Widely linear estimation with complex data — **Source link** ↗

Bernard Picinbono, Pascal Chevalier

Institutions: STMicroelectronics

Related papers:

- Second-order analysis of improper complex random vectors and processes

- Proper complex random processes with applications to information theory

- On circularity

- Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models

- Complex Valued Nonlinear Adaptive Filters

Widely Linear Estimation with Complex Data

Bernard Picinbono, Pascal Chevalier

Widely Linear Estimation

with Complex Data

Bernard Picinbono, *Fellow, IEEE* and Pascal Chevalier

Abstract

Mean square estimntioa of complex and normal data is not linear as in the real case but widely linear. The purpose of this correspondence is to calculate the optimum widely linear mean square estimate and to present its main properties. The advantage with respect to linear procedure is espedplly analyzed.

I. INTRODUCTION

Mean square estimation (MSE) is one of the most fundamental techniques of statistical signal processing. The basic problem can be stated as follows: Let y be a scalar random variable to be estimated (estimandum) in terms of an observation that is a random vector $\mathbf{x}$. The estimate $\hat{y}$ that minimizes the MS error is then the regression or the conditional expectation value $E[y|\mathbf{x}]$. This result is usually given

when $\mathbf{x}$ and y are real. However, it remains valid when these quantities are complex valued. If $\mathbf{x}$ and y are jointly normal with zero mean value and are real, then the regression is *linear*. However, this is no longer true for normal complex data, where the regression is linear both in $\mathbf{x}$ and $\mathbf{x}^*$ and is called widely linear (WL). It is interesting to study the properties of WL systems for MSE without introducing the normal assumption. In LMSE, the problem is to find an estimate written as

$$\hat{y} = \mathbf{h}^H \mathbf{x}, \tag{1}$$

where H means the complex conjugation and transposition (or Hermitian transposition). Because $\hat{y}$ is a scalar product, it results from the definition of such a product that $\hat{y}$ is a linear function of the vector $\mathbf{x}$, as defined in classical linear algebra. Consider now the scalar y' defined by

$$y' = \mathbf{h}^H \mathbf{x} + \mathbf{g}^H \mathbf{x}^*, \tag{2}$$

where $\mathbf{x}^*$ is the complex conjugate of $\mathbf{x}$, and $\mathbf{g}$ is another complex vector. This is the general form of the regression for complex normal random variables. It is clear that y' is not a linear function of x. However, the moment of order k of y' is completely defined from the moments of order k of $\mathbf{x}$ and $\mathbf{x}^*$, which characterizes a form of linearity. This is why (2) will be called a wide sense linear filter or system.

The general purpose of this correspondence is to show that, taking into consideration WL systems defined by (2) instead of strictly linear ones defined by (1) can yield significant improvements in estimation problems using complex data. This result can appear rather natural and was indicated under restrictive conditions in [l] and [2], and the principles of a more general presentation appear in [3, p. 4131 without the detailed analysis discussed hereafter. In reality, there is at least one reason why (2) is not widely used. This is due to the fact that in almost all calculations using complex Gaussian distribution, the assumption of circularity is explicitly (or implicitly) introduced (see [3, p. 1181). This assumption is valid in many practical situations, and in this case, the last term of (2) disappears in such a way that complex signals and systems can be treated as if they were real. However, this assumption, which is strongly connected with stationarity [4], has no reason for being general. This justifies the analysis of this correspondence.

II. WL MEAN SQUARE ESTIMATION

The problem of WL mean square estimation (WLMSE) is to find the vectors $\mathbf{u}$ and $\mathbf{v}$ in such a way that

$$\hat{y} = \mathbf{u}^H \mathbf{x} + \mathbf{v}^H \mathbf{x}^* \tag{3}$$

gives the minimum mse $E[|y - \hat{y}|^2]$. For this purpose, the first point to note is that the set of scalar complex random variables $z(\omega)$ in the form $z(\omega) = \mathbf{a}^H \mathbf{x}(\omega) + \mathbf{b}^H \mathbf{x}^*(\omega)$, where $\mathbf{a}$ and $\mathbf{b}$ belong to $\mathbb{C}^N$

constitutes a linear space. It becomes a Hilbert subspace with the scalar product $(zz') = E(z^*z')$. As a result $\hat{y}$ is the projection of y onto this subspace and is characterized by the orthogonality principle

$$(y - \hat{y}) \perp\!\!\!\perp \mathbf{x} \; ; \; (y - \hat{y}) \perp\!\!\!\perp \mathbf{x}^*. \tag{4}$$

