Characterizing Higher-Order Tensors by Means of Subspaces

Tech. Report No. 11-32, SCD-SISTA, Dept. Electrical Engineering (ESAT), K.U.Leuven, 2011

Lieven De Lathauwer

February 21, 2011

Lemma 1. Consider $\mathbf{A} \in \mathbb{K}^{N \times N}$ with EVD $\mathbf{A} = \mathbf{E} \cdot \mathbf{D}_A \cdot \mathbf{E}^{-1}$. Assume that $\mathbf{B} \in \mathbb{K}^{N \times N}$ commutes with $\mathbf{A}$. Then $\mathbf{B}$ can be decomposed as $\mathbf{B} = \mathbf{E} \cdot \mathbf{D}_B \cdot \mathbf{E}^{-1}$, in which $\mathbf{D}_B \in \mathbb{K}^{N \times N}$ is block-diagonal, where the blocks correspond to the different eigenvalues of $\mathbf{A}$. The size of a block is equal to the algebraic multiplicity of the corresponding eigenvalue of $\mathbf{A}$. Assume that the diagonal block of $\mathbf{D}_A$ corresponding to eigenvalue λ is given by $\text{blockdiag}(\lambda \mathbf{I}_P, \mathbf{J}(\lambda, L_1), \ldots, \mathbf{J}(\lambda, L_Q))$, in which $\mathbf{J}(\lambda, L_q)$ denotes a Jordan cell of size $(L_q \times L_q)$ associated with λ. Then the corresponding diagonal block of $\mathbf{D}_B$ has the structure $\text{blockdiag}(\mathbf{M}_P, \tilde{\mathbf{J}}(\mu_1, L_1), \ldots, \tilde{\mathbf{J}}(\mu_Q, L_Q))$, in which $\mathbf{M}_P \in \mathbb{K}^{P \times P}$ and in which $\tilde{\mathbf{J}}(\mu_q, L_q)$ is a Jordan cell of size $(L_q \times L_q)$ associated with μ_q or a block diagonal matrix of size $(L_q \times L_q)$ consisting of several Jordan cells, each associated with μ_q.

Consider two matrices $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{N \times N}$. Assume that $\mathbf{A}$ has full column and row rank. Knowledge that $\mathbf{B}$ has the same column space and row space as $\mathbf{A}$ obviously does not give us much information about the structure of $\mathbf{B}$, besides the fact that it has full rank. However, the situation turns out to be very different for higher-order tensors. The "row space", "column space" and "mode-3 space" of a third-order tensor tell a lot about its structure.

1. We consider $\mathcal{A}, \mathcal{B} \in \mathbb{K}^{N \times N \times N}$. Tensor $\mathcal{A}$ has full multilinear rank. We know that

$$\text{span}(\mathbf{A}_{[2,3;1]}) = \text{span}(\mathbf{B}_{[2,3;1]}) \quad \text{and} \quad \text{span}(\mathbf{A}_{[1,3;2]}) = \text{span}(\mathbf{B}_{[1,3;2]}). \tag{1}$$

We examine how $\mathcal{A}$ and $\mathcal{B}$ are related.

From $\text{span}(\mathbf{A}_{[2,3;1]}) = \text{span}(\mathbf{B}_{[2,3;1]})$ we know that $\mathcal{B} = \mathcal{A} \cdot_1 \mathbf{M}_1$ for some nonsingular matrix $\mathbf{M}_1 \in \mathbb{K}^{N \times N}$. From $\text{span}(\mathbf{A}_{[1,3;2]}) = \text{span}(\mathbf{B}_{[1,3;2]})$ we know that $\mathcal{B} = \mathcal{A} \cdot_2 \mathbf{M}_2$ for some nonsingular matrix $\mathbf{M}_2 \in \mathbb{K}^{N \times N}$. Hence, we have that $\mathcal{A} \cdot_1 \mathbf{M}_1 = \mathcal{A} \cdot_2 \mathbf{M}_2$ for some nonsingular matrices $\mathbf{M}_1$ and $\mathbf{M}_2$. This is equivalent with

$$\mathbf{M}_1 \cdot \mathbf{A}_{:,:,n} = \mathbf{A}_{:,:,n} \cdot \mathbf{M}_2^T \quad \forall n. \tag{2}$$

