Clarification on Tucker decomposition

Miaoyan Wang, Dec 21, 2019

Proposition 1 (Clustering Issue)**.** *Let $\mathcal{T} \in \mathbb{R}^{d_1 \times \cdots \times d_K}$ be an order-K $(d_1, \ldots, d_K)$-dimensional tensor. Suppose $\mathcal{T}$ admits a rank-$(r_1, \ldots, r_K)$ Tucker decomposition:*

$$\mathcal{T} = \mathcal{C} \times_1 \boldsymbol{M}_1 \times_2 \cdots \times_K \boldsymbol{M}_K.$$

Then, tensor slides are equal if and only if the corresponding rows in factor matrices are equal; i.e.,

$$\mathcal{T}(i, \ldots) = \mathcal{T}(j, \ldots) \Longleftrightarrow \boldsymbol{M}_1(i, :) = \boldsymbol{M}_1(j, :).$$

Proof. We prove the necessity, i.e. "equal tensor slices $\Rightarrow$ equal factor rows".

Suppose $\mathcal{T}(i, \ldots) = \mathcal{T}(j, \ldots)$ for some $i \neq j$ and $i, j \in [d_1]$. Let $\text{Unfold}_1(\mathcal{T}) \in \mathbb{R}^{d_1 \times d_{-1}}$ denote the mode-1 unfolding of $\mathcal{T}$, where $d_{-1} \stackrel{\text{def}}{=} \prod_{j \neq 1} d_j$. Then, the tensor slices being equal implies the rows being equal in the unfolded matrix $\text{Unfold}_1(\mathcal{T})$, i.e.,

$$\text{Unfold}_1(\mathcal{T})(i, \ldots) = \text{Unfold}_1(\mathcal{T})(j, \ldots). \tag{1}$$

Based on the Tucker decomposition of $\mathcal{T}$, we have

$$\text{Unfold}_1(\mathcal{T})(i, \ldots) = \boldsymbol{M}_1(i, :)\boldsymbol{B}, \quad \text{and} \quad \text{Unfold}_1(\mathcal{T})(j, \ldots) = \boldsymbol{M}_1(j, :)\boldsymbol{B}, \tag{2}$$

where $\boldsymbol{B} = \text{Unfold}_1(\mathcal{C} \times_2 \boldsymbol{M}_2 \times \cdots \boldsymbol{M}_K) \in \mathbb{R}^{r_1 \times d_{-1}}$, and $\boldsymbol{M}_1(i, :)$ (respectively, $\boldsymbol{M}_1(j, :)$) denotes the i-th (respectively, j-th) row in the factor matrix $\boldsymbol{M}_1$. Combining (1) and (2) yileds

$$\begin{aligned} & \boldsymbol{M}_1(i, :)\boldsymbol{B} = \boldsymbol{M}_1(j, :)\boldsymbol{B} \\ \Longleftrightarrow\ & \boldsymbol{M}_1(i, :)\boldsymbol{B}\boldsymbol{B}^T = \boldsymbol{M}_1(j, :)\boldsymbol{B}\boldsymbol{B}^T \\ \Longleftrightarrow\ & \boldsymbol{M}_1(i, :) = \boldsymbol{M}_1(j, :) \quad \text{provided that } \boldsymbol{B} \text{ has full row rank.} \end{aligned}$$

$\square$

Proposition 2 (Symmetry Issue)**.** *Let $\mathcal{T} \in \mathbb{R}^{d_1 \times \cdots \times d_K}$ be an order-K $(d_1, \ldots, d_K)$-dimensional tensor. Suppose that $\mathcal{T}$ is symmetric with respect to the first two modes; that is, $d_1 = d_2$, and*

$$\mathcal{T}(:, :, i_3, \ldots, i_K) = [\mathcal{T}(:, :, i_3, \ldots, i_K)]^T, \quad \textit{for all } (i_3, \ldots, i_K) \in [d_3] \times \cdots \times [d_K].$$

Let $\mathcal{T} = \mathcal{C} \times_1 \boldsymbol{M}_1 \times_2 \cdots \times_K \boldsymbol{M}_K$ denote the Tucker decomposition of $\mathcal{T}$. Then, the factor matrices $\boldsymbol{M}_1, \boldsymbol{M}_2$ satisfy

$$\boldsymbol{M}_1\boldsymbol{M}_1^T = \boldsymbol{M}_2\boldsymbol{M}_2^T.$$

In general, the symmetry between modes does not necessarily imply $\boldsymbol{M}_1 = \boldsymbol{M}_2$.