The symbol $\perp\!\!\!\perp$ means that all the components of $\mathbf{x}$ or $\mathbf{x}^*$ are orthogonal to $(y - \hat{y})$ with the previous scalar product. As a consequence, these equations can be written in terms of expectations, which yields

$$E(\hat{y}^*\mathbf{x})E(y^*\mathbf{x}) \; ; \; E(\hat{y}^*\mathbf{x}^*)E(y^*\mathbf{x}^*). \tag{5}$$

Replacing $\hat{y}$ with (3) gives

$$\mathbf{\Gamma u} + \mathbf{Cv} = \mathbf{r}, \tag{6}$$

$$\mathbf{C}^*\mathbf{u} + \mathbf{\Gamma}^*\mathbf{v} = \mathbf{s}^*, \tag{7}$$

where

$$\mathbf{\Gamma} = E(\mathbf{xx}^H) \; ; \; \mathbf{C} = E(\mathbf{xx}^T) \; ; \; \mathbf{r} = E(y^*\mathbf{x}) \; ; \; \mathbf{s} = E(y\mathbf{x}). \tag{8}$$

From (6) and (7), we find the solutions expressed as

$$\mathbf{u} = [\mathbf{\Gamma} - \mathbf{C\Gamma}^{-1*}\mathbf{C}^*]^{-1}[\mathbf{r} - \mathbf{C\Gamma}^{-1*}\mathbf{s}^*], \tag{9}$$

$$\mathbf{v} = [\mathbf{\Gamma}^* - \mathbf{C}^*\mathbf{\Gamma}^{-1}\mathbf{C}]^{-1}[\mathbf{s}^* - \mathbf{C}^*\mathbf{\Gamma}^{-1}\mathbf{r}]. \tag{10}$$

The corresponding mse is also deduced from the projection theo- and by using (3), (6), and (7) we obtain

$$\epsilon^2 = E[|y|^2] - (\mathbf{u}^H\mathbf{r} + \mathbf{v}^H\mathbf{s}^*). \tag{11}$$

This error is smaller than ϵ_L^2, which is the error that is obtained with a strictly LMSE-like (1) and equal to $E[|y|^2] - \mathbf{r}^H\mathbf{\Gamma}^{-1}\mathbf{r}$. The advantage of the WLMSE procedure over the LMSE is characterized by the quantity $\delta\epsilon_L^2 = \epsilon^2 - \epsilon^2$ which can be expressed as

$$\delta\epsilon^2 = [\mathbf{s}^* - \mathbf{C}^*\mathbf{\Gamma}^{-1}\mathbf{r}]^H[\mathbf{\Gamma}^* - \mathbf{C}^*\mathbf{\Gamma}^{-1}\mathbf{C}]^{-1}[\mathbf{s}^* - \mathbf{C}^*\mathbf{\Gamma}^{-1}\mathbf{r}]. \tag{12}$$

It is always nonnegative because the matrix $\mathbf{\Gamma}^* - \mathbf{C}^*\mathbf{\Gamma}^{-1}\mathbf{C}$ is positive definite, and consequently, $\delta\epsilon^2 = 0$ only when $\mathbf{s}^* - \mathbf{C}^*\mathbf{\Gamma}^{-1}\mathbf{r} = \mathbf{0}$.

At this step, it is worth pointing out that all the previous calculations could be realized by using only real quantities. However, by doing so, the compact expression of the complex quantity $\hat{y}$ is less obvious, and (12) takes a much more complex form. Furthermore, the comparison with the strictly linear procedure characterized by $\mathbf{v} = \mathbf{0}$ is less obvious with real quantities than with (3). Finally, in the strictly linear case, nobody has the idea to transform the complex Wiener-Hopf equation $\mathbf{\Gamma u} = \mathbf{c}$ in a set of two real equations.

