Algebraic Characterization of the Cost Function for Discrete Transversal Filters

Rafael de Carvalho Bluhm and Charles Casimiro Cavalcante

***Abstract*— This article deals with the algebraic structure of the real case cost function, present in the analysis of Wiener filters. The correlation written as a decomposable tensor comes from the isomorphism of the multiplication of inner products and linear operators, and for general bases does not have the same representation in terms of the components. This difference is the object of this study.**

***Keywords*— Cost Function, Correlation, Bilinear Transform, Tensor Product**

I. Introduction

In signal processing, filtering is an important technique for extracting signal information or allowing such signals to be suitable for certain applications. Most of these signals are treated stochastically due to the presence of noise.

One way of designing an optimum filter, specially in linear cases, is by the minimization of the mean square error, which is usually called the cost function. In this approach, the signal is compared to a desired signal and the filter adjusted so we can minimize the effects of the noise.

This work proposes a general algebraic study of the cost function considering real-valued signals, which is used to characterize the mean square error as a function of the coefficients, done by the study of bilinear forms and their relations with tensor products. This will be done in order to verify the algebraic consistency between notations in a simple problem, but which lacks a more formal treatment, sometimes not so obvious.

II. System Model

A transversal discrete filter of finite size N with external signal entries $\mathbf{u} = \{u_1, u_2, ..., u_k, ...u_N\}$ can be characterized by the coefficients w_i, with $u_i = u(n-i)$ and $\mathbf{w} = \{w_1, \ldots, w_N\}$, where $\mathbf{u}, \mathbf{w} \in \mathbb{V}$ and $\dim(\mathbb{V}) = N$. The linear time invariant response of the filter can be expressed as [1]:

$$y(n) = \sum_{k=1}^{N} u(n-k)w_k = \sum_{k=1}^{N} u_k w_k = \langle \mathbf{u}, \mathbf{w} \rangle, \qquad (1)$$

where $\langle \cdot, \cdot \rangle$ is the canonical inner product. Note that $\mathbf{u} : \Omega \to \mathbb{V}$ can be a random variable mapped into the vector space, $\mathbf{w}$ is a deterministic vector and the scalar $y(n)$ is the output signal at every discrete time instant n.

One possible optimization method to find the filter coefficients is by means of the minimization of the mean square error [1]. The quadratic error is calculated as

$$\varepsilon(n) = |e(n)|^2 = |d(n) - y(n)|^2,$$

where $d = d(n)$ is the desired signal sample for a particular time n. By replacing (1) in the previous expression we obtain

$$\varepsilon(n) = |d|^2 - 2\langle \mathbf{w}, d\mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle \langle \mathbf{u}, \mathbf{w} \rangle. \qquad (2)$$

Equation (2) provides the structure of the quadratic error in terms of the inner products with respect to the filter parameters.

In expression (2), the mean is calculated by applying the expectation operator and minimized with respect to the vector $\mathbf{w}$.

But how to calculate the expected value of the last term of the general expression (2)? This is done by the generalized study of the quadratic form $\langle \mathbf{w}, \mathbf{u} \rangle \langle \mathbf{u}, \mathbf{w} \rangle$ which is a particular case of the general bilinear $\langle \mathbf{y}, \mathbf{x} \rangle \langle \mathbf{v}, \mathbf{u} \rangle$ expression.

III. The General Bilinear Form of $\langle \mathbf{y}, \mathbf{x} \rangle \langle \mathbf{w}, \mathbf{u} \rangle$

From now on, the notation that differentiates upper and lower indexes will be used to distinguish vectors in their dual basis from the standard ones.

Given a dual vector $\mathbf{x}^* = \sum_i x_i \mathbf{e}^i \in \mathbb{V}^*$ in dual space (also known as linear functional space) of the standard representation $\mathbf{x} = \sum_i x^i \mathbf{e}_i \in \mathbb{V}$, where the sum occurs up to the dimension of the space. The dual basis has the following identity relation: $\mathbf{e}^i(\mathbf{e}_j) = 1$ if $i = j$ and $= 0$ for $i \neq j$.

A. Inner product representations

The inner product can be expressed as the application of linear functional vectors, given by

$$\langle \mathbf{y}, \mathbf{x} \rangle \langle \mathbf{w}, \mathbf{u} \rangle = \mathbf{x}^*(\mathbf{y})\mathbf{u}^*(\mathbf{w}) = \sum_{i,j} x_i y^i u_j w^j,$$

where $\mathbf{x}^*, \mathbf{u}^* \in \mathbb{V}^*$ and $\mathbf{y}, \mathbf{w} \in \mathbb{V}$. For the pair of inner multiplication we have

$$\langle \mathbf{y}, \mathbf{x} \rangle \langle \mathbf{w}, \mathbf{u} \rangle = \sum_{i,j} x_i u_j y^i w^j = \sum_{i,j} x_i u_j \mathbf{e}^i(\mathbf{y})\mathbf{e}^j(\mathbf{w}) \, . \qquad (3)$$