Proof. It is easy to check that $\text{Unfold}_1(\mathcal{T})$ and $\text{Unfold}_2(\mathcal{T})$ span the same column spaces. Therefore, $\boldsymbol{M}_1\boldsymbol{M}_1^T = \boldsymbol{M}_2\boldsymbol{M}_2^T$. To show the possibility of $\boldsymbol{M}_1 \neq \boldsymbol{M}_2$, we construct an order-2 tensor (i.e., matrix) $\mathcal{T}$ with the following Tucker decomposition:

$$\mathcal{T} = \underbrace{\begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 \end{bmatrix}}_{\boldsymbol{M}_1} \underbrace{\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}}_{\mathcal{C}} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}}_{\boldsymbol{M}_2^T} = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Clearly, $\boldsymbol{M}_1 \neq \boldsymbol{M}_2$ while $\mathcal{T}$ is symmetric. $\square$

Remark 1 (Weighting in the Clustering)**.** Let $\mathcal{T} = \mathcal{C} \times_1 \boldsymbol{M}_1 \times_2 \cdots \times_K \boldsymbol{M}_K$ denote the Tucker decomposition of $\mathcal{T}$. Consider the task of clustering tensor slides along the mode-1. Let $\boldsymbol{M}_1 \in \mathbb{R}^{d_1 \times r_1}$ denote the mode-1 factor matrix and $\mathcal{C} \in \mathbb{R}^{r_1 \times \cdots \times r_K}$ the core tensor. One possibility is to apply K-means to weighted rows of $\boldsymbol{M}_1$. Specifically, let $\boldsymbol{M}_1 = [\![m_{ij}]\!]$ and

$$\text{Unfold}_1(\mathcal{C}) = \begin{bmatrix} — & \boldsymbol{a}_1^T & — \\ — & \boldsymbol{a}_2^T & — \\ & \vdots & \\ — & \boldsymbol{a}_{r_1}^T & — \end{bmatrix},$$

where $\boldsymbol{a}_i^T$ is a length-$(r_2 \dots r_K)$ vector for $i = 1, \dots, r_1$. We construct a new matrix $\boldsymbol{M}_1^*$ whose columns are weighted by (multipliers of) $\|\boldsymbol{a}_i\|_2$:

$$\boldsymbol{M}_1^* = \begin{bmatrix} \|\boldsymbol{a}_1\|_2^{1/K} m_{11} & \|\boldsymbol{a}_2\|_2^{1/K} m_{12} & \cdots & \|\boldsymbol{a}_r\|_2^{1/K} m_{1r} \\ \|\boldsymbol{a}_1\|_2^{1/K} m_{21} & \|\boldsymbol{a}_2\|_2^{1/K} m_{22} & \cdots & \|\boldsymbol{a}_r\|_2^{1/K} m_{2r} \\ \vdots & \vdots & \vdots & \vdots \\ \|\boldsymbol{a}_1\|_2^{1/K} m_{d1} & \|\boldsymbol{a}_2\|_2^{1/K} m_{d2} & \cdots & \|\boldsymbol{a}_r\|_2^{1/K} m_{dr} \end{bmatrix},$$

where K is the order of the tensor and $\|\boldsymbol{a}_i\|_2$ is the vector norm. The reason for choosing $1/K$ as the exponent is to "evenly distribute the energy" of core tensor over K factors. In the special case when $K = 2$, our scheme is equivalent to weighting the singular vectors by squared roots of the singular values. Any multivariate clustering algorithms (e.g., K-means) can then be applied to the matrix $\boldsymbol{M}_1^*$.