III. EXAMPLES

A. Jointly Circular Case

This situation is characterized by

$$\mathbf{C} = \mathbf{r} \; ; \;\; \mathbf{s} = \mathbf{0}. \tag{13}$$

This assumption is well known in the normal case (see [3, p. 1181). In reality, it is sometimes used in the definition of complex normal random vectors, [5], [6, p. 1281, where the term "strongly normal" is used and is used in [7] as well. In particular, one can show that under some conditions, the Fourier components of stationary signals are complex circular random variables. The analytic signal of a real stationary signal is also second-order circular. The term of circularity comes from the fact that if (13) holds, the random vectors $\mathbf{x}$ and $\mathbf{x} \exp(j\alpha)$ have the same second-order properties for any α. Note that (13) characterizes only second-order circularity, and the concept must be extended when using higher order statistics. Note also that (13) means a joint circularity and is then an assumption on $\mathbf{x}$ and y.

It immediately results from (10) that (13) implies $\mathbf{v} = \mathbf{0}$. Similarly, (9) gives $\mathbf{u} = \mathbf{\Gamma}^{-1}\mathbf{r}$. Thus, the assumption of joint circularity implies that the WLMSE (3) takes the form (1) and is strictly linear. It is also clear that (12) gives $\delta\epsilon^2 = 0$, and the conclusion is that in the case of a joint circularity, the strictly linear system (1) is sufficient to reach the best performance. This is one of the arguments justifying the interest of circularity. However, even if circularity appears in many practical situations [7], there are cases where it cannot be introduced.

B. Circular Observation

Suppose now that the second assumption (13) is deleted. This means that circularity is only valid for the observation and is chmcterized by $\mathbf{C} = \mathbf{0}$, whereas no specific assumption is introduced forthe estimandum y. In this case, (9) and (10) are greatly simplified and become

$$\mathbf{u} = \mathbf{\Gamma}^{-1}\mathbf{r} \; ; \;\; \mathbf{v}^* = \mathbf{\Gamma}^{-1}\mathbf{s}. \tag{14}$$

his means that the term $\mathbf{u}^H\mathbf{x}$ (3) is the same as the one obtained when using strictly linear estimation. This fact can be explained by noting that the circularity assumption implies that the vectors $\mathbf{x}$ and are uncorrelated. Thus, the Hilbert subspaces generated by x and $\mathbf{x}^*$ are orthogonal, and taking into account $\mathbf{x}^*$ does not change the term coming from $\mathbf{x}$ only. This also explains the simplification of (12) that becomes

$$\delta\epsilon^2 = \mathbf{s}^H\mathbf{\Gamma}^{-1}\mathbf{s}. \tag{15}$$

Thus, a nonzero vector $\mathbf{s}$ necessarily implies an increment of the performance of estimation when using the structure (2) instead of (1).

C. Case of a Real Estimandum y

The estimation of a real quantity from complex data appears in many situations as, for example, when the observation comes from Fourier components of a real signal, some real parameters of which have to be estimated. Suppose then that y is real, $\mathbf{x}$ still being complex. This obviously implies that $\mathbf{r} = \mathbf{s}$ in (8). It results from (9) and (10) that $\mathbf{u} = \mathbf{v}^*$, and consequently,

$$\hat{y} = 2\mathrm{Re}(\mathbf{u}^H \mathbf{x}) \tag{16}$$

Similarly, the estimation error takes the form

$$\epsilon^2 = E(y^2) - 2\mathrm{Re}(\mathbf{u}^H \mathbf{r}). \tag{17}$$

The main property of the estimate (16) is that it is real, although there is no reason for the strictly linear estimate to be real, which is not convenient when estimating a real quantity.