Since this condition holds for all slices, it also holds for a generic linear combination $\mathbf{A}_{c_1}$ of them, which by construction has full rank:

$$\mathbf{M}_1 = \mathbf{A}_{c_1} \cdot \mathbf{M}_2^T \cdot \mathbf{A}_{c_1}^{-1}. \tag{3}$$

Substitution of (3) in (2) yields:

$$\mathbf{M}_2^T \cdot \mathbf{A}_{c_1}^{-1} \cdot \mathbf{A}_{:,:,n} = \mathbf{A}_{c_1}^{-1} \cdot \mathbf{A}_{:,:,n} \cdot \mathbf{M}_2^T \quad \forall n. \tag{4}$$

In other words, $\mathbf{M}_2^T$ commutes with $\mathbf{A}_{c_1}^{-1} \cdot \mathbf{A}_{:,:,n}$, $1 \leqslant n \leqslant N$, and with all linear combinations of these matrices. Hence, $\mathbf{M}_2^T$ commutes with $\mathbf{A}_{c_1}^{-1} \cdot \mathbf{A}_{c_2}$, in which $\mathbf{A}_{c_2}$ is a second generic linear combination of the slices $\mathbf{A}_{:,:,n}$. If the EVD of $\mathbf{A}_{c_1}^{-1} \cdot \mathbf{A}_{c_2}$ is given by $\mathbf{E} \cdot \mathbf{S} \cdot \mathbf{E}^{-1}$, then we have that $\mathbf{M}_2^T = \mathbf{E} \cdot \tilde{\mathbf{S}} \cdot \mathbf{E}^{-1}$, in which $\tilde{\mathbf{S}}$ is structured as explained in Lemma 1. On the other hand, since $\mathbf{A}_{c_1}^{-1} \cdot \mathbf{A}_{:,:,n}$ commutes with $\mathbf{M}_2^T$, it has the form

$$\mathbf{A}_{c_1}^{-1} \cdot \mathbf{A}_{:,:,n} = \mathbf{E} \cdot \mathbf{D}_n \cdot \mathbf{E}^{-1}, \tag{5}$$

in which $\mathbf{D}_n$ is structured as explained in Lemma 1, $1 \leqslant n \leqslant N$. From (5) we have that

$$\mathbf{A}_{:,:,n} = (\mathbf{A}_{c_1} \cdot \mathbf{E}) \cdot \mathbf{D}_n \cdot \mathbf{E}^{-1} \quad \forall n. \tag{6}$$

We distinguish different cases, corresponding to the different possibilities for the structure of $\tilde{\mathbf{S}}$:

(i) All eigenvalues of $\mathbf{M}_2$ are identical and $\mathbf{M}_2$ can be diagonalized, i.e., $\tilde{\mathbf{S}} = \alpha\,\mathbf{I}_N$. In this case, $\mathbf{M}_2 = \alpha\,\mathbf{I}_N$ and, consequently, $\mathcal{A}$ and $\mathcal{B}$ are scaled versions of each other. Note that also $\mathbf{M}_1 = \alpha\,\mathbf{I}_N$.

