TENSOR RANK, SIMULTANEOUSLY DIAGONALIZATION, AND SOME RELATED MATRIX VARIETIES

Shuchao Bi
University of California, Berkeley

Abstract

In this article, I will first give a criterion for a generic $m \times n \times n$ tensor to have rank n using simultaneously diagonalization of certain set of matrices. Then, in Section 3, I will study some properties (like irreducibility and diemnsion) of the varieties defined by the above simultaneous diagonalization condition which are closely related to the commuting varieties. This criterion also provides a possible way to attack the "Salmon Conjecture".

1. Introduction

Fix a field K, a d dimensional tensor over K is just a d dimensional table $T \in K^{n_1 \times \cdots \times n_d}$ where $n_k \in \mathbb{N}$. A d dimensional tensor has rank 1 (or is decomposable) if it can be written as a product of d vectors $v_k \in K^{n_k}$, i.e., $T_{i_1,\ldots,i_d} = (v_1)_{i_1} \cdots (v_d)_{i_d}$ for any indices $i_k \in [n_k]$. In algebraic geometry language, the rank 1 tensors are the image of the Segre embedding

$$K^{n_1} \times \cdots \times K^{n_d} \longrightarrow K^{n_1 \times \cdots \times n_d}.$$

The rank of a tensor T is the least number r such that T can be written as a sum of r rank 1 tensors. Determining the rank of a tensor has always been an interesting and important problem in algebraic complexity theory[23], algebraic statistics[6][24], engineering[12] and algebraic geometry[21].

2. The Key Criterion

In this section, we are working over any fixed field K. For a $m \times n \times n$ tensor X, let X_1, X_2, $\cdots$, X_m (which are $n \times n$ matrices) denote the slices in the first direction.

Theorem 1. Let X be a $m \times n \times n$ tensor with X_1 nonsingular. Then X has rank n if and only if the set of matrices

$$\{X_j X_1^{-1} : j = 2, ...m\}$$

can be diagonalized simultaneously.

Remark. The condition of the theorem can be weakened. In fact, if there exists a nonsingular linear combination of slices $X_1, X_2, \cdots, X_m$, then we can just replace X_1 by that linear combination. This operation doesn't change the rank at all. Note also that all linear combinations of X_i being singular is an algebraic condition, it amounts to say that $\det(\sum_{i=1}^n \lambda_i X_i) \equiv 0$ for all $\lambda_i \in \mathbb{C}$, i.e. all the coefficients of $\prod_{i=1}^n \lambda_i^{\alpha_i}$ where $\sum \alpha_i = n$ are zero. These are algebraic conditions on the entries of X.

Proof. "$\Rightarrow$": Suppose X has rank n, i.e. there exist $x_i \in K^m$, a_j, b_k $\in K^n$ for i, j, k= 1, ... ,n, such that

$$X = x_1 \otimes a_1 \otimes b_1 + x_2 \otimes a_2 \otimes b_2 + ... + x_n \otimes a_n \otimes b_n,$$

where the (i, j, k)-th spot of the decomposable tensor $x_t \otimes a_t \otimes b_t$ is $x_{ti} a_{tj} b_{tk}$ for any t=1,...,n.

Then we have

$$X_1 = x_{11} a_1 b_1' + x_{21} a_2 b_2' + ... + x_{n1} a_n b_n', \tag{1}$$
$$X_2 = x_{12} a_1 b_1' + x_{22} a_2 b_2' + ... + x_{n2} a_n b_n',$$
$$\vdots$$
$$X_m = x_{1m} a_1 b_1' + x_{2m} a_2 b_2' + ... + x_{nm} a_n b_n',$$

where b_i' denote the transpose of b_i.

If $a_1, a_2, ..., a_n$ are linearly dependent. Then from equation (1), we see that X_1 will has rank $\leq n-1$. This contradicts the hypothesis that X_1 is nonsingular. It follows that $a_1, a_2, ..., a_n$ are linearly independent. Similarly, $b_1, b_2, ..., b_n$ are linearly independent. So we can find invertible matrices P and Q such that $Pa_i = e_i$ and $Qb_i = e_i$, where $\{e_i$=(0...1...0) : i= 1,...,n $\}$ is the standard basis for K^n. Multiply P from the left and Q^T from the right to matrices $X_1, X_2, \cdots, X_m$, we get

$$PX_1Q^T = x_{11} Pa_1 b_1' Q^T + x_{21} Pa_2 b_2' Q^T + ... + x_{n1} Pa_n b_n' Q^T,$$
$$PX_2Q^T = x_{12} Pa_1 b_1' Q^T + x_{22} Pa_2 b_2' Q^T + ... + x_{n2} Pa_n b_n' Q^T,$$

$$\vdots$$
$$PX_mQ^T = x_{1m}Pa_1b_1'Q^T + x_{2m}Pa_2b_2'Q^T + ... + x_{nm}Pa_nb_n'Q^T,$$

that is

$$\begin{aligned} PX_1Q^T &= x_{11}e_1e_1' + x_{21}e_2e_2' + ... + x_{n1}e_ne_n', \\ PX_2Q^T &= x_{12}e_1e_1' + x_{22}e_2e_2' + ... + x_{n2}e_ne_n', \\ &\vdots \\ PX_mQ^T &= x_{1m}e_1e_1' + x_{2m}e_2e_2' + ... + x_{nm}e_ne_n', \end{aligned}$$

In other words, PX_iQ^T is the following diagonal matrix

$$\begin{pmatrix} x_{i1} & 0 & \cdots & 0 \\ 0 & x_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_{in} \end{pmatrix}$$

We denote this matrix by D_i. Then

$$\begin{aligned} PX_iQ^T(PX_1Q^T)^{-1} &= PX_iQ^T(Q^T)^{-1}X_1^{-1}P^{-1} \\ &= PX_iX_1^{-1}P^{-1} = D_iD_1^{-1}. \end{aligned}$$

From the last equality, we see the set of matrices

$$\{X_iX_1^{-1} : i = 2, ..., n\}$$

can be diagonalized simultaneously.

