A COMMENT ON THE BEST RANK-1 APPROXIMATION OF A SYMMETRIC TENSOR

MOODY T. CHU*

Abstract. There has been considerable interest in the best rank-1 approximation to a symmetric tensor. In a recent paper, it was shown that the best symmetric rank-1 approximation to a symmetric tensor is its best rank-1 approximation. The proof was based on the mathematical induction. This paper offers a much simpler argument.

Key words. symmetric tensor, best rank-1 approximation, singular value decomposition

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. A tensor of the form

$$\mathbf{u}^{(1)} \otimes \ldots \otimes \mathbf{u}^{(k)} := \left[u_{i_1}^{(1)} \ldots u_{i_k}^{(k)} \right],$$

where elements are the products of entries from vectors $\mathbf{u}^{(j)} \in \mathbb{R}^{I_j}$, $j = 1, \ldots, k$, is said to be of rank one. When $I_1 = \ldots = I_k$, we have a square tensor. An order-k square tensor T is said to be symmetric 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 $\{1, \ldots k\}$. A symmetric rank-1 tensor therefore necessarily implies that $\mathbf{u}^{(1)} = \ldots = \mathbf{u}^{(k)}$.

The problem of finding a best rank-1 approximation to T is to determine unit vectors $\mathbf{u}^{(j)} \in \mathbb{R}^{I_j}$, $j = 1, \ldots k$ and a scalar λ such that the functional

$$f\left(\lambda, \mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(k)}\right) := \left\| T - \lambda \mathbf{u}^{(1)} \otimes \ldots \otimes \mathbf{u}^{(k)} \right\|_F^2 = \sum_{i_1, i_2, \ldots, i_k} \left(t_{i_1,\ldots,i_k} - \lambda u_{i_1}^{(1)} \ldots u_{i_k}^{(k)} \right)^2 \tag{1.2}$$

is minimized. It is not difficult to see that for any fixed unit vectors $\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(k)}$, the optimal value of λ for (1.2) is given by

$$\lambda = \lambda\left(\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(k)}\right) = \left\langle T, \mathbf{u}^{(1)} \otimes \ldots \otimes \mathbf{u}^{(k)} \right\rangle. \tag{1.3}$$

In other words, by regarding the set of order-k tensors as a vector space equipped with the Frobenius inner product, the expression for λ in (1.3) is precisely the length of the projection of the "vector" T onto the direction of the "unit vector" $\mathbf{u}^{(1)} \otimes \ldots \otimes \mathbf{u}^{(k)}$. Thus, minimizing the orthogonal component of T, as is desired in (1.2), is equivalent to maximizing the parallel component λ. In [13], the expression (1.3) is called the generalized Rayleigh quotient. Many approaches for finding the maximizer in (1.3) have been proposed in the literature. See, for example, [2, 4, 5, 7, 8, 9].

It has been conjectured [11] and recently proved that the best rank-1 approximation to a symmetric tensor is symmetric [14, Theorem 2.1]. The proof was by induction. See related discussions in [1, 3, 6, 7, 10]. We think that such a result can easily be understood with the notion of conventional singular value decomposition for matrices. All we need is the following fundamental fact.

*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.

LEMMA 1.1. *Given a matrix $A \in \mathbb{R}^{m\times n}$, the global maximum of the generalized Rayleigh quotient*

$$\max_{\substack{\mathbf{y}\in\mathbb{R}^m,\,\|\mathbf{y}\|=1\\ \mathbf{z}\in\mathbb{R}^n,\,\|\mathbf{z}\|=1}} \mathbf{y}^\top A\mathbf{z} \tag{1.4}$$

is precisely the largest singular value of A, where the global maximizer $(\mathbf{y},\mathbf{z})$ consists of precisely the corresponding left and right singular vectors. In the event that $A \in \mathbb{R}^{m\times m}$ is symmetric, then $\mathbf{y}=\mathbf{z}$ and, hence, the best rank-1 approximation to A is symmetric.

2. Tensor as a linear transformation. Suppose that the set $\{1,\dots,k\}$ is partitioned as the union of two disjoint nonempty subsets $\boldsymbol{\alpha}=\{\alpha_1,\dots,\alpha_s\}$ and $\boldsymbol{\beta}=\{\beta_1,\dots,\beta_t\}$, where $s+t=k$. For convenience, an element $\tau_{i1,\dots,i_k}$ will be marked as $\tau_{[i_{\boldsymbol{\alpha}};i_{\boldsymbol{\beta}}]}$ where $i_{\boldsymbol{\alpha}}$ and $i_{\boldsymbol{\beta}}$ contain those indices at locations $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, respectively. For example, if $\boldsymbol{\alpha}=\{2,4\}$ and $\boldsymbol{\beta}=\{1,3,5,6\}$, then $\tau_{2;3;1;5;6;4}=\tau_{[3,5;2,1,6,4]}$. We shall regard an order-k tensor $T\in\mathbf{R}^{I_1\times\dots\times I_k}$ as a "matrix representation" of a linear operator mapping order-s tensors to order-t tensors [12]. Specifically, we identify T with the map

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

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

$$\mathscr{T}_{\boldsymbol{\beta}}(S) := T\circledast_{\boldsymbol{\beta}} S = \left[\langle \tau_{[:;\ell_1,\dots,\ell_t]}, S\rangle\right] \in \mathbb{R}^{I_{\beta_1}\times\dots\times I_{\beta_t}}. \tag{2.2}$$

