A THIRD-ORDER GENERALIZATION OF THE MATRIX SVD

Paper Report
CHAU Dang Minh

Summary. We have studied several SVD generalizations for tensors (CP and the Tuckers), which express a tensor as a sum of rank-one tensors. In the approach in [2], the authors defined a ring of tensors under the original addition and the to-be-defined product which can be reduced to that of invertible matrices. This allows a generalization for third-order tensors given as a product of three tensors.

1 Preliminaries

In this section, we provide necessary operations for the definition of this type of SVD. Firstly, recall the notation of a p-order tensor

$$\mathcal{A} = (a_{i_1 \dots i_p}) \in \mathbb{R}^{n_1 \times \dots \times n_p}. \tag{1}$$

For a third-order tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the *unfold map* is a generalization of matricization, where

- The first parameter is the tensor.
- The optional second parameter is the dimension, defaulted to be 1.
- The optional third parameter taking an array value, specifying the order of fibers to be organized.

For example, let

$$\mathcal{A} = \left[\begin{array}{cc|cc} a_{111} & a_{121} & a_{112} & a_{122} \\ a_{211} & a_{221} & a_{212} & a_{222} \end{array}\right] = \left[\begin{array}{cc|cc} 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \end{array}\right].$$

We have

$$\operatorname{unfold}(\mathcal{A}) = \operatorname{unfold}(\mathcal{A}, 1) = \begin{bmatrix} \mathcal{A}_{:,:,1} \\ \mathcal{A}_{:,:,2} \end{bmatrix} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \\ 5 & 7 \\ 6 & 8 \end{pmatrix},$$

$$\operatorname{unfold}(\mathcal{A}, 2) = \begin{bmatrix} \mathcal{A}_{:,1,:} \\ \mathcal{A}_{:,2,:} \end{bmatrix} = \begin{pmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{pmatrix}, \operatorname{unfold}(\mathcal{A}, 2, [2,1]) = \begin{bmatrix} \mathcal{A}_{:,2,:} \\ \mathcal{A}_{:,1,:} \end{bmatrix} = \begin{pmatrix} 3 & 7 \\ 4 & 8 \\ 1 & 5 \\ 2 & 6 \end{pmatrix}.$$

Inversely, we have the *fold map*, which transforms a matrix to a tensor. The parameters of the unfold map are the same as those of the unfold map. The *circular map* is defined for each matrix $A = \begin{bmatrix} A_1 & \cdots & A_{n_3} \end{bmatrix} \in \mathbb{R}^{n_1 n_3 \times n_2}$ as

$$\operatorname{circ}(A) = \begin{bmatrix} A_1 & A_{n_3} & \cdots & A_2 \\ A_2 & A_1 & \cdots & A_3 \\ \vdots & \vdots & \ddots & \vdots \\ A_{n_3} & A_{n_3-1} & \cdots & A_1 \end{bmatrix}. \tag{2}$$

The $n \times n$ discrete Fourier transform matrix $F_n = (f_{jk})$ is defined [1] by

$$f_{jk} = w_n^{(j-1)(k-1)}, \tag{3}$$

where

$$w_n = \exp\left(\frac{-2\pi i}{n}\right) = \cos\left(\frac{2\pi i}{n}\right) - i \sin\left(\frac{2\pi i}{n}\right).$$

For a vector $x \in \mathbb{C}^n$, $\mathrm{dft}(x) = F_n x$ is called the discrete Fourier transform (DFT) of x.

For example, we have

$$F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & w_4^1 & w_4^2 & w_4^3 \\ 1 & w_4^2 & w_4^4 & w_4^6 \\ 1 & w_4^3 & w_4^6 & w_4^9 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix}.$$

Proposition 1. *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and F_{n_3} is the $n_3 \times n_3$ matrix, we have*

$$(F_{n_3} \otimes I_{n_1}) \cdot \mathrm{circ}(\mathrm{unfold}(\mathcal{A})) \cdot (F_{n_3}^* \otimes I_{n_2}) = \begin{pmatrix} D_1 & & \\ & \ddots & \\ & & D_{n_3} \end{pmatrix}, \tag{4}$$

where D_ℓ are the faces of the tensor $\mathcal{D}$ computed by applying DFT along each fiber $A_{i,j,:}$ of $\mathcal{A}$.

