ON THE ADJOINT OF TENSORS

MOODY T. CHU*

Abstract. The notion of adjoint is studied when a tensor is regarded as a linear transformation between tensor spaces.

Key words. adjoint, linear transformation

AMS subject classifications. 15A15, 15A09, 15A23

1. Introduction. A tensor of order k can be represented by a k-way array

$$T = [\tau_{i_1,\ldots,i_k}] \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_k}$$

with elements $t_{i_1,\ldots,i_k}$ accessed via k indices. When $I_1 = \ldots = I_k$, we have a square tensor. An order-k square tensor T is said to be super-symmetric [1, 2, 3, 5] if

$$\tau_{i_1,\ldots,i_k} = \tau_{i_{\sigma(1)},\ldots,i_{\sigma(k)}} \tag{1.1}$$

with respect to all possible permutations σ over the integers $\{1, \ldots k\}$. Such a total immunity to permutations of indices in a super-symmetric tensor appears to be an ultimate generalization of the symmetric matrices, but is it merely a straightforward structural emulation or does it carry a similar meaning of a self-adjoint operator?

In linear algebra, the transpose $A^\top$ of a matrix A usually results from the adjoint of the linear transformation represented by A. Since a matrix involves only two indices, the adjoint operator can easily be obtained by merely swapping the order of the indices. What is meant by physically transposing a multi-dimensional array?

2. Tensor as a linear transformation. Suppose that the set $\{1, \ldots, k\}$ is partitioned as the union of two disjoint nonempty subsets $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_s\}$ and $\boldsymbol{\beta} = \{\beta_1, \ldots, \beta_t\}$, where $s + t = k$. For convenience, an element $\tau_{i1,\ldots,i_k}$ will be marked as $\tau^{(\boldsymbol{\alpha},\boldsymbol{\beta})}_{[\mathcal{I}|\mathcal{J}]}$ where $\mathcal{I} := (i_1, \ldots, i_s)$ and $\mathcal{J} := (j_1, \ldots, j_t)$ contain those indices at locations $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, respectively. Each index in the arrays $\mathcal{I}$ and $\mathcal{J}$ should be within the corresponding range of integers, e.g., $i_1 \in \{1, \ldots, I_{\alpha_1}\}$ and so on. If the reference to a specific partitioning $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is clear, without causing ambiguity, we abbreviate the element as $\tau_{[\mathcal{I}|\mathcal{J}]}$. For example, if $\boldsymbol{\alpha} = \{2, 4\}$ and $\boldsymbol{\beta} = \{1, 3, 5, 6\}$, then $\tau_{2,3,1,5,6,4} = \tau_{[3,5|2,1,6,4]}$. Because there is no preference of $\boldsymbol{\alpha}$ over $\boldsymbol{\beta}$, we should have $\tau^{(\boldsymbol{\alpha},\boldsymbol{\beta})}_{[\mathcal{I}|\mathcal{J}]} = \tau^{(\boldsymbol{\beta},\boldsymbol{\alpha})}_{[\mathcal{J}|\mathcal{I}]}$ for any fixed partitioning $\{1, \ldots, k\} = \boldsymbol{\alpha} \cup \boldsymbol{\beta}$. We stress that such a symmetry is with respect to the partitioning only and it does not indicate any symmetry with respect to the indices at all. Choosing various ways to partition $\{1, \ldots, k\}$ offers us a tool to slice into a high-dimensional T and exam its cross-sections from different perspectives in the following sense.

Given a fixed partitioning $\{1, \ldots, k\} = \boldsymbol{\alpha} \cup \boldsymbol{\beta}$, we shall regard an order-k tensor $T \in \mathbb{R}^{I_1 \times \ldots \times I_k}$ as a "matrix representation" of a linear operator mapping order-s tensors to order-t tensors [6]. Specifically, we identify T with the map

$$\mathscr{T}_{\boldsymbol{\beta}} : \mathbb{R}^{I_{\alpha_1} \times \ldots \times I_{\alpha_s}} \to \mathbb{R}^{I_{\beta_1} \times \ldots \times I_{\beta_t}}, \tag{2.1}$$

where, for any $A \in \mathbb{R}^{I_{\alpha_1} \times \ldots \times I_{\alpha_s}}$, we have

$$\mathscr{T}_{\boldsymbol{\beta}}(A) := T \circledast_{\boldsymbol{\beta}} A = \left[\langle \tau_{[:|\mathcal{J}]}, A \rangle\right] \in \mathbb{R}^{I_{\beta_1} \times \ldots \times I_{\beta_t}}. \tag{2.2}$$