In the above, $\tau_{[:;\ell_1,\dots,\ell_t]}$ denotes the $(\ell_1,\dots,\ell_t)$-th "slice" in the $\boldsymbol{\beta}$ direction of the tensor T, that is, ℓ_j occurs at the β_j-th location in the array $\{1,\dots,k\}$ and assumes values $\ell_j=1,\dots I_{\beta_j}$, for $j=1,\dots t$; whereas ":" denotes a wild card at the $(\alpha_1,\dots,\alpha_s)$ location to be summed over and

$$\langle \tau_{[:;\ell_1,\dots,\ell_t]}, S\rangle := \sum_{i_1=1}^{I_{\alpha_1}}\cdots\sum_{i_s=1}^{I_{\alpha_s}} \tau_{[i_1,\dots,i_s;\ell_1,\dots,\ell_t]} s_{i_1,\dots,i_s} \tag{2.3}$$

is the Frobenius inner product generalized to multi-dimensional arrays. The following results will be extremely useful in the subsequent discussion.

LEMMA 2.1. *Given an order-k tensor T and any partitioning $\{1,\dots,k\}=\boldsymbol{\alpha}\cup\boldsymbol{\beta}$, it is true that*

$$\left\langle T, \mathbf{u}^{(1)}\otimes\dots\otimes\mathbf{u}^{(k)}\right\rangle = \left\langle T\circledast_{\boldsymbol{\beta}} \mathbf{u}^{(\alpha_1)}\otimes\dots\otimes\mathbf{u}^{(\alpha_s)}, \mathbf{u}^{(\beta_1)}\otimes\dots\otimes\mathbf{u}^{(\beta_t)}\right\rangle. \tag{2.4}$$

Proof. Based on the definition (2.2), the right hand side of (2.4) is simply a rearrangement of terms in the summation by the associative law. □

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 $\mathbf{u}^{(\alpha_1)},\dots,\mathbf{u}^{(\alpha_{k-2})}\in\mathbb{R}^n$, the product $T\circledast_{\boldsymbol{\beta}}\mathbf{u}^{(\alpha_1)}\otimes\dots\otimes\mathbf{u}^{(\alpha_{k-2})}$ is a symmetric matrix.*

Proof. For convenience, write $S:=\mathbf{u}^{(\alpha_1)}\otimes\dots\otimes\mathbf{u}^{(\alpha_{k-2})}$ 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. □

2.1. Best rank-1 approximation. Suppose that $\mathbf{u}^{(1)}\otimes\dots\otimes\mathbf{u}^{(k)}$ is the best rank-1 approximation to a given order-k symmetric tensor T. By (1.3), the generalized Rayleigh quotient $\lambda=\left\langle T,\mathbf{u}^{(1)}\otimes\dots\otimes\mathbf{u}^{(k)}\right\rangle$ is maximal. Consider the case $\boldsymbol{\beta}=\{1,2\}$. By Lemma 2.1, we can write

$$\lambda=\left\langle T\circledast_{\boldsymbol{\beta}}\mathbf{u}^{(3)}\otimes\dots\otimes\mathbf{u}^{(k)},\mathbf{u}^{(1)}\otimes\mathbf{u}^{(2)}\right\rangle.$$

By Lemma 2.2, $T\circledast_{\boldsymbol{\beta}}\mathbf{u}^{(3)}\otimes\dots\otimes\mathbf{u}^{(k)}$ is a symmetric matrix. By Lemma 1.1, we conclude that $\mathbf{u}^{(1)}=\mathbf{u}^{(2)}$. Moving to the choice $\boldsymbol{\beta}=\{2,3\}$, we than have $\mathbf{u}^{(2)}=\mathbf{u}^{(3)}$. Continuing this process, we conclude that $\mathbf{u}^{(1)}=\dots=\mathbf{u}^{(k)}$. So the best rank-1 approximation to a symmetric tensor is necessarily symmetric.

REFERENCES

[1] K. BATSELIER, H. LIU, AND N. WONG, *A constructive algorithm for decomposing a tensor into a finite sum of orthonormal rank-1 terms*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 1315–1337.

[2] J. BRACHAT, P. COMON, B. MOURRAIN, AND E. TSIGARIDAS, *Symmetric tensor decomposition*, Linear Algebra Appl., 433 (2010), pp. 1851–1872.

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

[4] S. FRIEDLAND, V. MEHRMANN, R. PAJAROLA, AND S. K. SUTER, *On best rank one approximation of tensors*, Numer. Linear Algebra Appl., 20 (2013), pp. 942–955.

[5] M. ISHTEVA, P.-A. ABSIL, AND P. VAN DOOREN, *Jacobi algorithm for the best low multilinear rank approximation of symmetric tensors*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 651–672.

[6] Y.-L. JIANG AND X. KONG, *On the uniqueness and perturbation to the best rank-one approximation of a tensor*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 775–792.

[7] E. KOFIDIS AND P. A. REGALIA, *On the best rank-1 approximation of higher-order supersymmetric tensors*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 863–884.

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

[9] T. G. KOLDA AND B. W. BADER, *Tensor decompositions and applications*, SIAM Rev., 51 (2009), pp. 455–500.

[10] G. NI AND Y. WANG, *On the best rank-1 approximation to higher-order symmetric tensors*, Math. Comput. Modelling, 46 (2007), pp. 1345–1352.

[11] L. QI, *The best rank-one approximation ratio of a tensor space*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 430–442.

[12] 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.

[13] T. ZHANG AND G. H. GOLUB, *Rank-one approximation to high order tensors*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 534–550 (electronic).

[14] X. ZHANG, C. LING, AND L. QI, *The best rank-1 approximation of a symmetric tensor and related spherical optimization problems*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 806–821.