathbf{x}^*\right] + W\mathrm{E}\left[\mathbf{y}\mathbf{y}^*\right]W^*.$$

Thus

$$S = B\mathrm{E}\left[\mathbf{y}\mathbf{y}^*\right]B^*, \tag{10}$$

with

$$B = \mathrm{E}\left[\mathbf{x}\mathbf{y}^*\right]\left(\mathrm{E}\left[\mathbf{y}\mathbf{y}^*\right]\right)^{-1} - W.$$

From the expression given in (10) for the matrix S, we can now show that S is positive semidefinite:

$$v^*Sv = \mathrm{E}\left[v^*B\mathbf{y}\mathbf{y}^*B^*v\right] = \mathrm{E}\left[|v^*B\mathbf{y}|^2\right] \geq 0 \quad \forall v \in \mathbb{C}^{N\times 1}.$$

2.4 MMSE

If A is an $N \times N$ diagonalizable matrix, then $\mathrm{tr}(A) = \sum_{k=1}^{N} \lambda_k(A)$, where tr denotes trace and $\lambda_1(A), \ldots, \lambda_N(A)$ are the eigenvalues of A counted with their multiplicities. Since S is a Hermitian positive semidefinite matrix, S is diagonalizable and all its eigenvalues are non-negative, and consequently, $\mathrm{tr}(S) \geq 0$. Therefore using that $\mathrm{tr}(S) = \mathrm{tr}(R) - \mathrm{tr}(R_0)$ we conclude that

$$\mathrm{tr}(R) \geq \mathrm{tr}(R_0). \tag{11}$$

Thus we have proved that the minimum MSE (MMSE) is

$$\mathrm{MMSE} = \mathrm{tr}\left(R_0\right).$$

2.5 Geometric MMSE

In this subsection we assume that the Hermitian positive semidefinite matrix R_0 is invertible (or equivalently, positive definite) and we show that (11) is also true when trace is replaced by determinant.

[1] Any correlation matrix is Hermitian and positive semidefinite.

Let $R_0 = V \operatorname{diag}\left(\lambda_1\left(R_0\right), \ldots, \lambda_N\left(R_0\right)\right) V^*$ be an eigenvalue decomposition of the Hermitian matrix R_0. Since R_0 is positive definite, all its eigenvalues are positive, and consequently, we can define the following two Hermitian matrices:

$$R_0^{\frac{1}{2}} := V \operatorname{diag}\left(\left(\lambda_1\left(R_0\right)\right)^{\frac{1}{2}}, \ldots, \left(\lambda_N\left(R_0\right)\right)^{\frac{1}{2}}\right) V^*,$$

and

$$R_0^{-\frac{1}{2}} := V \operatorname{diag}\left(\left(\lambda_1\left(R_0\right)\right)^{-\frac{1}{2}}, \ldots, \left(\lambda_N\left(R_0\right)\right)^{-\frac{1}{2}}\right) V^* = \left(R_0^{\frac{1}{2}}\right)^{-1}.$$

Hence the determinant of the matrix R can be expressed as

$$\begin{aligned}
\det(R) &= \det\left(R_0 + S\right) \\
&= \det\left(R_0^{\frac{1}{2}}\left(R_0 + S\right) R_0^{-\frac{1}{2}}\right) \\
&= \det\left(R_0 + R_0^{\frac{1}{2}} S R_0^{-\frac{1}{2}}\right) \\
&= \det\left(R_0\left(I_N + R_0^{-\frac{1}{2}} S R_0^{-\frac{1}{2}}\right)\right) \\
&= \det\left(R_0\right) \det\left(I_N + R_0^{-\frac{1}{2}} S \left(R_0^{-\frac{1}{2}}\right)^*\right).
\end{aligned}$$

