Trilinear Wiener Filtering: Application to Equalization Problems

Lucas N. Ribeiro, João C. M. Mota, André L. F. de Almeida

Abstract—This paper presents a trilinear equalizer that is based on the classical Wiener filter. We review the theory of multiway arrays and then apply it in constructing a model of operation for the filter. However, the cost function of the estimation problem in the model is nonlinear. This problem overcame by using a property of the tensor versus vector product, allowing us to divide the nonlinear optimization problem in three linear problems. Then we develop an iterative algorithm for obtaining the optima filters for the problems. The performance of the filter is evaluated in two applications problems.

Keywords—Wiener Filtering, Multilinear Algebra, Alternating Least-Squares, PARAFAC.

I. INTRODUCTION

In the past years, multidimensional models have been used to solve signal processing problems applied in many fields like psychometrics [7], data analysis [11], chemometrics [16], educametrics [15] and communications systems [5]. Sidiropoulos on his paper [14] described a blind PARAFAC receiver that explored the diversity of DS-CDMA systems for estimating the factor matrices of the transmitted datacubes. In his paper [13], Muti et al. considered multiway filtering, using the subspace method and multimode Principal Component Analysis (PCA). His second approach on filtering was extending classical Wiener to tensor data. He applied the multiway Wiener filter on digital image processing and multicomponent seismic data.

In his application of the multiway Wiener filter, a colored image is modelled as a third order tensor \(\mathcal{X}\). Given a received tensor \(\mathcal{R}\), an estimation \(\hat{\mathcal{X}}\) is generated by successively filtering the received tensor by \(N\) \(n\)-mode filters. The filtering is performed by multiplying a filter matrix by each slice of the input tensor. The optimization criterion used to determine the optimal \(n\)-mode filters is the minimization of the mean-squared error (MSE) between the desired signal \(\mathcal{X}\) and the estimation \(\hat{\mathcal{X}}\).

O. Filiz [6] proposed an alternating adaptive algorithm for estimating rank constrained spatial-temporal filters, since the solution for the optimization problem has not closed form. Using this same idea of dividing a complex optimization problem in simple problems, we propose a trilinear filter that instead of estimating the factor matrices of information about the channel, we devise the optimum trilinear filter that minimizes the error in the mean square sense. The three factors vectors of the filter will be obtained by solving the classical Wiener-Hopf equations, as we will see in the development.

The paper is organized as follows: section II presents some concepts about tensors, in section III, we describe a model for transmitting signal throughout trilinear systems and noise. Section IV describes the algebraic development to obtain the Wiener solutions and V describes the trilinear alternating least-squares algorithm for obtaining the optimal solutions. Finally, section VI shows some applications and we conclude the paper on section VII.

II. MULTI-WAY ARRAYS

A \(N\)th order tensor is a multidimensional array whose entries are accessed via \(N\) indices [13]. It can be represented by \(\mathcal{U} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}\) for example. If \(\mathcal{U}\) has rank one, it can be written as the outer product "\(\circ\)" of \(N\) vectors, i.e.:

\[
\mathcal{U} = u_1 \circ u_2 \circ \ldots \circ u_N
\]

where \(u_1 \in \mathbb{R}^{I_1}, u_2 \in \mathbb{R}^{I_2}, \ldots, u_N \in \mathbb{R}^{I_N}\) are its factor vectors. The rank of a tensor \(\mathcal{X}\) is defined as the smallest number of rank-one tensors that generates \(\mathcal{X}\) as their sum [9]. For example, consider that \(\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}\) has rank \(R\), then we can decompose it as:

\[
\mathcal{X} = \sum_{r=1}^{R} x_{1,r} \circ x_{2,r} \circ x_{3,r}
\]

where \(R\) is a positive integer and \(x_{1,r}, x_{2,r}, x_{3,r}\) are the factor vectors of the \(r\)th rank-one tensor. The equation (2) is known as the PARAFAC (parallel factors) decomposition [2].