In the above, $\tau_{[:|\mathcal{J}]} = \tau_{[:|j_1,\ldots,j_t]}$ denotes the $(j_1, \ldots, j_t)$-th "slice" in the $\boldsymbol{\beta}$ direction of the tensor T, that is, for $\ell = 1, \ldots t$, j_ℓ occurs at the β_ℓ-th location in the array $\{1, \ldots, k\}$ and assumes values $j_\ell = 1, \ldots I_{\beta_\ell}$, whereas

*Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205. (chu@math.ncsu.edu) This research was supported in part by the National Science Foundation under grant DMS-1316779.

the notation ":" represents a wild card at the $(\alpha_1, \ldots, \alpha_s)$ locationand

$$\langle \tau_{[:|\mathcal{J}]}, A \rangle := \sum_{i_1=1}^{I_{\alpha_1}} \cdots \sum_{i_s=1}^{I_{\alpha_s}} \tau_{[i_1,\ldots,i_s|\mathcal{J}]} a_{i_1,\ldots,i_s} \tag{2.3}$$

is the Frobenius inner product generalized to multi-dimensional arrays. Indeed, (2.3) can be abbreviated as

$$(\mathscr{T}_{\boldsymbol{\beta}}(A))_{\mathcal{J}} = \sum_{\mathcal{I}} \tau_{[\mathcal{I}|\mathcal{J}]} A_{\mathcal{I}}, \tag{2.4}$$

where the summation of $\mathcal{I}$ runs through appropriate ranges of the indices $i_1, \ldots, i_{\alpha_s}$. In terms of this multi-index notation, the operation (2.2) resembles the usual matrix-to-vector multiplication.

The adjoint of $\mathscr{T}_{\boldsymbol{\beta}}$ should be a linear transformation

$$\mathscr{T}^*_{\boldsymbol{\beta}} : \mathbb{R}^{I_{\beta_1} \times \ldots \times I_{\beta_t}} \to \mathbb{R}^{I_{\alpha_1} \times \ldots \times I_{\alpha_s}} \tag{2.5}$$

satisfying the Lagrange's identity [4]

$$\langle \mathscr{T}_{\boldsymbol{\beta}}(A), B \rangle = \left\langle A, \mathscr{T}^*_{\boldsymbol{\beta}}(B) \right\rangle \tag{2.6}$$

for all $A \in \mathbb{R}^{I_{\alpha_1} \times \ldots \times I_{\alpha_s}}$ and $B \in \mathbb{R}^{I_{\beta_1} \times \ldots \times I_{\beta_t}}$.

LEMMA 2.1. *Given a fixed partitioning $\{1, \ldots, k\} = \boldsymbol{\alpha} \cup \boldsymbol{\beta}$, it is true that*

$$\mathscr{T}^*_{\boldsymbol{\beta}} = \mathscr{T}_{\boldsymbol{\alpha}}. \tag{2.7}$$

Proof. Observe that

$$\langle \mathscr{T}_{\boldsymbol{\beta}}(A), B \rangle = \sum_{\mathcal{J}} \left(\sum_{\mathcal{I}} \tau_{[\mathcal{I}|\mathcal{J}]} a_{\mathcal{I}} \right) b_{\mathcal{J}} = \sum_{\mathcal{I}} \left(\sum_{\mathcal{J}} \tau_{[\mathcal{I}|\mathcal{J}]} b_{\mathcal{J}} \right) a_{\mathcal{I}}.$$

It follows that

$$(\mathscr{T}^*_{\boldsymbol{\beta}}(B))_{\mathcal{I}} = \sum_{\mathcal{J}} \tau_{[\mathcal{I}|\mathcal{J}]} b_{\mathcal{J}}. \tag{2.8}$$

Recall that $\tau_{[\mathcal{I}|\mathcal{J}]} = \tau^{(\boldsymbol{\alpha},\boldsymbol{\beta})}_{[\mathcal{I}|\mathcal{J}]} = \tau^{(\boldsymbol{\beta},\boldsymbol{\alpha})}_{[\mathcal{J}|\mathcal{I}]}$. By the definition of (2.2), the action in (2.8) is the same as that of $\mathscr{T}_{\boldsymbol{\alpha}}$. □

In the case $k = 2$ and $T \in \mathbb{R}^{m \times n}$, then $\mathscr{T}_1 : \mathbb{R}^n \to \mathbb{R}^m$ and its matrix representation is precisely T, while $\mathscr{T}_2 : \mathbb{R}^m \to \mathbb{R}^n$ and is represented by $T^\top$. It is clear that $\mathscr{T}_2 = \mathscr{T}_1^* = T^\top$. Note that the 2-D arrays T and $T^\top$ contain the same set of data. Their structural difference is only a matter of being perceived from different perspectives as rows and columns. Lemma 2.1 conveniently generalizes the notion of adjoint to tensors — The "transpose" of a tensor T, viewed from the $\boldsymbol{\beta}$ point of view, is the very same T viewed from the $\boldsymbol{\alpha}$ point of view, and vice versa. The notion of an adjoint is $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ dependent.