Therefore if $R_0^{-\frac{1}{2}} S \left(R_0^{-\frac{1}{2}}\right)^* = UDU^*$ is an eigenvalue decomposition of the Hermitian matrix $R_0^{-\frac{1}{2}} S \left(R_0^{-\frac{1}{2}}\right)^*$, then

$$\begin{aligned}
\frac{\det(R)}{\det(R_0)} &= \det\left(I_N + UDU^*\right) \\
&= \det\left(UU^* + UDU^*\right) \\
&= \det\left(U(I_N + D)U^*\right) \\
&= \det\left(I_N + D\right) \\
&= \prod_{k=1}^{N}\left(1 + \lambda_k\left(R_0^{-\frac{1}{2}} S \left(R_0^{-\frac{1}{2}}\right)^*\right)\right).
\end{aligned} \tag{12}$$

Since S is positive semidefinite, $R_0^{-\frac{1}{2}} S \left(R_0^{-\frac{1}{2}}\right)^*$ is also positive semidefinite:

$$v^* R_0^{-\frac{1}{2}} S \left(R_0^{-\frac{1}{2}}\right)^* v = \left(\left(R_0^{-\frac{1}{2}}\right)^* v\right)^* S \left(R_0^{-\frac{1}{2}}\right)^* v \geq 0 \qquad \forall v \in \mathbb{C}^{N \times 1}.$$

Consequently, all the eigenvalues of $R_0^{-\frac{1}{2}} S \left(R_0^{-\frac{1}{2}}\right)^*$ are non-negative, and from (12) we obtain

$$\frac{\det(R)}{\det\left(R_0\right)} \geq 1.$$

Finally, since $\det\left(R_0\right) = \prod_{k=1}^{N} \lambda_k(R_0) > 0$ we conclude that

$$\det(R) \geq \det\left(R_0\right).$$

Thus we have proved that the geometric MMSE (GMMSE) is

$$\text{GMMSE} = \det(R_0).$$

Acknowledgments. The work of Jesús Gutiérrez-Gutiérrez and Javier Del Ser was supported in part by the Basque Government through the MIMONET project (PC2009-27B), and by the Spanish Ministry of Science and Innovation through the projects COSIMA (TEC2010-19545-C04-02), COMONSENS (CSD2008-00010), and a Torres-Quevedo grant (PTQ-09-01-00740).

Appendix A: Some Specific Cases

A.1 Vector Linear Interpolation and Prediction

Consider $m, d \in \mathbb{N}$, with $1 \leq d \leq m$. Let $\{\mathbf{x}_n; n \in \mathbb{Z}\}$ be a complex N-dimensional random process. We estimate $\mathbf{x}_d$ from $\mathbf{x}_1, \dots, \mathbf{x}_{d-1}, \mathbf{x}_{d+1}, \dots, \mathbf{x}_m$ in the following manner:

$$\hat{\mathbf{x}}_d = \sum_{\substack{1 \leq k \leq m \\ k \neq d}} W_k \mathbf{x}_k, \tag{13}$$

where $W_k \in \mathbb{C}^{N \times N}$.

The estimator given in (13) is called linear interpolator when $1 < d < m$, forward linear predictor when $d = m$, and backward linear predictor when $d = 1$.

The estimation scheme (13) can be obtained as a particular case of (1) by taking:

$$\begin{cases} \hat{\mathbf{x}} = \hat{\mathbf{x}}_d, \\ \mathbf{y} = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{d-1} \\ \mathbf{x}_{d+1} \\ \vdots \\ \mathbf{x}_m \end{pmatrix}, \\ W = \begin{pmatrix} W_1 & \dots & W_{d-1} & W_{d+1} & \dots & W_m \end{pmatrix}. \end{cases}$$

A.2 Vector Wiener Filtering

Let $\{\mathbf{x}_n; n \in \mathbb{Z}\}$ and $\{\mathbf{y}_n; n \in \mathbb{Z}\}$ be two complex vector random processes of dimensions N and M, respectively. Given $m \in \mathbb{N}$ we estimate $\mathbf{x}_n$ from $\mathbf{y}_n, \mathbf{y}_{n-1}, \dots, \mathbf{y}_{n-m+1}$ in the following manner:

$$\hat{\mathbf{x}}_n = \sum_{k=1}^{m} W_k \mathbf{y}_{n-k+1}, \tag{14}$$

where $W_k \in \mathbb{C}^{N \times M}$.

The filter given in (14) is called the vector Wiener filter when it minimizes the MSE. We have proved in this paper that the vector Wiener filter can also be defined as the filter that minimizes the GMSE.

