A Link between Decomposition of a Tensor in Multilinear Rank-$(L, L, 1)$ Terms and Simultaneous Block Diagonalization

Lieven De Lathauwer

Tech. Report, ESAT-SISTA, KU Leuven

March 31, 2012

Decomposition

$$\mathcal{T} = \sum_{r=1}^{R} (\mathbf{A}_r \cdot \mathbf{B}_r^T) \circ \mathbf{c}_r. \tag{1}$$

$$\mathbf{T}_{KI \times J} = (\mathbf{C} \otimes \mathbf{A}) \cdot \mathbf{B}^T. \tag{2}$$

Condition

We use the same bilinear mapping Φ as in [1]. The decomposition is essentially unique if the following three conditions are satisfied:

1. $LR \leqslant J$ and $\mathbf{B}$ has full column rank.

2. $LR \leqslant IJ$ and $\mathbf{C} \otimes \mathbf{A}$ has full column rank.

3. the tensors $\Phi\left(\mathbf{c}_\alpha \mathbf{A}_\alpha(:,l)^T,\ \mathbf{c}_\beta \mathbf{A}_\beta(:,m)^T\right)$, $1 \leqslant \alpha < \beta \leqslant R$, $1 \leqslant l, m \leqslant L$, are linearly independent.

If these conditions are satisfied, then it is clear from (2) that the rank of $\mathbf{T}_{KI\times J}$ is equal to LR.

Formulation as simultaneous block diagonalization

Economy-size SVD:

$$\mathbf{T}_{KI\times J} = \mathbf{U}\cdot\mathbf{\Sigma}\cdot\mathbf{V}^H = \mathbf{E}\cdot\mathbf{V}^H, \tag{3}$$

(Note that we split up in $(\mathbf{U}\cdot\mathbf{\Sigma})\cdot\mathbf{V}^H$ and not $\mathbf{U}\cdot(\mathbf{\Sigma}\cdot\mathbf{V}^H)$ or $(\mathbf{U}\cdot\mathbf{\Sigma}^{\frac{1}{2}})\cdot(\mathbf{\Sigma}^{\frac{1}{2}}\cdot\mathbf{V}^H)$. The reason is that if the components of $\mathbf{B}$ are uncorrelated (which is often the case in the long mode), matrix $\mathbf{W}$ below is in principle orthogonal. Hence, the SD problem below is well-conditioned.)

Combination of (2) and (3) yields:

$$\begin{aligned} \mathbf{E} &= (\mathbf{C}\otimes\mathbf{A})\cdot\mathbf{W}^{-1} && (4)\\ \mathbf{B} &= \mathbf{V}^*\cdot\mathbf{W}^{-T} && (5) \end{aligned}$$

in which $\mathbf{W} \in \mathbb{C}^{LR\times LR}$ is nonsingular. It is clear that $\mathbf{W}$ can only be determined up to right multiplication with a block diagonal matrix involving nonsingular $(L\times L)$ blocks. This is consistent with the ambiguities in $\mathbf{A}$ and $\mathbf{B}$. Consider the partitioning $\mathbf{W} = [\mathbf{W}_1 \ldots \mathbf{W}_R]$, in which $\mathbf{W}_r \in \mathbb{C}^{LR\times L}$.

We compute, for all $r, s \in [1, LR]$:

$$\begin{aligned} \mathcal{P}_{rs} &= \Phi(\mathbf{E}_r, \mathbf{E}_s) \\ &= \sum_{u,v=1}^{LR} (\mathbf{W}^{-1})_{ur}(\mathbf{W}^{-1})_{vs}\Phi\left([\mathbf{C}\otimes\mathbf{A}](:,u), [\mathbf{C}\otimes\mathbf{A}](:,v)\right). \end{aligned}$$

We introduce new indices $\alpha, \beta \in [1, R]$ and $l, m \in [1, L]$ such that $u = (\alpha - 1)L + l$ and $v = (\beta - 1)L + m$. Then we have:

$$\mathcal{P}_{rs} = \sum_{l,m=1}^{L} \sum_{\alpha,\beta=1}^{R} (\mathbf{W}^{-1})_{(\alpha-1)L+l,r} (\mathbf{W}^{-1})_{(\beta-1)L+m,s} \Phi\left(\mathbf{c}_\alpha \mathbf{A}_\alpha(:,l)^T,\ \mathbf{c}_\beta \mathbf{A}_\beta(:,m)^T\right). \tag{6}$$