Example 1. It might be instructive to display these structures by an order-3 tensor, although such an attempt to visualize higher-order tensors is infeasible (and useless). Consider the partitioning with $\boldsymbol{\alpha} = \{2, 3\}$ and $\boldsymbol{\beta} = \{1\}$. The order-3 tensor T may be visualized as an $I_1 \times 1$ block matrix of which each block is an $I_2 \times I_3$ matrix, whereas the action $\circledast_1$ is a 2-dimensional contraction defined by

$$\mathscr{T}_1(X) \equiv \underbrace{\begin{bmatrix} \boxed{\tau_{1,:}} \\ \boxed{\tau_{2,:}} \\ \vdots \\ \boxed{\tau_{I_1,:}} \end{bmatrix}}_{T} \circledast_1 X = [\langle \tau_{i,:}, X \rangle] \in \mathbb{R}^{I_1},$$

for $X \in \mathbb{R}^{I_2 \times I_3}$. The adjoint $\mathscr{T}_1^* = \mathscr{T}_{2,3}$ can be visualized as an $I_2 \times I_3$ block matrix of which each block is an $I_1 \times 1$ column vector with action defined by

$$\mathscr{T}_{2,3}(X) \equiv \underbrace{\begin{bmatrix} \boxed{\tau_{:,11}} & \boxed{\tau_{:,12}} & \cdots & \boxed{\tau_{:,1I_3}} \\ \boxed{\tau_{:,21}} & & \cdots & \boxed{\tau_{:,2I_3}} \\ \vdots & & & \vdots \\ \boxed{\tau_{:,I_21}} & & & \boxed{\tau_{:,I_2I_3}} \end{bmatrix}}_{T^*} \circledast_{23} \mathbf{x} = [\langle \tau_{:,ij}, \mathbf{x} \rangle] \in \mathbb{R}^{I_2 \times I_3},$$

for $\mathbf{x} \in \mathbb{R}^{I_1}$. Both T and T^* contain the same set of data but are organized differently, depending on $(\boldsymbol{\alpha}, \boldsymbol{\beta})$.

To mimic the notion of symmetric matrices under the context of adjoint, we need

$$\mathscr{T}_{\boldsymbol{\alpha}} = \mathscr{T}_{\boldsymbol{\alpha}}^* \tag{2.9}$$

as operators. In turn, by Lemma 2.1, we need $\mathscr{T}_{\boldsymbol{\alpha}} = \mathscr{T}_{\boldsymbol{\beta}}$. It thus becomes necessary that $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ have the same cardinality, say, s, and $I_{\alpha_\ell} = I_{\beta_\ell}$, for $\ell = 1, \ldots, s$. The condition (2.9) also requires that

$$\langle \tau_{[:|\mathcal{J}]}^{(\boldsymbol{\alpha},\boldsymbol{\beta})}, A \rangle = \langle \tau_{[:|\mathcal{J}]}^{(\boldsymbol{\beta},\boldsymbol{\alpha})}, A \rangle \tag{2.10}$$

for all order-s tensors $A \in \mathbb{R}^{I_{\alpha_1} \times \ldots \times I_{\alpha_s}}$ and all multi-indices $\mathcal{J} \in I_{\alpha_1} \times \ldots \times I_{\alpha_s}$. Therefore, an order-$2s$ tensor T is "symmetric" with respect to the partitioning $(\boldsymbol{\beta}, \boldsymbol{\alpha})$ if and only if

$$\tau_{[\mathcal{I}|\mathcal{J}]}^{(\boldsymbol{\alpha},\boldsymbol{\beta})} = \tau_{[\mathcal{I}|\mathcal{J}]}^{(\boldsymbol{\beta},\boldsymbol{\alpha})} = \tau_{[\mathcal{J}|\mathcal{I}]}^{(\boldsymbol{\alpha},\boldsymbol{\beta})} \tag{2.11}$$

for all multi-indices $\mathcal{I}, \mathcal{J} \in I_{\alpha_1} \times \ldots \times I_{\alpha_s}$. We note that an odd order tensor can never be self-adjoint regardless how the partitioning is taken.