By the definition of tensor product of functional linear basis which states that $(\mathbf{e}^i \otimes \mathbf{e}^j)(\mathbf{y}, \mathbf{w}) \equiv \mathbf{e}^i(\mathbf{y})\mathbf{e}^j(\mathbf{w})$ [2], the expression in (3) can be written as

$$\langle \mathbf{y}, \mathbf{x} \rangle \langle \mathbf{w}, \mathbf{u} \rangle = \sum_{i,j} x_i u_j (\mathbf{e}^i \otimes \mathbf{e}^j)(\mathbf{y}, \mathbf{w}) = (\mathbf{x}^* \otimes \mathbf{u}^*)(\mathbf{y}, \mathbf{w}), \qquad (4)$$

where $\otimes$ stands for the outer product.

In order to write $\mathbf{x}^* \otimes \mathbf{u}^* \in \mathbb{V}^* \otimes \mathbb{V}^*$ as a linear transformation the following isomorphism is used

$$b(\mathbf{y}, \mathbf{w}) = \langle \mathbf{y}, \mathbf{B}\mathbf{w} \rangle$$

Rafael de Carvalho Bluhm. Center of Telecommunications Engineering and Informatics, Federal University of Ceará (UFC), Fortaleza-CE, Brazil, E-mails: {rafaelbluhm,charles}@gtel.ufc.br.

where the matrix $\mathbf{B}$ for the basis is given by $\mathbf{B}(\mathbf{e}_j) = \sum_k b_j^k \mathbf{e}_k$, where each element is represented by

$$b(\mathbf{e}_i, \mathbf{e}_j) = \langle \mathbf{e}_i, \sum_k b_j^k \mathbf{e}_k \rangle = \sum_k b_j^k \langle \mathbf{e}_i, \mathbf{e}_k \rangle = \sum_k b_j^k g_{ik} \ .$$

The inner product $\langle \mathbf{e}_i, \mathbf{e}_k \rangle \equiv g_{ik}$ is called *metric tensor* and the equivalent representation for dual space $\langle \mathbf{e}^i, \mathbf{e}^k \rangle \equiv g^{ik}$ is called *inverse metric tensor*, given by the identity of relation between the original space and its dual. Also $b_{ij} = x_i u_j$ and $b_j^k = x^k u_j$ which implies that

$$x_i = \sum_k g_{ik} x^k \tag{5}$$

So, we can write

$$\mathbf{x} \otimes \mathbf{u}^* \equiv \sum_{i,j} x^i u_j \mathbf{e}_i \otimes \ \mathbf{e}^j,$$

that belongs to $\mathbb{V}^* \otimes \mathbb{V}$ which is isomorphic to the linear transformation space $\mathfrak{L} : \mathbb{V} \rightarrow \mathbb{V}$ [3]. We can then conclude that

$$\langle \mathbf{y}, \mathbf{x} \rangle \langle \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{y}, \mathbf{x} \otimes \mathbf{u}^*(\mathbf{w}) \rangle, \tag{6}$$

where the functional $\mathbf{u}^*$ acts on $\mathbf{w}$.

The expressions (4) and (6) show that, for both forms to represent the same product, there is a difference in the representation of its elements. This equivalence is given by the presence of the metric acting in each component like in (5).

B. Matrix Representations

For a vector $\mathbf{x} = \sum_{i=1}^N x^i \mathbf{e}_i \ \in \mathbb{V}$ and $\mathbf{u}^* = \sum_{i=1}^N u_i \mathbf{e}^i \ \in \mathbb{V}^*$, the component operations can be given by matrix operations for a given basis $\alpha = \{\mathbf{e}_1, ..., \mathbf{e}_N\}$, expressed as

$$[\mathbf{x} \otimes \mathbf{u}^*]_\alpha = \begin{bmatrix} x^1 \\ \vdots \\ x^N \end{bmatrix} \begin{bmatrix} u_1 & \dots & u_N \end{bmatrix} = \begin{bmatrix} x^1 u_1 & \dots & x^1 u_N \\ \vdots & \ddots & \vdots \\ x^N u_1 & \dots & x^N u_N \end{bmatrix} .$$

We can write that in compact notation $[\mathbf{x} \otimes \mathbf{u}^*]_\alpha = \mathbf{X}\mathbf{U}^* = [x^i u_j]_{N \times N}$, where bold capital letters represent the matrix coordinates for a given basis. The upper index denotes the row and lower one the column, i.e., $\mathbf{X} = [x^i]$ is a column vector and $\mathbf{U}^* = [u_j]$ is a row vector.

