BEST WIDELY LINEAR UNBIASED ESTIMATOR FOR REAL VALUED PARAMETER VECTORS

O. Lang and M. Huemer

Johannes Kepler University Linz
Institute of Signal Processing
Altenbergerstraße 69, 4040 Linz, Austria

ABSTRACT
For classical estimation with an underlying linear model the best linear unbiased estimator (BLUE) is usually utilized for estimating the deterministic but unknown parameter vector. In the case of real valued parameter vectors but complex valued measurement matrices and noise vectors, the BLUE results in complex valued estimates, introducing a systematic error. In recent years widely linear estimators have been investigated for complex estimation. In this work a novel widely linear classical estimator is derived which incorporates the knowledge that the parameter vector is real valued. The proposed estimator is unbiased in the classical sense and it outperforms the BLUE and the best widely linear unbiased estimator (BWLUE) in terms of the variances of the vector estimator’s elements.

Index Terms— Classical estimation, BLUE, BWLUE, augmented form, widely linear.

1. INTRODUCTION

In this work we consider the task of estimating a real valued parameter vector based on complex valued measurements. The measurements are assumed to be connected with the parameters via the linear model

\[ y = Hx + n, \]

where \( x \in \mathbb{R}^{N_x \times 1} \) is a real valued parameter vector, \( y \in \mathbb{C}^{N_y \times 1} \) is a complex valued measurement vector, \( H \in \mathbb{C}^{N_y \times N_x} \) is the complex valued measurement matrix, and \( n \in \mathbb{C}^{N_y \times 1} \) is a complex valued zero mean random noise vector.

In a Bayesian interpretation, the real valued parameter vector is improper and the application of widely linear estimators is obvious. For the definition of propriety we refer to [1] and Sec. 2. A widely linear Bayesian estimator suitable for the described problem is the widely linear minimum mean square error (WLMMSE) estimator. The WLMMSE estimator requires first and second order statistics about \( x \). For zero mean \( x \), the WLMMSE estimator is of the form

\[ \hat{x} = Ey + Fy^*, \]

where \((\cdot)^*\) denotes the complex conjugate. Isolating the \( i \)th row of (2) yields

\[ \hat{x}_i = e_i^H y + f_i^H y^*, \]

where \( e_i^H \) and \( f_i^H \) are the \( i \)th rows of \( E \) and \( F \), respectively, and \((\cdot)^H\) denotes the conjugate transposition. For real valued parameter vectors, the WLMMSE estimator fulfills the property

\[ e_i^H = f_i^T, \]

with \((\cdot)^T\) denoting transposition. Consequently, in this case the WLMMSE estimator produces real valued estimates.

Other estimators such as the Bayesian LMMSE estimator [2], the best linear unbiased estimator (BLUE) [2], or the best widely linear unbiased estimator (BWLUE) [3] do not result in real valued estimates \( \hat{x}_i \) for real parameters \( x_i \) in general; an exception is the case \( H, x, n \in \mathbb{R} \). In fact, the classical estimators BLUE and BWLUE do not allow the utilization of statistics about \( x \). These classical estimators are employed, e.g., when no statistics about \( x \) are available or when it is desired that the estimator is unbiased in the classical sense, i.e., it shall fulfill \( E_n[\hat{x}] = x \), where the index denotes the averaging probability density function (PDF). Since the classical estimators BLUE and BWLUE applied on the model in (1) do not result in real valued estimates, a systematic error is introduced.