(ii) All eigenvalues of $\mathbf{M}_2$ are distinct. In this case, all matrices $\mathbf{D}_n$ are diagonal and (6) is the conventional slice-wise representation of a CP decomposition of $\mathcal{A}$. Let us write:

$$\mathcal{A} = \sum_{n=1}^{N} \mathbf{u}_n \otimes \mathbf{v}_n \otimes \mathbf{w}_n, \tag{7}$$

in which $\otimes$ denotes the tensor outer product and $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_N] = \mathbf{A}_{c_1} \cdot \mathbf{E}$, $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_N] = \mathbf{E}^{-T}$ and $\mathbf{W} = [\mathbf{w}_1 \cdots \mathbf{w}_N] = [\mathrm{diag}(\mathbf{D}_n) \cdots \mathrm{diag}(\mathbf{D}_n)]^T$. The latter three matrices have full rank, otherwise $\mathcal{A}$ would not have full multilinear rank. We have:

$$\mathcal{B} = \sum_{n=1}^{N} (\mathbf{M}_1 \mathbf{u}_n) \otimes \mathbf{v}_n \otimes \mathbf{w}_n = \sum_{n=1}^{N} \mathbf{u}_n \otimes (\mathbf{M}_2 \mathbf{v}_n) \otimes \mathbf{w}_n. \tag{8}$$

Since sufficient conditions for CP essential uniqueness are satisfied [5], (8) implies that $\mathbf{M}_1 = \mathbf{U} \cdot \mathbf{\Lambda}_1 \cdot \mathbf{U}^{-1}$ and $\mathbf{M}_2 = \mathbf{V} \cdot \mathbf{\Lambda}_2 \cdot \mathbf{V}^{-1}$, in which $\mathbf{\Lambda}_1, \mathbf{\Lambda}_2 \in \mathbb{K}^{N \times N}$ are nonsingular diagonal. In other words, $\mathcal{A}$ and $\mathcal{B}$ consist of the same rank-1 terms, possibly differently scaled.

(iii) $\mathbf{M}_2$ can be diagonalized and has eigenvalues α_r with multiplicity L_r, $1 \leqslant r \leqslant R$:

$$\mathbf{M}_2 = \mathbf{V} \cdot \mathrm{blockdiag}(\alpha_1 \mathbf{I}_{L_1}, \ldots, \alpha_R \mathbf{I}_{L_R})) \cdot \mathbf{V}^{-1}, \tag{9}$$

in which $\mathbf{V} = \mathbf{E}^{-T}$. In this case, the matrices $\mathbf{D}_n$ are block-diagonal with blocks of size $(L_r \times L_r)$, $1 \leqslant r \leqslant R$. Eq. (6) is the slice-wise representation of a decomposition of $\mathcal{A}$ in rank-$(L_r, L_r, \cdot)$ terms [2, 3, 4]. We have:

$$\mathbf{B}_{:,:,n} = (\mathbf{A}_{c_1} \cdot \mathbf{E}) \cdot \mathbf{D}_n \cdot (\mathbf{E}^{-1} \cdot \mathbf{M}_2^T) \quad \forall n. \tag{10}$$

Substitution of the EVD of $\mathbf{M}_2^T$ yields:

$$\mathbf{B}_{:,:,n} = (\mathbf{A}_{c_1} \cdot \mathbf{E}) \cdot (\mathbf{D}_n \cdot \mathrm{blockdiag}(\alpha_1 \mathbf{I}_{L_1}, \ldots, \alpha_R \mathbf{I}_{L_R})) \cdot \mathbf{E}^{-1} \quad \forall n. \tag{11}$$

In other words, $\mathcal{A}$ and $\mathcal{B}$ consist of the same rank-$(L_r, L_r, \cdot)$ terms, possibly differently scaled. Note that

$$\mathbf{M}_1 = \mathbf{U} \cdot \mathrm{blockdiag}(\alpha_1 \mathbf{I}_{L_1}, \ldots, \alpha_R \mathbf{I}_{L_R})) \cdot \mathbf{U}^{-1}, \tag{12}$$

in which $\mathbf{U} = \mathbf{A}_{c_1} \cdot \mathbf{E}$.