The advantage of the structure (3) with respect to (2) is even more clear when the observation x is circular. In fact, as seen previously in Section III-B, the vector $\mathbf{u}$ is the same as the one that must be used to realize the LMSE of y with (1). Thus, by using this vector, the two estimators (1) and (2) become

$$\hat{y}_L = \mathbf{u}^H \mathbf{x} \ ; \ \ \hat{y}_{WL}^2 = 2\mathrm{Re}(\mathbf{u}^H \mathbf{x}), \tag{18}$$

and the corresponding errors are

$$\epsilon^2 = E(y^2) - (\mathbf{u}^H \mathbf{r}) \ ; \ \ \epsilon_{WL}^2 = E(y^2) - 2(\mathbf{u}^H \mathbf{r}). \tag{19}$$

Note that the quantity $\mathbf{u}^H \mathbf{r}$ is positive because it is equal to $\mathbf{u}^H \mathbf{\Gamma} \mathbf{u}$ and $\mathbf{\Gamma}$ is a positive definite matrix. In conclusion, the wide sense linear estimator (18) provides a real estimate and a decrease of the error that is twice as great as the strictly linear estimate, which in general is complex. It is also clear from (19) that $\delta\epsilon^2 = \mathbf{u}^H \mathbf{r}$.

D. Singular Estimation

The estimation is singular when the mse is zero. If the wide sense linear mean square error (11) is zero, the estimandum y belongs to the Hilbert subspace introduced after (3) and can then be written as

$$y = \mathbf{a}^H \mathbf{x} + \mathbf{b}^H \mathbf{x}^* \tag{20}$$

It is now interesting to study the behavior of the strictly LMSE when (20) holds or in the case of singular wide sense LMSE. Note first that if $\mathbf{b} = \mathbf{0}$, (20) becomes strictly linear. In this case, singular estimation appears equally well, with the two forms of linear procedures.

Let us now investigate the complete opposite situation. It corresponds to the case where the strictly linear procedure provides a zero estimation. This means that the meas that the mse ϵ_L^2 is equal to $E[|y|^2]$. This situation appears when $\mathbf{r}$ defined by (8) is zero, which means that the estimandum y and the observation $\mathbf{x}$ are uncorrelated. By replacing y given by (20) in (8), we obtain

$$\mathbf{r} = \mathbf{\Gamma a} + \mathbf{Cb}. \tag{21}$$

As $\mathbf{\Gamma}$ is positive definite, the condition $\mathbf{r} = \mathbf{0}$ is equivalent to

$$\mathbf{a} = -\mathbf{\Gamma}^{-1}\mathbf{Cb}. \tag{22}$$

Then, if $\mathbf{C} \neq \mathbf{0}$ or if $\mathbf{x}$ is not circular, it is possible to associate to any nonzero vector $\mathbf{b}$ another nonzero vector $\mathbf{a}$ given by (22) and such that y given by (20) is uncorrelated with $\mathbf{x}$. This implies a zero LMSE. On the other hand, because of (20), $\hat{y}$ given by (3) is equal to y, and the mse zero, which means singular estimation. We then have zero estimation with the strictly linear procedure and perfect estimation with the wide sense liiear procedure.

IV. CONCLUSION

We have presented a compact approach of the general problem of linear estimation with complex data. This problem is usually presented as a straightforward extension of the results obtained in the real case. By doing so, it is in general impossible to reach the optimum performance that can be deduced from the second-order statistics. This requires the introduction of widely linear systems. The structure of WLMSE has been determined. From this result, we have shown that widely linear systems can yield significant improvements in estimation performance with respect to strictly linear systems generally used, except when the circularity assumption is introduced. It is even possible to reach singular WL estimation while strictly linear systems give a zero estimation.

REFERENCES

[1] W. M. Brown and R. B. Crawe, "Conjugate linear filtering," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 462-465, 1969.

[2] W. A. Gardner, "Cyclic Wiener filtering: Theory and methods," *IEEE Trans. Trans. Commun.*, vol. 41, pp. 151-163, 1993.

[3] B. Picinbono, *Random Signals and Systems*, Englewood Cliffs: N.J. : Prentice Hall, 1993.

[4] B. Picinbono, "On circularity," *IEEE Trans. Signal Processing*, vol. 42, pp. 3473-3482, 1994.

[5] K. Miller, *Multidimensional Gaussian Distributions*, New York: Wiley, 1964.

[6] M. Loève, *Probability Theory, vol. 2*, New York: Springer Verlag, 1978.

[7] S. Kay, *Modem Spectral Analysis, Theory and Applications*, Englewood Cliffs; NJ Prentice-Hall, 1988.