In this work, a novel classical estimator is derived which allows to incorporate the fact that \( x \) is real valued. This estimator is of widely linear form and unbiased in the classical sense. The derivation is executed by minimizing the variance of the vector estimator’s elements subject to an unbiasedness constraint. This unbiasedness constraint incorporates the fact that \( x \) is real valued. Furthermore, in order to yield real valued estimates, \( e_i^H = f_i^T \) is enforced during the derivation. The resulting widely linear estimator is termed BWLUE for real valued parameter vectors and it in general outperforms the BLUE and the BWLUE in terms of the variances of the vector estimator’s elements. A special case is investigated where the proposed estimator coincides with the real part of the BLUE. The estimator’s performance is validated with a simulation example.
The remainder of this paper is organized as follows: In Sec. 2 the basic concepts necessary for applying widely linear estimators are briefly recapitulated. Sec. 3 contains the derivation of the BWLUE for real parameter vectors, where we focus on proper noise statistics. A discussion about the proposed estimator is presented in Sec. 4 and a simulation example in which the proposed estimator is compared with several other estimators is provided in Sec. 5.

2. PRELIMINARIES FOR WIDELY LINEAR ESTIMATORS

In this section we recapitulate the preliminaries required to apply the widely linear estimators used in this work. This section is more or less a shortened version of the corresponding parts in [1].

We start by constructing the complex augmented vector \( \tilde{a} \) of a vector \( a \in \mathbb{C}^{N_a \times 1} \) by stacking \( a \) on top of its complex conjugate \( a^* \), i.e.

\[
\tilde{a} = \begin{bmatrix} a \\ a^* \end{bmatrix} \in \mathbb{C}^{2N_a \times 1}.
\] (5)

In order to characterize the second-order statistical properties of \( \tilde{a} \) we consider the augmented covariance matrix

\[
C_{\tilde{a}\tilde{a}} = E[(\tilde{a} - E[\tilde{a}])(\tilde{a} - E[\tilde{a}])^H]
\] (6)

\[
= \begin{bmatrix} C_a & C_{a*} \\ C_{a*}^H & C_{a\tilde{a}} \end{bmatrix} = \begin{bmatrix} C_a & C_{a,\tilde{a}} \\ C_{a,\tilde{a}}^H & C_{a\tilde{a}} \end{bmatrix} \in \mathbb{C}^{2N_a \times 2N_a},
\] (7)

with \( C_{\tilde{a}\tilde{a}} = E_{\tilde{a}}[(\tilde{a} - E_{\tilde{a}}[\tilde{a}])(\tilde{a} - E_{\tilde{a}}[\tilde{a}])^H] \) as the (Hermitian and positive semi-definite) covariance matrix and \( C_{\tilde{a}\tilde{a}} = E_{\tilde{a}}[(\tilde{a} - E_{\tilde{a}}[\tilde{a}])(\tilde{a} - E_{\tilde{a}}[\tilde{a}])^T] \) as the complementary covariance matrix. For \( C_{\tilde{a}\tilde{a}} \) and \( C_{a\tilde{a}} \) we have \( C_{aa} = C_{a\tilde{a}}^H \) and \( C_{a\tilde{a}} = C_{\tilde{a}a} \), respectively.

\( C_{\tilde{a}\tilde{a}} \) is sometimes also referred to as pseudo-covariance matrix or conjugate covariance matrix. If \( C_{aa} = 0 \), then the vector \( a \) is called proper, otherwise improper [4]. For scalar Gaussian random variables, properness means that the equipotential lines of its PDF plotted in the complex plane are circles. If those equipotential lines are elliptical, then the scalar Gaussian random variable is improper. If \( a \) is real valued, then \( C_{\tilde{a}\tilde{a}} = C_{a\tilde{a}} \).

Let \( x \) be the parameter vector to be estimated and \( y \) be the measurement vector, then a general widely linear estimator takes on the form in (2) (or an affine version of (2)). In general, widely linear estimators are superior to their linear counterparts as soon as the measurements \( y \) turn improper. Applications for widely linear estimators are investigated in [3][8][10]. Another way to express the estimator in (2) is its augmented version

\[
\hat{x} = \begin{bmatrix} E & F \\ F^* & E^* \end{bmatrix} \begin{bmatrix} y \\ y^* \end{bmatrix} = E \Sigma y.
\] (8)