(iv) $\mathbf{M}_2$ has eigenvalues α_r with multiplicity L_r, $1 \leqslant r \leqslant R$. For α_2, ..., α_R the geometric multiplicity equals the algebraic multiplicity. The geometric multiplicity of α_1 is equal to 1. In this case, $\tilde{\mathbf{S}}$ is diagonal, up to the upper left diagonal block, which is a Jordan cell of size $(L_1 \times L_1)$. We have:

$$\mathbf{M}_2^T = \mathbf{E} \cdot \text{blockdiag}(\mathbf{J}(\alpha_1, L_1), \ldots, \alpha_R \mathbf{I}_{L_R})) \cdot \mathbf{E}^{-1}. \tag{13}$$

The matrices $\mathbf{D}_n$ are block-diagonal with blocks of size $(L_r \times L_r)$, $1 \leqslant r \leqslant R$. The first $(L_1 \times L_1)$ block is a Jordan cell or a block diagonal matrix consisting of several smaller Jordan cells, all associated with the same eigenvalue. The possible splitting in smaller Jordan cells and the eigenvalue can be different for different n. We can again interpret (6) as the slice-wise representation of a decomposition of $\mathcal{A}$ in rank-$(L_r, L_r, \cdot)$ terms, where the components related to α_1 are grouped in the first term. We have:

$$\mathbf{B}_{:,:,n} = (\mathbf{A}_{c_1} \cdot \mathbf{E}) \cdot \mathbf{D}_n \cdot (\mathbf{E}^{-1} \cdot \mathbf{M}_2^T) \quad \forall n. \tag{14}$$

Substitution of the EVD of $\mathbf{M}_2^T$ yields:

$$\mathbf{B}_{:,:,n} = (\mathbf{A}_{c_1} \cdot \mathbf{E}) \cdot (\mathbf{D}_n \cdot \text{blockdiag}(\mathbf{J}(\alpha_1, L_1), \ldots, \alpha_R \mathbf{I}_{L_R})) \cdot \mathbf{E}^{-1} \quad \forall n. \tag{15}$$

The first block term in the expansion of $\mathcal{A}$ and $\mathcal{B}$ is the same up to multiplication of the core by a Jordan cell. The other block terms are the same up to scaling. Note that:

$$\mathbf{M}_1 = \mathbf{U} \cdot \text{blockdiag}(\mathbf{J}(\alpha_1, L_1), \ldots, \alpha_R \mathbf{I}_{L_R})) \cdot \mathbf{U}^{-1}, \tag{16}$$

in which $\mathbf{U} = \mathbf{A}_{c_1} \cdot \mathbf{E}$.

(v) ...

The conclusion is the following. Apart from the specific situation with Jordan cells, $\mathcal{A}$ and $\mathcal{B}$ are possibly different linear combinations of the same rank-$(L_r, L_r, \cdot)$ terms, where the size of the blocks depends on the multiplicities of the eigenvalues of $\mathbf{M}_2$. (It is also possible to start from $\mathbf{M}_1$.) In the case $L_1 = L_2 = \ldots = L_N = 1$, $\mathcal{A}$ and $\mathcal{B}$ are linear combinations of the same rank-1 terms. In the case $L_1 = N$, $\mathcal{A}$ and $\mathcal{B}$ are equal up to scaling.

For $N \geqslant 3$, all decompositions in two or more rank-$(L_r, L_r, \cdot)$ terms are subgeneric, as can easily be verified by counting degrees of freedom. This means that, for a generic tensor $\mathcal{A} \in \mathbb{K}^{N \times N \times N}$, with $N \geqslant 3$, the conditions (1) imply that $\mathcal{A}$ and $\mathcal{B}$ are the same up to scaling. Over $\mathbb{C}$, the generic rank of $(2 \times 2 \times 2)$-tensors is 2. Over $\mathbb{R}$, the Lebesgue measure of both the set of rank-2 $(2 \times 2 \times 2)$-tensors and the set of rank-3 $(2 \times 2 \times 2)$-tensors is different from zero. See [1] and references therein. Hence, over $\mathbb{C}$, conditions (1) for $N = 2$ generically imply that $\mathcal{A}$ and $\mathcal{B}$ are possibly different linear combinations of the same two rank-1 terms. Over $\mathbb{R}$, the same holds true for $(2 \times 2 \times 2)$-tensors on the rank-2 manifold. For tensors on the rank-3 manifold conditions (1) for $N = 2$ imply equality up to scaling.