Hereafter, the notation for the \(n\)th element of a certain vector \(z\), will be denoted by \(z^{(n)}\). Then, the element of a rank \(R\) third order tensor \(\mathcal{X}\) indexed by \((i,j,k)\):

\[
x_{ijk} = \sum_{r=1}^{R} a_{1,r}^{(i)} a_{2,r}^{(j)} a_{3,r}^{(k)}
\]

A tensor can be also decomposed in slices, which are bidimensional sections defined by fixing all but two indices. Consider a third order tensor \(\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}\) that can be composed like this: \(\mathcal{X} = x_1 \circ x_2 \circ x_3\). Its slices through the three modes are: \(X_{.:i,.:} \in \mathbb{R}^{I_2 \times I_3}\), \(X_{.:j,.:} \in \mathbb{R}^{I_1 \times I_3}\) and \(X_{.:k,.:} \in \mathbb{R}^{I_1 \times I_2}\). Then, the unfolded of a third order tensor, in the following way:

\[
X_{(1)} = [X_{1,.,.,}, \ldots, X_{I_1,.,.}]^T \in \mathbb{R}^{I_1 I_2 \times I_3}
\]

\[
X_{(2)} = [X_{.,1,.,}, \ldots, X_{.,I_2,}]^T \in \mathbb{R}^{I_1 I_3 \times I_2}
\]
These unfolded matrices can be represented by the following equations:

\[
X_{(1)} = \begin{pmatrix} x_1 & x_2 \end{pmatrix}^T, \\
X_{(2)} = \begin{pmatrix} x_2 & x_3 \end{pmatrix}^T, \\
X_{(3)} = \begin{pmatrix} x_3 & x_1 \end{pmatrix}^T
\]

where \(X_{(n)}\) is the tensor \(X\) unfolded in the mode-\(n\).

These unfolded matrices can be represented by the following equations:

\[
X_{(1)} = (x_1 \otimes x_2)x_3^T, \\
X_{(2)} = (x_2 \otimes x_3)x_1^T, \\
X_{(3)} = (x_3 \otimes x_1)x_2^T
\]

where \(\otimes\) denotes the Kronecker product.

The \(n\)-mode product of a \(N\)th order tensor \(X\) \(R \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}\) with a vector \(v \in \mathbb{R}^I\), denoted by \(X \times v\), results into a tensor \(Y\) of order \(N-1\) whose size is \(I_1 \times \ldots \times I_{n-1} \times I_{n+1} \times \ldots \times I_N\) [1]. Its elementwise representation is:

\[
y_{i_1 \ldots i_{n-1}i_{n+1} \ldots i_N} = \sum_{i_n=1}^{I_N} T_{i_1i_2 \ldots i_N} v_{i_n}
\]

In this paper, we will utilize the consecutive \(n\)-mode vector product between a rank \(R\) third order tensor \(X \in \mathbb{R}^{I_1 \times I_2 \times I_3}\) and three vectors \(v_1 \in \mathbb{R}^{I_1}\), \(v_2 \in \mathbb{R}^{I_2}\) and \(v_3 \in \mathbb{R}^{I_3}\). Since each \(n\)-mode vector product decreases one order of the resulting tensor, multiplying a tensor by three vectors results into a zeroth order tensor, i.e. a scalar. This consecutive product is written as:

\[
y = X \times_1 v_1 \times_2 v_2 \times_3 v_3
\]

where \(y\) is a real scalar. Using equation (10), we can write the elementwise representation of \(X_{ijk}\), equation (3):

\[
y = \sum_{i=1}^{I_1} \sum_{j=1}^{I_2} \sum_{k=1}^{I_3} v_1^{(i)} v_2^{(j)} v_3^{(k)} x_{ijk}
\]

Substituting the equation of the elementwise representation for \(X\), equation 3, into equation (12):

\[
y = \sum_{r=1}^{R} \sum_{i=1}^{I_1} \sum_{j=1}^{I_2} \sum_{k=1}^{I_3} v_1^{(i)} v_2^{(j)} v_3^{(k)} x_{1r}^{(i)} x_{2r}^{(j)} x_{3r}^{(k)}
\]

\[
y = \sum_{r=1}^{R} (v_1^T x_{1r}) (v_2^T x_{2r}) (v_3^T x_{3r})
\]