2 The Product Operation

Having developed relevant operations which imply Proposition 1, we now come to our main product operation.

Definition 1. *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\mathcal{B} \in \mathbb{R}^{n_2 \times \ell \times n_3}$. The product $\mathcal{A} * B$ is the $n_1 \times \ell \times n_3$ tensor*

$$\mathcal{A} * \mathcal{B} = \mathrm{fold}(\mathrm{circ}(\mathrm{unfold}(\mathcal{A})) \cdot \mathrm{unfold}(\mathcal{B})). \tag{5}$$

Next, we develop the notions of inverse and transpose.

Definition 2. *Let $\mathcal{A}$ be a three-order tensor.*

1. *The identity tensor $\mathcal{I} \in \mathbb{R}^{n \times n \times \ell}$ is the tensor whose the front face $\mathcal{I}_{:,:,1}$ is the identity matrix and other faces are all zeros.*
2. *The transpose $\mathcal{A}^\top$ of $\mathcal{A}$ is defined by*

$$\mathcal{A}^\top = \mathrm{fold}(\mathrm{unfold}(\mathcal{A}, 1, [1, n_3 : -1 : 2])). \tag{6}$$

3. *A tensor $\mathcal{B}$ is called the inverse of $\mathcal{A}$ if $\mathcal{A} * \mathcal{B} = \mathcal{B} * \mathcal{A} = \mathcal{I}$.*
4. *A tensor $\mathcal{Q} \in \mathbb{R}^{n \times n \times \ell}$ is orthogonal if $\mathcal{Q} * \mathcal{Q}^\top = \mathcal{Q}^\top * \mathcal{Q} = \mathcal{I}$.*

Certain properties of the product as those of matrices are proven.

Proposition 2. *Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be tensors of appropriate sizes. We have*

1. *(Associativity) $(\mathcal{A} * B) * C = \mathcal{A} * (B * C)$.*
2. *$(\mathcal{A} * \mathcal{B})^\top = \mathcal{B}^\top * A^\top$.*
3. *If $\mathcal{Q}$ is orthogonal, then $\|\mathcal{Q} * \mathcal{A}\|_F = \|\mathcal{A}\|_F$.*

Proposition 3. *Let $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ have appropriate sizes whose the last dimension is n_3 satisfying $\mathcal{A} = \mathcal{B} * \mathcal{C}$. Then*

$$\sum_{\ell=1}^{n_3} \mathcal{A}_{:,:,\ell} = \left(\sum_{\ell=1}^{n_3} \mathcal{B}_{:,:,\ell}\right)\left(\sum_{\ell=1}^{n_3} \mathcal{C}_{:,:,\ell}\right). \tag{7}$$

3 The Generalized SVD

In this section, we give a proof on the existence of a generalized SVD, called the T-SVD, some applications on data compression and further discussion.

Theorem 1. *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. Then $\mathcal{A}$ can be factorized as*

$$\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^\top,$$

where $\mathcal{U} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$ and $\mathcal{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ are orthogonal, and $\mathcal{S} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ has diagonal frontal faces.

Proof. By Proposition 1, we have

$$(F_{n_3} \otimes I_{n_1}) \cdot \mathrm{circ}(\mathrm{unfold}(\mathcal{A})) \cdot (F_{n_3}^* \otimes I_{n_2}) = \mathrm{diag}(F_{n_3}\mathcal{A}_{:,:,1}, \ldots, F_{n_3}\mathcal{A}_{:,:,n_3}).$$

Compute the SVD of each D_j as $D_j = U_i \Sigma_i V_i^\top$, we have

$$\begin{pmatrix} D_1 & & \\ & \ddots & \\ & & D_{n_3} \end{pmatrix} = \begin{pmatrix} U_1 & & \\ & \ddots & \\ & & U_{n_3} \end{pmatrix} \begin{pmatrix} \Sigma_1 & & \\ & \ddots & \\ & & \Sigma_{n_3} \end{pmatrix} \begin{pmatrix} V_1^\top & & \\ & \ddots & \\ & & V_{n_3}^\top \end{pmatrix}.$$