Now assume that there exists a symmetric matrix $\mathbf{M} \in \mathbb{C}^{LR \times LR}$ such that

$$\sum_{r,s=1}^{LR} m_{rs} \mathcal{P}_{rs} = \mathcal{O}. \tag{7}$$

Substitution of (6) in (7) yields:

$$\sum_{r,s=1}^{LR} \sum_{l,m=1}^{L} \sum_{\alpha,\beta=1}^{R} (\mathbf{W}^{-1})_{(\alpha-1)L+l,r} (\mathbf{W}^{-1})_{(\beta-1)L+m,s} m_{rs} \Phi\left(\mathbf{c}_\alpha \mathbf{A}_\alpha(:,l)^T,\ \mathbf{c}_\beta \mathbf{A}_\beta(:,m)^T\right) = \mathcal{O}. \tag{8}$$

We have that $\Phi\left(\mathbf{c}_\alpha \mathbf{A}_\alpha(:,l)^T,\ \mathbf{c}_\beta \mathbf{A}_\beta(:,m)^T\right) = \mathcal{O}$ if $\alpha = \beta$, regardless of the values of l and m. Our third working condition now implies that

$$\mathbf{W}^{-1} \cdot \mathbf{M} \cdot \mathbf{W}^{-T} = \mathbf{D}, \tag{9}$$

in which $\mathbf{D}$ is a block diagonal matrix involving $(L \times L)$ blocks. Equivalently,

$$\mathbf{M} = \mathbf{W} \cdot \mathbf{D} \cdot \mathbf{W}^T. \tag{10}$$

It can be verified that any matrix $\mathbf{M}$ that has the structure (10), satisfies (7). Hence, there are L^2R symmetric matrices in the kernel of (7), which can all be decomposed as in (10):

$$\begin{aligned} \mathbf{M}_1 &= \mathbf{W} \cdot \mathbf{D}_1 \cdot \mathbf{W}^T \\ &\vdots \\ \mathbf{M}_{L^2R} &= \mathbf{W} \cdot \mathbf{D}_{L^2R} \cdot \mathbf{W}^T. \end{aligned} \tag{11}$$

Let $\mathbf{D}_n = \text{blockdiag}(\mathbf{D}_n^{(1)}, \ldots, \mathbf{D}_n^{(R)})$, where $\mathbf{D}_n^{(r)} \in \mathbb{C}^{L \times L}$, $1 \leqslant n \leqslant L^2R$, $1 \leqslant r \leqslant R$. The formulation in terms of a block diagonalization problem is consistent with the fact that $\mathbf{W}$ can only be determined up to right multiplication with a block diagonal matrix. The advantage of (11) over (1) is

that (11) is overdetermined, while (1) is possibly underdetermined. It can be verified that the decomposition in rank-$(L_r, L_r, 1)$ terms can be computed by means of simultaneous block diagonalization of $\sum_{r=1}^{R} L_r^2$ matrices, where the blocks have dimension $(L_r \times L_r)$.

Stack all the matrices $\mathbf{M}_n$ in a tensor $\mathcal{M} \in \mathbb{C}^{LR \times LR \times L^2 R}$. The simultaneous block diagonalization problem is equivalent with the decomposition of $\mathcal{M}$ in rank-$(L, L, \cdot)$ terms, as defined in [2]. The decomposition of $\mathcal{T}$ in rank-$(L_r, L_r, 1)$ terms can be reformulated as the decomposition of a tensor $\mathcal{M} \in \mathbb{C}^{(\sum_{r=1}^{R} L_r) \times (\sum_{r=1}^{R} L_r) \times (\sum_{r=1}^{R} L_r^2)}$ in rank-$(L_r, L_r, \cdot)$ terms.

Solution by means of simultaneous generalized block Schur decomposition

Substitute in (11) the QR-factorization $\mathbf{W} = \mathbf{Q}^H \cdot \mathbf{R}'$ and the RQ-factorization $\mathbf{W}^T = \mathbf{R}'' \cdot \mathbf{Z}^H$. Denote the diagonal $(L \times L)$ blocks of $\mathbf{R}'$ and $\mathbf{R}''$ by $\mathbf{R}^{(r)'}$ and $\mathbf{R}^{(r)''}$, respectively, $1 \leqslant r \leqslant R$. Since $\mathbf{W}$ is nonsingular, $\mathbf{R}'$ and $\mathbf{R}''$ are nonsingular. This, in turn, implies that all $\mathbf{R}^{(r)'}$ and $\mathbf{R}^{(r)''}$ are nonsingular. We obtain:

$$\begin{aligned} \mathbf{Q} \cdot \mathbf{M}_1 \cdot \mathbf{Z} &= \mathbf{R}' \cdot \mathbf{D}_1 \cdot \mathbf{R}'' = \mathbf{R}_1 \\ &\vdots \\ \mathbf{Q} \cdot \mathbf{M}_{L^2R} \cdot \mathbf{Z} &= \mathbf{R}' \cdot \mathbf{D}_{L^2R} \cdot \mathbf{R}'' = \mathbf{R}_{L^2R}, \end{aligned} \tag{12}$$

in which $\mathbf{R}_n$ is block upper triangular. This simultaneous generalized block Schur decomposition can be computed by means of a block version of the extended QZ-iteration in [3]. The matrices $\mathbf{Q}$ and $\mathbf{Z}$ should be updated block by block.

The question is now how $\mathbf{W}$ can be found, once (12) has been computed.

Stack $\mathcal{M}$ in an $(L^2R^2 \times L^2R)$ matrix $\tilde{\mathbf{M}}$. From (11) we have:

$$\tilde{\mathbf{M}} = [\mathbf{W}_1 \otimes \mathbf{W}_1 \ \ldots \ \mathbf{W}_R \otimes \mathbf{W}_R] \cdot \tilde{\mathbf{D}}, \tag{13}$$

where $\tilde{\mathbf{D}} \in \mathbb{C}^{L^2R \times L^2R}$ is defined by

$$\tilde{\mathbf{D}} = \begin{bmatrix} \mathrm{vec}(\mathbf{D}_1^{(1)}) & \cdots & \mathrm{vec}(\mathbf{D}_{L^2R}^{(1)}) \\ \vdots & & \vdots \\ \mathrm{vec}(\mathbf{D}_1^{(R)}) & \cdots & \mathrm{vec}(\mathbf{D}_{L^2R}^{(R)}) \end{bmatrix}. \tag{14}$$

The matrix $\tilde{\mathbf{M}}$ is full column rank since $\mathbf{M}_1, \ldots, \mathbf{M}_{L^2R}$ are linearly independent. The matrix $[\mathbf{W}_1 \otimes \mathbf{W}_1 \ \ldots \ \mathbf{W}_R \otimes \mathbf{W}_R]$ consists of columns of $\mathbf{W} \otimes \mathbf{W}$, which is full rank since $\mathbf{W}$ is full rank. Hence, $\tilde{\mathbf{D}}$ is full rank.

Denote the diagonal $(L \times L)$ blocks of $\mathbf{R}_n$ by $\mathbf{R}_n^{(r)}$, $1 \leqslant n \leqslant L^2R$, $1 \leqslant r \leqslant R$. We have:

$$\mathbf{R}_n^{(r)} = \mathbf{R}^{(r)'} \cdot \mathbf{D}_n^{(r)} \cdot \mathbf{R}^{(r)''}. \tag{15}$$

In vectorized form we have:

$$\mathrm{vec}(\mathbf{R}_n^{(r)}) = \left((\mathbf{R}^{(r)''})^T \otimes \mathbf{R}^{(r)'} \right) \cdot \mathrm{vec}(\mathbf{D}_n^{(r)}). \tag{16}$$

Stack the block diagonals of $\mathbf{R}_n$ in an $(L^2R \times L^2R)$ matrix $\tilde{\mathbf{R}}$, the columns corresponding to different values of n. From (16) we have:

$$\tilde{\mathbf{R}} = \mathrm{blockdiag}((\mathbf{R}^{(1)''})^T \otimes \mathbf{R}^{(1)'}, \ldots, (\mathbf{R}^{(R)''})^T \otimes \mathbf{R}^{(R)'}) \cdot \tilde{\mathbf{D}}, \tag{17}$$