The estimation scheme (14) can be obtained as a particular case of (1) by taking:

$$\begin{cases} \hat{\mathbf{x}} &= \hat{\mathbf{x}}_n, \\ \mathbf{y} &= \begin{pmatrix} \mathbf{y}_n \\ \vdots \\ \mathbf{y}_{n-m+1} \end{pmatrix}, \\ W &= \begin{pmatrix} W_1 & \ldots & W_m \end{pmatrix}. \end{cases}$$

Observe that (1) can also be obtained as a particular case of (14) by taking:

$$\begin{cases} \hat{\mathbf{x}}_n &= \hat{\mathbf{x}}, \\ m &= 1, \\ W_1 &= W, \\ \mathbf{y}_n &= \mathbf{y}. \end{cases}$$

That is, the estimation schemes (1) and (14) are equivalent.

A.3 MIMO DFE

We consider a MIMO communication system given by

$$\mathbf{y}_n = \sum_{k=0}^{p} H_k \mathbf{x}_{n-k} + \mathbf{s}_n,$$

where p is a non-negative integer, $H_k \in \mathbb{C}^{M\times N}$, the input $\{\mathbf{x}_n; n \in \mathbb{Z}\}$ is a discrete complex vector random process of dimension N, and the noise $\{\mathbf{s}_n; n \in \mathbb{Z}\}$ is a complex vector random process of dimension M.

We estimate $\mathbf{x}_{n-d}$ from m output vectors $\mathbf{y}_n, \mathbf{y}_{n-1}, \ldots, \mathbf{y}_{n-m+1}$, and l input vectors $\mathbf{x}_{n-1-d}, \mathbf{x}_{n-2-d}, \ldots, \mathbf{x}_{n-l-d}$ in the following manner:

$$\hat{\mathbf{x}}_{n-d} = \sum_{k=1}^{m} W_k \mathbf{y}_{n-k+1} + \sum_{k=1}^{l} V_k \mathbf{x}_{n-k-d}, \tag{15}$$

where $W_k \in \mathbb{C}^{N\times M}$, $V_k \in \mathbb{C}^{N\times N}$, d is the decision delay, and $m, l \in \mathbb{N}$ should satisfy $d + l + 1 = m + p$.

The estimator given in (15) is called MIMO decision-feedback equalizer except when $V_k = 0$ with $1 \leq k \leq l$. In that other case it is called MIMO linear equalizer.

The estimation scheme (15) can be obtained as a particular case of (1) by taking:

$$\begin{cases} \hat{\mathbf{x}} &= \hat{\mathbf{x}}_{n-d}, \\ \mathbf{y} &= \begin{pmatrix} \mathbf{y}_n \\ \vdots \\ \mathbf{y}_{n-m+1} \\ \mathbf{x}_{n-1-d} \\ \vdots \\ \mathbf{x}_{n-l-d} \end{pmatrix}, \\ W = \begin{pmatrix} W_1 & \ldots & W_m & V_1 & \ldots & V_l \end{pmatrix}. \end{cases}$$

References

1. Vaidyanathan, P.P.: The Theory of Linear Prediction. Morgan & Claypool (2008)
2. Benesty, J., Huang, Y., Chen, J.: Wiener and adaptive filters. In: Benesty, J., Sondhi, M.M., Huang, Y. (eds.) Springer Handbook of Speech Processing. Springer, Heidelberg (2008)
3. Yang, J., Roy, S.: Joint transmitter-receiver optimization for multi-input multi-output systems with decision feedback. IEEE Trans. Inf. Theory 40(5), 1334–1347 (1994)
4. Al-Dhahir, N., Sayed, A.H.: The finite-length multi-input multi-output MMSE-DFE. IEEE Trans. Signal Process. 48(10), 2921–2936 (2000)
5. Gutiérrez-Gutiérrez, J., Crespo, P.M.: Asymptotically equivalent sequences of matrices and Hermitian block Toeplitz matrices with continuous symbols: applications to MIMO systems. IEEE Trans. Inf. Theory 54(12), 5671–5680 (2008)