For $\mathbf{Y} = [y^i]_{N \times 1}$ and $\mathbf{W} = [w^i]_{N \times 1}$ we write

$$\langle \mathbf{y}, \mathbf{x} \otimes \mathbf{u}^*(\mathbf{w}) \rangle = \langle \mathbf{Y}, \mathbf{X}\mathbf{U}^*\mathbf{W} \rangle,$$

or in the dual representation, by the property that $(\mathbf{A}\mathbf{B})^* = \mathbf{B}^*\mathbf{A}^*$ [2]:

$$\langle \mathbf{y}, \mathbf{x} \otimes \mathbf{u}^*(\mathbf{w}) \rangle = \mathbf{W}^*\mathbf{U}\mathbf{X}^*\mathbf{Y}.$$

Furthermore, for the real symmetry inner product (can be done for the complex case with Hermitian symmetry) we have

$$\langle \mathbf{y}, \mathbf{x} \otimes \mathbf{u}^*(\mathbf{w}) \rangle = \langle \mathbf{x} \otimes \mathbf{u}^*(\mathbf{w}), \mathbf{y} \rangle = \mathbf{Y}^*\mathbf{X}\mathbf{U}^*\mathbf{W}.$$

For the real case, both expressions have the same value, which is

$$\mathbf{W}^*\mathbf{U}\mathbf{X}^*\mathbf{Y} = \mathbf{Y}^*\mathbf{X}\mathbf{U}^*\mathbf{W} = (\mathbf{W}^*\mathbf{U}\mathbf{X}^*\mathbf{Y})^*.$$

IV. Quadratic error

The cost function is just a particular situation of the previous analysis, when $\mathbf{x} = \mathbf{u}$ and $\mathbf{y} = \mathbf{w}$.

From this, the quadratic form is computed and gives the filter last term of expression (2), it means,

$$\langle \mathbf{w}, \mathbf{u} \rangle \langle \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{u} \otimes \mathbf{u}^*(\mathbf{w}) \rangle = \mathbf{W}^*\mathbf{U}\mathbf{U}^*\mathbf{W}$$

which relates the equivalent representations.

A. Mean calculation of quadratic error

Calculating the mean of Equation (1) we have the following expression of the cost function:

$$J = E[\varepsilon] = E[|d|^2] - 2\langle \mathbf{w}, E[d\mathbf{u}] \rangle + E[\langle \mathbf{w}, \mathbf{u} \rangle \langle \mathbf{u}, \mathbf{w} \rangle]. \tag{7}$$

The last (quadratic) term is given by:

$$E[\langle \mathbf{w}, \mathbf{u} \rangle \langle \mathbf{w}, \mathbf{u} \rangle] = \langle \mathbf{w}, E[\mathbf{u} \otimes \mathbf{u}^*](\mathbf{w}) \rangle = \mathbf{W}^* E[\mathbf{U}\mathbf{U}^*]\mathbf{W}. \tag{8}$$

Equation (8) in component terms is given by the second order symmetric tensor called *auto correlation* which can be written as

$$R_{\mathbf{u}\mathbf{u}^*} = E[\mathbf{u} \otimes \mathbf{u}^*] = \sum_{i,j=1}^N E[u^i u_j] \mathbf{e}_i \otimes \mathbf{e}^j \ .$$

The auto correlation written in matrix form is given as

$$R_{\mathbf{u}\mathbf{u}^*} = \begin{bmatrix} E[u^1 u_1] & \dots & E[u^1 u_N] \\ \vdots & \ddots & \vdots \\ E[u^N u_1] & \dots & E[u^N u_N] \end{bmatrix} . \tag{9}$$

A relation for expressions in (4) and in (9) is given by the metric so that

$$R_{\mathbf{u}^*\mathbf{u}^*} = \mathbf{G} R_{\mathbf{u}\mathbf{u}^*},$$

where $\mathbf{G} = [g_{ij}]_{N \times N}$ is the matrix form of the metric tensor and $\mathbf{G}^{-1} = [g^{ij}]_{N \times N}$.

Another relation of the auto correlation can be obtained by the action of the metric matrix, as in $R_{\mathbf{u}^*\mathbf{u}^*} = \mathbf{G} R_{\mathbf{u}\mathbf{u}} \mathbf{G}$. The final expression or the cost function can be expressed as

$$J = E[|d(n)|^2] - 2\langle \mathbf{w}, \mathbf{p} \rangle + \langle \mathbf{w}, R_{\mathbf{u}\mathbf{u}^*} \mathbf{w} \rangle,$$

where $\mathbf{p} = E[d\mathbf{u}]$. From the expression above, the optimal $\mathbf{w}$ can be derived, which is out of the scope of this work.

V. Conclusions

From the study of bilinear forms and its isomorphisms with inner products and linear transformations, one can see the algebraic formulation of the cost function is important in the filter analysis. The formulation allows to explicit the symmetric tensor of the auto correlation, which is important in the analysis of discrete filters, specially in the cases the cost function is the mean square error. Finally, a generalized representation of the filter coefficients in a general base system was provided.

References

[1] Simon Haykin, *Adaptive Filter Theory*. Pretice Hall Information and System Sciences Series, Third Edition.

[2] Elon Lages Lima, *Álgebra Linear*. IMPA, 9ª edição.

[3] Elon Lages Lima, *Cálculo Tensorial*. IMPA, 3ª edição.