The block diagonal factor is nonsingular because all $\mathbf{R}^{(r)'}$ and $\mathbf{R}^{(r)''}$ are nonsingular. Combination of (13) and (17) yields:

$$\begin{gathered} \tilde{\mathbf{M}} \cdot \tilde{\mathbf{R}}^{-1} = \\ [\mathbf{W}_1 \otimes \mathbf{W}_1 \ \ldots \ \mathbf{W}_R \otimes \mathbf{W}_R] \cdot [\mathrm{blockdiag}((\mathbf{R}^{(1)''})^T \otimes \mathbf{R}^{(1)'}, \ldots, (\mathbf{R}^{(R)''})^T \otimes \mathbf{R}^{(R)'})]^{-1} = \\ [(\mathbf{W}_1 \otimes \mathbf{W}_1) \cdot ((\mathbf{R}^{(1)''})^{-T} \otimes (\mathbf{R}^{(1)'})^{-1}) \ \ldots \ (\mathbf{W}_R \otimes \mathbf{W}_R) \cdot ((\mathbf{R}^{(R)''})^{-T} \otimes (\mathbf{R}^{(R)'})^{-1})] \end{gathered} \tag{18}$$

Actually, the right-hand side of (18) should be obtained by solving the linear set of equations $\tilde{\mathbf{M}} = \mathbf{Y} \cdot \tilde{\mathbf{R}}$ for $\mathbf{Y} \in \mathbb{C}^{L^2R^2 \times L^2R}$. We map the rth $(L^2R^2 \times L^2)$ submatrix of this partitioned matrix to $\mathcal{W}_r \in \mathbb{C}^{LR \times LR \times L^2}$, $1 \leqslant r \leqslant R$. Denote the pth column of $(\mathbf{R}^{(r)''})^{-T} \otimes (\mathbf{R}^{(r)'})^{-1}$ by $\mathbf{x}_p^{(r)}$, $1 \leqslant r \leqslant R$, $1 \leqslant p \leqslant L^2$. Define $\tilde{\mathcal{X}}_r$ as the $(L \times L \times L^2)$ tensor in which the matrices $\mathrm{unvec}(\mathbf{x}_1^{(r)}), \ldots, \mathrm{unvec}(\mathbf{x}_{L^2}^{(r)})$ are stacked. Then we have:

$$\mathcal{W}_r = \tilde{\mathcal{X}}_r \bullet_1 \mathbf{W}_r \bullet_2 \mathbf{W}_r. \tag{19}$$

The column space of $\mathbf{W}_r$ equals the mode-1 (and mode-2) vector space of $\mathcal{W}_r$. It can be obtained by truncated Multilinear Singular Value Decomposition, or, more accurately, by best multilinear rank-(L, L, L^2) approximation.

After computation of $\mathbf{W}$, $\mathbf{B}$ follows from (5). Matrices $\mathbf{C}$ and $\mathbf{A}$ follow from (4). We compute $\mathbf{E} \cdot \mathbf{W} = [\mathbf{F}_1 \ldots \mathbf{F}_R]$, where $\mathbf{F}_r \in \mathbb{C}^{IJ \times L}$, $1 \leqslant r \leqslant R$. Stack $\mathbf{F}_r$ in $\tilde{\mathbf{F}}_r \in \mathbb{C}^{LI \times J}$. Then the best rank-1 approximation of $\tilde{\mathbf{F}}_r$ yields an estimate of $\mathbf{c}_r$ and $\mathrm{vec}(\mathbf{A}_r)$, $1 \leqslant r \leqslant R$.

Note that one can detect the value of L (and R) by clustering. The first L^2 columns of $\mathbf{Y}$ correspond to $(LR \times LR)$ matrices of which the column (and row) space is a subspace of the same L-dimensional vector space, namely, the column space of $\mathbf{W}_1$. The next L^2 columns of $\mathbf{Y}$ correspond to $(LR \times LR)$ matrices of which the column (and row) space lies in the L-dimensional column space of $\mathbf{W}_2$, and so on.

These results can be generalized for the decomposition of $\mathcal{T}$ in rank-$(L_r, L_r, 1)$ terms.

References

[1] L. De Lathauwer, "A Link between the Canonical Decomposition in Multilinear Algebra and Simultaneous Matrix Diagonalization", *SIAM J. Matrix Anal. Appl.*, Vol. 28, No. 3, 2006, pp. 642–666.

[2] L. De Lathauwer, "Decompositions of a Higher-Order Tensor in Block Terms — Part II: Definitions and Uniqueness", *SIAM J. Matrix Anal. Appl.*, Special Issue *Tensor Decompositions and Applications*, Vol. 30, No. 3, 2008, pp. 1033–1066.

[3] A.-J. van der Veen and A. Paulraj, *An analytical constant modulus algorithm*, IEEE Trans. Signal Processing, 44 (1996), pp. 1136–1155.