This work focuses on the BLUE [2], the BWLUE [3] and the WLMSE estimator [3][11] for zero mean parameter vectors given in (9)–(11), respectively.

\[
\hat{x}_B = (H^H C_{nn}^{-1} H)^{-1} H^H C_{nn}^{-1} y,
\] (9)

\[
\hat{x}_{WB} = (H^H C_{nn}^{-1} H)^{-1} H^H C_{nn}^{-1} y,
\] (10)

\[
\hat{x}_{WL} = C_{xx} H^H (H C_{xx} H^H + C_{nn})^{-1} y,
\] (11)

where

\[
H = \begin{bmatrix} H & 0 \\ 0 & H^* \end{bmatrix}.
\] (12)

Note that the BWLUE reduces to the BLUE for proper noise.

3. DERIVATION

In the following, the derivation of the BWLUE for real valued parameter vectors is given, assuming proper noise statistics (\( C_{nn} = 0 \)). An extension to improper noise will be handled in an upcoming paper. We start with some arguments about the BWLUE. The BWLUE for the \( i \)th element of the parameter vector written in the form

\[
\hat{x}_i = e_i^H y + f_i^H y^*
\] (13)

\[
= \begin{bmatrix} e_i^H & f_i^H \end{bmatrix} \begin{bmatrix} y \\ y^* \end{bmatrix}
= w_i^H \Sigma y
\] (14)

\[
= w_i H \Sigma y
\] (15)

can be derived by minimizing the cost function [3]

\[
J = w_i^H C_{nn} w_i,
\] (16)

s.t. the usual unbiasedness constraint \( E_{\Sigma}[\hat{x}_i] = x_i \) which simply follows to

\[
w_i H \Sigma H = u_i^T,
\] (17)

where \( u_i^T \) is a zero row vector of size \( 1 \times 2N_x \) with a ’1’ at its \( i \)th position. To obtain our targeted BWLUE for real parameter vectors we in contrast to the ordinary BWLUE enforce

\[
\Re\{\hat{x}_i\} = 0
\] (18)

\[
E_{\Sigma}[\Re\{\hat{x}_i\}] = E_{\Sigma}[\hat{x}_i] = x_i,
\] (19)

where \( \Re\{\cdot\} \) and \( \Im\{\cdot\} \) denote the real and imaginary part, respectively. From (18) one can easily show that the choice \( e_i^H = f_i^T \) is necessary and sufficient to make \( \hat{x}_i \) real valued independent of the concrete realization of \( y \). Incorporating this result into (19) leads to

\[
E_{\Sigma}[\hat{x}_i] = E_{\Sigma}[e_i^H y + e_i^T y^*]
= e_i^H H x + e_i^T H^* x
\] (20)