2. We consider $\mathcal{A}, \mathcal{B} \in \mathbb{K}^{N \times N \times N}$. Tensor $\mathcal{A}$ has full multilinear rank. We know that

$$\begin{aligned} \operatorname{span}(\mathbf{A}_{[2,3;1]}) = \operatorname{span}(\mathbf{B}_{[2,3;1]}), \quad \operatorname{span}(\mathbf{A}_{[1,3;2]}) = \operatorname{span}(\mathbf{B}_{[1,3;2]}) \\ \text{and} \quad \operatorname{span}(\mathbf{A}_{[1,2;3]}) = \operatorname{span}(\mathbf{B}_{[1,2;3]}). \end{aligned} \tag{17}$$

We examine how $\mathcal{A}$ and $\mathcal{B}$ are related.

The first two conditions yield the relations derived under *Problem 1*. The last two conditions yield similar relations, this time involving modes 2 and 3. Combining the relations, we obtain the following variants of the cases discussed under *Problem 1*:

(i) $\mathcal{A}$ and $\mathcal{B}$ are scaled versions of each other.

(ii) $\mathcal{A}$ and $\mathcal{B}$ consist of the same rank-1 terms, possibly differently scaled.

(iii) $\mathcal{A}$ and $\mathcal{B}$ consist of the same rank-(L_r, L_r, L_r) terms, possibly differently scaled.

(iv) $\mathcal{A}$ and $\mathcal{B}$ admit decompositions in rank-(L_r, L_r, L_r) terms that are related. The first block term, related to α_1, is the same up to multiplication of the core by a Jordan cell. It can be verified that the only possibility to satisfy (17) is a core $\mathcal{C}$ of the form

$$\mathbf{C}_{[1,2;3]} = \left(\begin{array}{cc|cc} \lambda & 0 & 0 & \lambda \\ 0 & \lambda & 0 & 0 \end{array} \right), \tag{18}$$

with $\lambda \neq 0$ and $L_1 = 2$. The other block terms are the same up to scaling.

(v) ...

Acknowledgements

Research supported by: (1) Research Council K.U.Leuven: GOA-MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC), CIF1 and STRT1/08/023 (2) F.W.O.: (a) project G.0427.10N, (b) Research Communities ICCoS, ANMMM and MLDM, (3) the Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, "Dynamical systems, control and optimization", 2007–2011), (4) EU: ERNSI.

References

[1] P. Comon, J.M.F. ten Berge, L. De Lathauwer, J. Castaing, "Generic and Typical Ranks of Multi-Way Arrays", *Lin. Alg. Appl.*, Vol. 430, No. 11–12, June 2009, pp. 2997–3007

[2] L. De Lathauwer, "Decompositions of a Higher-Order Tensor in Block Terms — Part I: Lemmas for Partitioned Matrices", *SIAM J. Matrix Anal. Appl.*, Special Issue *Tensor Decompositions and Applications*, Vol. 30, No. 3, 2008, pp. 1022–1032.

[3] 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.

[4] L. De Lathauwer, D. Nion, "Decompositions of a Higher-Order Tensor in Block Terms — Part III: Alternating Least Squares Algorithms", *SIAM J. Matrix Anal. Appl.*, Special Issue *Tensor Decompositions and Applications*, Vol. 30, No. 3, 2008, pp. 1067–1083.

[5] R.A. Harshman, *Foundations of the PARAFAC procedure*: *Model and conditions for an "explanatory" multi-mode factor analysis*, UCLA Working Papers in Phonetics, 16 (1970), pp. 1–84.