Therefore,

$$\begin{aligned} \mathrm{circ}(\mathrm{unfold}(\mathcal{A})) = (F_{n_3}^* \otimes I_{n_1}) &\begin{pmatrix} U_1 & & \\ & \ddots & \\ & & U_{n_3} \end{pmatrix} (F_{n_3} \otimes I_{n_1}) \\ (F_{n_3}^* \otimes I_{n_1}) &\begin{pmatrix} \Sigma_1 & & \\ & \ddots & \\ & & \Sigma_{n_3} \end{pmatrix} (F_{n_3} \otimes I_{n_2}) \\ (F_{n_3}^* \otimes I_{n_2}) &\begin{pmatrix} V_1^\top & & \\ & \ddots & \\ & & V_{n_3}^\top \end{pmatrix} (F_{n_3} \otimes I_{n_2}). \end{aligned}$$

The matrices in each row are all circular. Take the first column of the right-hand size and fold up the result using inverse DFT, we have $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^\top$, where $\mathcal{U}$ has $U_i, i = 1, \ldots, n_3$ to be frontal slides. Similarly for $\mathcal{S}$ and $\mathcal{V}$. By simple computations, we can show that the tensors are orthogonal. □

The proof is a construction of the SVD, which can be written as an algorithm.

Algorithm 1 T-SVD

Input: $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$.

Compute $\mathcal{D}$ whose fibers are DFT of fibers of $\mathcal{A}$.
for $n = 1, \ldots, n_3$ **do**
  $[\mathcal{U}_{:,:,n}, \mathcal{S}_{:,:,n}, \mathcal{V}_{:,:,n}] = \mathrm{SVD}(\mathcal{D}_{:,:,n})$.
end for
Apply inverse DFT to fibers of $\mathcal{U}, \mathcal{S}$ and $\mathcal{V}$.

The T-SVD is used in data compression. Note that

$$\mathcal{A} = \sum_{j=1}^{\min\{n_1,n_2\}} \mathcal{U}_{:,i,:} * \mathcal{S}_{i,i,:}\mathcal{V}_{:,i,:}, \tag{8}$$

we can approximate

$$\mathcal{A} \approx \sum_{j=1}^{k} \mathcal{U}_{:,i,:} * \mathcal{S}_{i,i,:} \mathcal{V}_{:,i,:} \tag{9}$$

for some $k < \min\{n_1, n_2\}$. Another strategy makes use of Proposition 3 and associativity. We have

$$\sum_{\ell=1}^{n_3} \mathcal{A}_{:,:,\ell} = \left(\sum_{\ell=1}^{n_3} \mathcal{U}_{:,:,\ell}\right) \left(\sum_{\ell=1}^{n_3} \mathcal{S}_{:,:,\ell}\right) \left(\sum_{\ell=1}^{n_3} \mathcal{V}^{\top}_{:,:,\ell}\right) := USV^{\top}.$$

For some $k_1 < n_1$ and $k_2 < n_2$, let $\tilde{U} = U_{:,1:k_1}$ and $\tilde{V} = V_{:,1:k_2}$. Let $\mathcal{T} \in \mathbb{R}^{k_1 \times k_2 \times n_3}$ such that

$$\mathcal{T}_{:,:,\ell} = \tilde{U}^{\top} A_{:,:,\ell} \tilde{V}.$$

We the approximate

$$\mathcal{A} \approx \sum_{j=1}^{k_1} \sum_{k=1}^{k_2} \tilde{U}_{:,j} \circ \tilde{V}_{:,k} \circ \mathcal{T}_{j,k,:}. \tag{10}$$

Using the T-SVD, we can further prove the existence of analogous QR decomposition and eigendecomposition for tensors.

References

[1] Gene H Golub and Charles F Van Loan. *Matrix computations*. JHU press, 2013.

[2] Misha E Kilmer, Carla D Martin, and Lisa Perrone. “A third-order generalization of the matrix svd as a product of third-order tensors”. In: *Tufts University, Department of Computer Science, Tech. Rep. TR-2008-4* (2008).