\[
= (e_i^H H + e_i^T H^*) x,
\] (21)
hence the unbiased constraint for this estimator is
\[ e_i^TH + e_i^T H^* = u_i^T, \]
with \( u_i^T \) being a zero row vector of size \( 1 \times N_x \) with a ‘1’ at its \( i \)th position. Note that \( (23) \) is a less stringent requirement than \( (17) \). Altogether this leads to the constrained optimization problem
\[
\min e_i^H H + e_i^T H^* = u_i^T, \tag{24}
\]
which can be solved by utilizing the Lagrange multiplier method. The Lagrange cost function follows to
\[
J' = \left[ e_i^H e_i^T \right]_{nn} \left[ e_i^H e_i^T \right]_{nn} + \lambda^T \left( H^H e_i + H^T e_i^* - u_i \right) \]
\[ = e_i^H C_{nn} e_i + e_i^T C_{nn} e_i^* + \lambda^T \left( H^H e_i + H^T e_i^* - u_i \right). \]
Note that the Lagrange multiplier \( \lambda \) is real valued since the constraint is real valued. Taking the partial derivative of \( J' \) w.r.t. \( e_i^T \) (using Wirtinger’s calculus \[12\], i.e. treating \( e_i \) as static) results in
\[
\frac{\partial J'}{\partial e_i^T} = C_{nn} e_i + C_{nn} e_i^* + H \lambda. \tag{25}
\]
Setting this result equal to zero allows to identify
\[
e_i = -\frac{1}{2} C_{nn}^{-1} H \lambda. \tag{26}
\]
Inserting \( (26) \) into the constraint in \( (24) \) yields
\[
\lambda^T = -2u_i^T \left( H^H C_{nn}^{-1} H + H^T (C_{nn}^{-1})^* H^* \right)^{-1}. \tag{27}
\]
Inserting \( (27) \) into \( (26) \) leads to
\[
e_i^H = u_i^T \left( H^H C_{nn}^{-1} H + H^T (C_{nn}^{-1})^* H^* \right)^{-1} H^H C_{nn}^{-1}. \tag{28}
\]
Combining \( (28) \) with \( (3) \) results in
\[
\hat{x}_i = u_i^T \left( H^H C_{nn}^{-1} H + H^T (C_{nn}^{-1})^* H^* \right)^{-1} H^H C_{nn}^{-1} y + u_i^T \left( H^H C_{nn}^{-1} H + H^T (C_{nn}^{-1})^* H^* \right)^{-1} H^T (C_{nn}^{-1})^* y^* \tag{29}
\]
The vector estimator immediately follows to
\[
\hat{x} = E y + E^* y^*, \tag{30}
\]
where the estimator matrix \( E \) follows from \( (28) \) as
\[
E = \left( H^H C_{nn}^{-1} H + H^T (C_{nn}^{-1})^* H^* \right)^{-1} H^H C_{nn}^{-1}. \tag{31}
\]
This is the final result for the BWLUE for real valued parameter vectors and proper noise.

\[
\text{4. DISCUSSION}
\]
In a first step we verify that this result yields an unbiased estimator. The expectation of \( \hat{x} \) follows to
\[
E_n [\hat{x}] = \left( H^H C_{nn}^{-1} H + H^T (C_{nn}^{-1})^* H^* \right)^{-1} H^H C_{nn}^{-1} H x + \left( H^H C_{nn}^{-1} H + H^T (C_{nn}^{-1})^* H^* \right)^{-1} H^T (C_{nn}^{-1})^* H^* x = x. \tag{32}
\]
which proofs that this estimator is unbiased.

Another formulation of \( (30)-(31) \) can be derived as
\[
\hat{x} = \left( 2 \Re \{ H^H C_{nn}^{-1} H \} \right)^{-1} \left( H^H C_{nn}^{-1} y + H^T (C_{nn}^{-1})^* y^* \right) = \left( \Re \{ H^H C_{nn}^{-1} H \} \right)^{-1} \Re \{ H^H C_{nn}^{-1} y \}. \tag{33}
\]
This is the most compact formulation of the BWLUE for real valued parameter vectors.

Assuming the special case where the term \( H^H C_{nn}^{-1} H \) is real valued we obtain
\[
\hat{x} = \left( \Re \{ H^H C_{nn}^{-1} H \} \right)^{-1} \Re \{ H^H C_{nn}^{-1} y \} \tag{34}
\]
\[
= \left( H^H C_{nn}^{-1} H \right)^{-1} \Re \{ H^H C_{nn}^{-1} y \} \tag{35}
\]
\[
= \Re \left( H^H C_{nn}^{-1} H \right)^{-1} \Re \{ H^H C_{nn}^{-1} y \}. \tag{36}
\]
In that case, the BWLUE for real parameter vectors coincides with the real part of the BLUE in \( (9) \). Furthermore, it also coincides with the real part of the BWLUE in \( (10) \) since the noise is assumed to be proper.