Applying the commutative property of the scalar product between the vectors and rearranging the factors:

\[
y = \sum_{r=1}^{R} (v_1^T v_2) (x_{1r}^T v_1)
\]

where \(U_{1r} = x_{1r}^T v_2\). Applying the \(\text{vec}()\) operator on equation (15), and knowing that [12]:

\[
\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B)
\]

we have that equation (15) becomes:

\[
y = \sum_{r=1}^{R} \text{vec}(v_1^T U_{1r} v_2) (X_{1r}^T v_1)
\]

Since \(\text{vec}(U_{1r}) = x_{2r} \otimes x_{3r}\), equation (17) turns into:

\[
y = \sum_{r=1}^{R} (v_2 \otimes v_3)^T (x_{2r} \otimes x_{3r}) X_{1r}^T v_1
\]

substituting equation (8), we obtain:

\[
y = \sum_{r=1}^{R} (v_2 \otimes v_3)^T X_{1r}^{(r)} v_1
\]

where \(X_{1r}^{(r)}\) is the mode 2 unfolding of the \(r\)th rank one tensor of the PARAFAC decomposition of \(X\). The equation (19) brings an important result, for it is the matricial product version of the \(n\)-mode vector product. If we initiate this development isolating either \((x_1^T v_2)\) or \((x_3^T v_3)\) on equation (14) and repeating the development, we will obtain respectively these equations:

\[
y = \sum_{r=1}^{R} (v_3 \otimes v_1)^T X_{1r}^{(r)} v_2
\]

\[
y = \sum_{r=1}^{R} (v_1 \otimes v_2)^T X_{1r}^{(r)} v_3
\]

The equations (19), (20) and (21) will be important in section IV.

### III. THE TRILINEAR FILTERS PROCESSING MODEL

Consider that we would like to transmit the real set of sequences \(\{s_1(n), s_2(n), \ldots, s_R(n)\}\) through the rank-one trilinear systems \(\{H_1, H_2, \ldots, H_R\}\) where \(H_r \in \mathbb{R}^{I_1 \times I_2 \times I_3}\). For each tensor system has unitary rank, it can be decomposed as:

\[
H_r = h_{a,r} \circ h_{b,r} \circ h_{c,r}
\]

where \(h_{a,r} \in \mathbb{R}^{I_1}\), \(h_{b,r} \in \mathbb{R}^{I_2}\) and \(h_{c,r} \in \mathbb{R}^{I_3}\) are the factor vectors of \(H_r\) and \(R\) denotes the number of sources. The output of the \(r\)th trilinear system is:

\[
U_r(n) = H_r s_r(n) = (h_{a,r} \circ h_{b,r} \circ h_{c,r}) s_r(n)
\]

The operation on equation (23) is a product between a scalar and a tensor, resulting into a tensorial sequence \(U_r(n) \in \mathbb{R}^{I_1 \times I_2 \times I_3}\).

Let us consider an additive white, independent-from-signal and Gaussian noise tensor \(B \in \mathbb{R}^{I_1 \times I_2 \times I_3}\). The white Gaussian noise assumption can be formulated as

\[
E[b_{i_1i_2i_3}] = \delta_{i_1j_1} \delta_{i_2j_2} \delta_{i_3j_3}
\]

where \(i_k\) and \(j_k\) \(\in \{1, \ldots, I_k\}\), \(k \in \{1, 2, 3\}\) and \(\delta\) is the Kronecker delta. The outputs of the trilinear systems are combined with the noise tensor \(B\). Then, the received signal \(X(n) \in \mathbb{R}^{I_1 \times I_2 \times I_3}\) is:

\[
X(n) = \sum_{r=1}^{R} s_r(n) H_r + B
\]
Observe that now $R$ represents both the number of sources and the rank of the tensor $\mathcal{X}$, whose elementwise representation is:

$$x_{ijk}(n) = \sum_{r=0}^{R} h_{a,r}^{(i)} h_{b,r}^{(j)} h_{c,r}^{(k)} s_r(n) + b_{ijk}$$