Example 2. Consider the order-4 tensor $T \in \mathbb{R}^{2\times3\times2\times3}$ represented index-wise by

$$T = \left[\begin{array}{|ccc|ccc|ccc|} \hline 1111 & 1112 & 1113 & 1211 & 1212 & 1213 & 1311 & 1312 & 1313 \\ 1121 & 1122 & 1123 & 1221 & 1222 & 1223 & 1321 & 1322 & 1323 \\ \hline 2111 & 2112 & 2113 & 2211 & 2212 & 2213 & 2311 & 2312 & 2313 \\ 2121 & 2122 & 2123 & 2221 & 2222 & 2223 & 2321 & 2322 & 2323 \\ \hline \end{array}\right],$$

where $\tau_{i,j,k,\ell}$ is located at the (k,ℓ)-th entry of the (i,j)-th block. Then T is symmetric with respect to the partitioning $\boldsymbol{\alpha} = \{1,2\}$ and $\boldsymbol{\beta} = \{3,4\}$ if $\tau_{i,j,k,\ell} = \tau_{k,\ell,i,j}$. We pair these relationships in colors. If we reshuffle the very same entries and fold them into a tensor $S \in \mathbb{R}^{2\times2\times3\times3}$ represented by

$$S = \left[\begin{array}{|ccc|ccc|} \hline 1111 & 1112 & 1113 & 1211 & 1212 & 1213 \\ 1121 & 1122 & 1123 & 1221 & 1222 & 1223 \\ 1131 & 1132 & 1133 & 1231 & 1232 & 1233 \\ \hline 2111 & 2112 & 2113 & 2211 & 2212 & 2213 \\ 2121 & 2122 & 2123 & 2221 & 2222 & 2223 \\ 2131 & 2132 & 2133 & 2231 & 2232 & 2233 \\ \hline \end{array}\right],$$

then S is symmetric with respect to either $(\{1,3\},\{2,4\})$ or $(\{1,4\},\{2,3\})$ if $\sigma_{i,j,k,\ell} = \sigma_{j,i,\ell,k}$. It is interesting to note that in this case, if S is flattened as shown above, then S is a symmetric matrix.

LEMMA 2.2. *Suppose that the given tensor T is symmetric and that $\boldsymbol{\beta}$ has cardinality 2. Then, with respect to any $k-2$ vectors $\mathbf{u}^{(\alpha_1)}, \ldots, \mathbf{u}^{(\alpha_{k-2})} \in \mathbb{R}^n$, the product $T \circledast_{\boldsymbol{\beta}} \mathbf{u}^{(\alpha_1)} \otimes \ldots \otimes \mathbf{u}^{(\alpha_{k-2})}$ is a symmetric matrix.*

Proof. For convenience, write $S := \mathbf{u}^{(\alpha_1)} \otimes \ldots \otimes \mathbf{u}^{(\alpha_{k-2})}$ and $M := T \circledast_{\boldsymbol{\beta}} S$. By the definition (2.2), M is a matrix. Observer that

$$m_{ij} = \langle \tau_{[:|i,j]}, S \rangle = \langle \tau_{[:|j,i]}, S \rangle = m_{ji},$$

because $\tau_{[:|i,j]} = \tau_{[:|j,i]}$ by the symmetry of T. □

REFERENCES

[1] P. COMON, G. GOLUB, L.-H. LIM, AND B. MOURRAIN, *Symmetric tensors and symmetric tensor rank*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1254–1279.

[2] S. FRIEDLAND, *Best rank one approximation of real symmetric tensors can be chosen symmetric*, Front. Math. China, 8 (2013), pp. 19–40.

[3] T. G. KOLDA, *Numerical optimization for symmetric tensor decomposition*, Math. Program., 151 (2015), pp. 225–248.

[4] G. I. MARCHUK, *Construction of adjoint operators in non-linear problems of mathematical physics*, Sbornik: Mathematics, 189 (1998), p. 1505.

[5] L. QI, *Eigenvalues of a real supersymmetric tensor*, J. Symb. Comput., 40 (2005), pp. 1302–1324.

[6] L. WANG AND M. T. CHU, *On the global convergence of the alternating least squares method for rank-one approximation to generic tensors*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 1058–1072.