The covariance matrix of \( \hat{x} \) can be shown to be
\[
C_{xx} = \Re \{ E C_{nn} E^H + E^* C_{nn} E^T \} \tag{37}
\]
\[
= 2 \Re \{ E C_{nn} E^H \} \tag{38}
\]
\[
= \left( H^H C_{nn}^{-1} H + H^T (C_{nn}^{-1})^* H^* \right)^{-1}. \tag{39}
\]
The expression for the BWLUE for real valued parameter vectors in \( (30)-(31) \) could also have been derived by minimizing the Bayesian mean square error (BMSE) cost function \( E_{y,x} [||\hat{x} - x||^2] \) subject to the same constraint as in \( (24) \), such that this estimator can also be interpreted in a Bayesian sense. In this sense, the Bayesian error covariance matrix \( C_{ee} \), where \( e = \hat{x} - x \), is equal to \( C_{xx} \) in \( (37)-(39) \) and the minimum BMSEs can be found on the main diagonal of \( C_{ee} \).

For every complex linear or widely linear estimator such as the WLMMESE estimator, there exists an equivalent real valued estimator derived from an extended real valued linear model. The formulation as complex linear model and according complex estimator has the advantage of being much more compact. The equivalent real valued model to \( (1) \) is
\[
\begin{bmatrix} \Re \{ y \} \\ \Im \{ y \} \end{bmatrix} = \begin{bmatrix} \Re \{ H \} \\ \Im \{ H \} \end{bmatrix} x + \begin{bmatrix} \Re \{ n \} \\ \Im \{ n \} \end{bmatrix}. \tag{40}
\]
The BLUE applied on this real valued linear model follows to
\[
\hat{x} = \left( \begin{bmatrix} \Re (\{H\}) \\ \Im (\{H\}) \end{bmatrix} \right)^H C_{nn}^{-1} \begin{bmatrix} \Re (\{H\}) \\ \Im (\{H\}) \end{bmatrix}^{-1} \left( \begin{bmatrix} \Re (\{y\}) \\ \Im (\{y\}) \end{bmatrix} \right)
\]

(41)

where \( C_{nn} \) is an appropriately modified covariance matrix of the noise vector in \( (40) \). Eq. (41) is equivalent to \( (30) \) and \( (31) \). However, comparing the matrices to be inverted in \( (33) \) and in \( (41) \) reveals the advantage of the proposed estimator. Hence, applying the proposed estimator allows for a significant complexity reduction.

5. SIMULATION RESULTS

To compare the performance of the derived estimator we consider the case where a discrete-time approximation of a real valued analogue impulse response is estimated based on spectral measurements of an analogue linear and time-invariant (LTI) system. The measurements are performed at equidistant frequency steps \( f = k\Delta f \) with \( k = 0, \ldots, 19 \). The complex measurements assembled in a vector are denoted as \( y \). These measurements represent the single sided frequency response of the analogue LTI system. An extension to a double sided frequency response causes the measurement noises to be perfectly correlated. This would further result in a singular \( C_{nn} \), preventing the application of the BLUE. The discrete-time impulse response \( x \), which is assumed to have a length of \( N_x = 5 \), is connected with the measurements via
\[
y = T_S H x + n.
\]

(42)

In \( (42) \), \( H \) is given by the first 5 columns and the first 20 rows of a DFT matrix of size \( 40 \times 40 \), and \( T_S \) is the sampling time of \( x \) and is necessary in order to combine analogue spectral measurements with a discrete-time impulse response \( (13) \). In the following, \( T_S \) is set to 1 for simplicity.

In order to compare the BWLUE for real valued parameter vectors with the Bayesian WLMMSE estimator, the simulations are carried out in a Bayesian framework where the impulse response \( x \) is randomly generated with \( E[x] = 0 \) and \( C_{xx} = C_{xx} = I \), \( n \) in \( (42) \) is assumed to be a zero mean and proper noise vector with covariance matrix \( C_{nn} = \sigma^2_n I \), and \( \sigma^2_n \) is varied between \( 10^{-3} \) and \( 10^2 \).