(26)

where $b_{ijk}$ is the elementwise representation of $\mathcal{B}$. We desire to recover the real sequence $s_r(n)$, then we will filter the signal $\mathcal{X}(n)$ with a rank-1 trilinear system

$$\mathcal{W} = w_a \circ w_b \circ w_c$$

(27)

where $w_a \in \mathbb{R}^{I_1}$, $w_b \in \mathbb{R}^{I_2}$ and $w_c \in \mathbb{R}^{I_3}$ are its factor vectors. It produces an output $\hat{s}_r(n) \in \mathbb{R}$ that is compared to the desired signal $s_r(n)$, generating an estimation error $e_r(n)$. The coefficients of the factor vectors of $\mathcal{W}$ are tuned according to an algorithm that minimizes an optimization criterion based on $e_r(n)$.

Since $\mathcal{W}$ has unitary rank, we define the output of the filter as the $n$-mode vector product between the input tensor and the factor vectors of the filter, resulting into the estimated signal:

$$\hat{s}_r(n) = \mathcal{X}(n) \times_1 w_a \times_2 w_b \times_3 w_c$$

(28)

IV. THE TRILINEAR WIENER FILTERS

The equation (28) can be expressed using the equations (19), (20) and (21), resulting into:

$$\hat{s}_r(n) = \sum_{r=1}^{R} (w_b \odot w_c)^T X_{(2)}^{(r)}(n) w_a$$

(32)

$$= \sum_{r=1}^{R} (w_c \odot w_a)^T X_{(3)}^{(r)}(n) w_b$$

(33)

$$= \sum_{r=1}^{R} (w_a \odot w_b)^T X_{(1)}^{(r)}(n) w_c$$

(34)

We can use the equations (32), (33) and (34) to obtain three linear problems from the problem (31). It can be done by fixing the two parameters in the Kronecker product. Since we fix two parameters, we affirm that the following vectors are equivalent different ways, as we have seen through equations (19), (20) and (21).

Rewriting the equations (32), (33) and (34) using the equations (35), (36) and (37), we can obtain three LS estimations:

$$\tilde{u}_a(n) = \sum_{r=1}^{R} \left( X_{(2)}^{(r)} \right)^T (w_b \odot w_c)$$

(35)

$$\tilde{u}_b(n) = \sum_{r=1}^{R} \left( X_{(3)}^{(r)} \right)^T (w_c \odot w_a)$$

(36)

$$\tilde{u}_c(n) = \sum_{r=1}^{R} \left( X_{(1)}^{(r)} \right)^T (w_a \odot w_b)$$

(37)

It happens that the equations (38), (39) and (40) are classical least squares (LS) estimations and its solutions are [8]:

$$\tilde{w}_a = R_a^{-1} p_a$$

(41)

$$\tilde{w}_b = R_b^{-1} p_b$$

(42)

$$\tilde{w}_c = R_c^{-1} p_c$$

(43)

where

$$R_a = E[\tilde{u}_a(n) \tilde{u}_a(n)^T]$$

(44)

$$R_b = E[\tilde{u}_b(n) \tilde{u}_b(n)^T]$$

(45)

$$R_c = E[\tilde{u}_c(n) \tilde{u}_c(n)^T]$$

(46)

are the covariance matrices of the input filters $w_a$, $w_b$, $w_c$ respectively and

$$p_a = E[u_a s_r(n)]$$

(47)

$$p_b = E[u_b s_r(n)]$$

(48)

$$p_c = E[u_c s_r(n)]$$

(49)

where $p_a$, $p_b$ and $p_c$ are the cross correlation vectors between the input of the filters $w_a$, $w_b$, $w_c$ and the desired signal $s_r(n)$.
V. TRILINEAR ALTERNATING LEAST-SQUARES (TALS)

In order to obtain the optimum filters, we solve, in an alternate way, each problem using the LS estimates (38), (39) and (40), fixing all parameters but the argument. This way, the sequences \( \{ \hat{w}_a^k \}_{k \in \mathbb{N}}, \{ \hat{w}_b^k \}_{k \in \mathbb{N}} \) and \( \{ \hat{w}_c^k \}_{k \in \mathbb{N}} \) may converge to its Wiener solution after \( k \) TALS steps. The condition for convergence is that the norm of the vector \( \Delta \) is less then a certain positive value \( \varepsilon \). The TALS algorithm is summarized in the following steps:

1) Initialize \( \hat{w}_a(0), \hat{w}_b(0), \hat{w}_c(0) \) and \( k = 1 \).
2) ALS loop: while \( ||\Delta|| > \varepsilon \)
   a) Using \( \hat{w}_b(k), \hat{w}_c(k) \), estimate \( \hat{w}_a(k) = \hat{R}_a^{-1} \hat{p}_{da} \)
   b) Using \( \hat{w}_c(k), \hat{w}_a(k) \), estimate \( \hat{w}_b(k) = \hat{R}_b^{-1} \hat{p}_{db} \)
   c) Using \( \hat{w}_a(k), \hat{w}_b(k) \), estimate \( \hat{w}_c(k) = \hat{R}_c^{-1} \hat{p}_{dc} \)
   d) Calculate \( \Delta = \mathcal{W}(k) - \mathcal{W}(k-1) \)
   e) \( k = k + 1 \)

where \( \mathcal{W} \) is given by equation (27). In order to calculate the covariance matrices and crosscorrelation vector estimates, the input vector is calculated (equations 35, 36, 37), and then the covariance matrix (equations 44, 45, 46) and the crosscorrelation vector are calculated (equations 47, 48, 49).

VI. APPLICATIONS

In this section, we expose two applications of the trilinear Wiener filtering in signal processing problems. The SNR is defined as

\[
SNR = 10 \log \left( \frac{E_s}{||B||^2_F} \right)
\]

where \( E_s \) is the energy of the transmitted signal and \( ||B||^2_F \) is the Frobenius norm of the noise tensor. The first application is the recovery of a sinusoid using the model described in the section III and the TALS algorithm (section V) with \( \varepsilon = 10^{-4} \). In the first simulation, each source \( r \) transmits a sinusoid whose frequency is \( (10 \cdot r) \) Hz. The recovery was done using a \( 4 \times 4 \times 4 \) trilinear filter in a multi-source environment of 4 sources. We have recovered the sinusoid of 30 Hz as depicted in figure (2). In the second simulation, we have analyzed the MSE vs. SNR performance of three filters, as shown on figure 3. It can be seen that as we increase the order of the filter, it performs better, since it can process more information in a single run.

The second application is the transmission of BPSK signals by \( R = 4 \) sources. The equalizer will try to estimate the transmitted signals using the TALS algorithm for \( \varepsilon = 10^{-4} \) and we will analyze the SNR vs. BER performance. Like in the first application, as we increase the order of the filter, it performs better in the SNR vs. BER. Through figure 5, we observe that as we increase the number of transmitting sources, the interference increases, then the filter performs worse.

VII. CONCLUSIONS

Multilinear algebra applied to signal processing has been a promising tool and has allowed the exploration of a new field of research and applications. In this paper, we managed to mix the classical Wiener filter with a multilinear structure. Instead of solving a non-linear problem, we have divided the problem in three linear problems that were solved using LS estimators. The trilinear Wiener filter works under the assumption that the received signal has been corrupted by a trilinear system and summed up to a trilinear noise. Under these assumptions, we have shown through simulations that it performs satisfactorily. As we increase the order of the equalizer, it performs better and better. But as we increase the rank of the received tensor signal, the equalizer needs to deal with more interference. There can be some improvements to the trilinear filter like increasing the rank of the equalizer filter and deriving an adaptive model for the trilinear vector.
Fig. 4. Simulation 3: BER vs. SNR performance for 4 users.

Fig. 5. Simulation 4: BER vs. SNR performance for a $10 \times 10 \times 10$ filter.

REFERENCES