The resulting average BMSE values, defined as the average BMSEs of the vector estimator’s elements, are presented in Fig. 1. There, the performance of the BLUE as well as the real part of the BLUE as in \( (36) \) is shown. By taking only the real part of the BLUE, the performance can be slightly increased. The BLUE and the BWLUE coincide since the noise is proper, this is why only the results of the BLUE are shown. A significant increase in performance is provided by the WLMMSE estimator, but the derived BWLUE for real parameter vectors which does not use prior statistical knowledge about the parameters almost reaches the performance of the WLMMSE estimator for noise variances up to \( 10^3 \). In this area, the prior knowledge in form of \( C_{xx} \) and \( C_{xx} \) is weak compared to the information content of the measurements. For larger noise variances, the prior knowledge begins to increase the performance of the WLMMSE estimator compared to the BWLUE for real valued parameter vectors. However, for all values of \( \sigma^2_n \), the proposed estimator outperforms the classical BLUE and BWLUE by approximately two orders of magnitude in BMSE.

6. CONCLUSION

In this work the task of estimating a real valued parameter vector based on complex measurements has been investigated. This task has been analyzed in the classical sense, i.e. for the assumption of a deterministic but unknown parameter vector. A novel classical estimator has been derived which can incorporate the knowledge that the parameter vector is real valued. This estimator outperforms other classical estimators such as the BLUE or the BWLUE. A simulation example was shown revealing that the proposed estimator can compete with the Bayesian WLMMSE estimator, which requires first and second order statistics about the parameter vector. Proper noise statistics have been assumed for the derivation. An extension to improper noise will be handled in upcoming work.
7. REFERENCES

[1] T. Adali, P. Schreier, and L. Scharf, “Complex-Valued Signal Processing: The Proper Way to Deal With Improperity,” In IEEE Transactions on Signal Processing, Vol. 59, No. 11, pp. 5101–5125, Nov 2011.

[2] S. M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, Vol. 1. Prentice Hall, 1993.

[3] P. Schreier and L. Scharf, Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals. Cambridge University Press, 2010.

[4] E. Ollila, “On the Circularity of a Complex Random Variable,” In IEEE Signal Processing Letters, Vol. 15, pp. 841–844, 2008.

[5] J. Eriksson, E. Ollila, and V. Koivunen, “Essential Statistics and Tools for Complex Random Variables,” In IEEE Transactions on Signal Processing, Vol. 58, No. 10, pp. 5400–5408, Oct 2010.

[6] P. J. Schreier and L. L. Scharf, “Second-order analysis of improper complex random vectors and processes,” In IEEE Transactions on Signal Processing, Vol. 51, No. 3, pp. 714–725, March 2003.

[7] M. Novey, T. Adali, and A. Roy, “A Complex Generalized Gaussian Distribution: Characterization, Generation, and Estimation,” In IEEE Transactions on Signal Processing, Vol. 58, No. 3, pp. 1427–1433, March 2010.

[8] J. Via, D. Ramirez, I. Santamaria, and L. Vielva, “Widely and semi-widely linear processing of quaternion vectors,” In Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2010), pp. 3946–3949, March 2010.

[9] P. Chevalier, J.-P. Delmas, and A. Oukaci, “Optimal widely linear MVDR beamforming for noncircular signals,” In Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009), pp. 3573–3576, April 2009.

[10] M. Huemer, O. Lang, and C. Hofbauer, “Component-Wise Conditionally Unbiased Widely Linear MMSE Estimation,” 2016. [Online]. Available: http://www.sigport.org/599

[11] B. Picinbono and P. Chevalier, “Widely linear estimation with complex data,” In IEEE Transactions on Signal Processing, Vol. 43, No. 8, pp. 2030–2033, Aug 1995.

[12] W. Wirtinger, “Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen,” In Mathematische Annalen, Vol. 97, No. 1, pp. 357–375, 1927.

[13] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, Discrete-time Signal Processing (2Nd Ed.). Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1999.