"$\Leftarrow$": Suppose the set of matrices

$$\{X_iX_1^{-1} : i = 2, ..., n\}$$

can be diagonalized simultaneously, that is, there exist invertible matrix P such that for $i = 2, \dots, n$,

$$PX_iX_1^{-1}P^{-1} = E_i$$

where E_is are diagonal matrices.

Suppose

$$E_i = \begin{pmatrix} e_1^i & 0 & \cdots & 0 \\ 0 & e_2^i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e_n^i \end{pmatrix}$$

For notation consistence, let $E_1 := I_n$. Then for $i = 1, ..., n$

$$X_i = P^{-1} E_i P X_1.$$

Suppose

$$P^{-1} = (a_1,\ a_2,\ ...,\ a_n)$$

and

$$PX_1 = \begin{pmatrix} b_1' \\ b_2' \\ \vdots \\ b_n' \end{pmatrix}$$

for a_i, $b_j \in K^n$.

Then

$$X_i = P^{-1} E_i P X_1 = (a_1,\ a_2,\ ...\ ,\ a_n) \begin{pmatrix} e_1^i & 0 & \cdots & 0 \\ 0 & e_2^i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e_n^i \end{pmatrix} \begin{pmatrix} b_1' \\ b_2' \\ \vdots \\ b_n' \end{pmatrix}$$

$$= e_1^i a_1 b_1' + e_2^i a_2 b_2' + ... + e_n^i a_n b_n'.$$

The above formula implies X has rank $\leq n$, but X_1 is nonsingular implies X has rank $\geq n$. So we have X has rank exactly n. □

3. Related Varieties

We know that the set of matrices

$$\{X_j X_1^{-1} : j = 2, ...m\}$$

can be diagonalized simultaneously if and only if each of them is diagonalizable and they commute. This set of matrices commute means for any i, j,

$$X_iX_1^{-1}X_jX_1^{-1} = X_jX_1^{-1}X_iX_1^{-1},$$

i.e.,

$$X_iX_1^{-1}X_j = X_jX_1^{-1}X_,$$

i.e.,

$$X_iad(X_1)X_j = X_jad(X_1)X_i,$$

where $\mathrm{ad}(X_1)$ denotes the adjoint (adjugate) of X_1. The above formulas are polynomial equations. These observations motivates me to study the algebraic variety defined by these algebraic equations. In fact, in the case $n = m = 4$, the study of the irreducible components of this variety is crucial to solve the so called "Salmon Conjecture".

From now on, let us work over the complex numbers $\mathbb{C}$. We are concerned with the varieties

$$V_{m,n} = \{(X_1, ..., X_m) \in \mathbb{C}^{m\times n^2} | X_iad(X_1)X_j = X_jad(X_1)X_i \text{ for any } i, j\}.$$

The irreducibility of $V_{m,n}$ is investigated in the following theorems.

Theorem 2. $V_{m,n}$ is reducible for $m \geq 5,\ n \geq 4$.

Proof. Let $V'_{m,n}$ denotes the subset of $V_{m,n}$ where the first slice X_1 is invertible. Let $C(m, n)$ denote the variety of m-tuple of commuting $n \times n$ matrices, i.e.,

$$C(m, n) = \{(A_1, ..., A_m) \in \mathbb{C}^{m\times n^2} | A_iA_j = A_jA_i \text{ for any } i, j\}.$$

Consider the map,

$$\varphi : V'_{m,n} \longrightarrow C(m - 1, n)$$

defined by

$$(X_1, ..., X_m) \longmapsto (X_2X_1^{-1}, ..., X_mX_1^{-1}).$$

This is a well-defined map by definition. For any $(A_1, ..., A_{m-1}) \in C(m, n)$, we have $(I, A_1, ..., A_{m-1}) \in V(m, n)'$ which maps to $(A_1, ..., A_{m-1})$ by φ, so φ is surjective. Gerstenhaber[7] proved that $C(m, n)$ is reducible for $m \geq 4,\ n \geq 4$. This implies that $V'_{m,n}$ is reducible for $m \geq 5,\ n \geq 4$, since otherwise $C(m - 1, n)$ would be irreducible as the image of a irreducible variety.

To prove $V_{m,n}$ is reducible, let

$$G = \{(X_1, ..., X_m) \in \mathbb{C}^{m\times n^2} | X_1 \text{ nonsingular}\}.$$

G is defined by non-vanishing of a polynomial, so it is open and dense in $\mathbb{C}^{m\times n^2}$. Suppose $V_{m,n}$ is irreducible, then $V'_{m,n} = V_{m,n} \cap G$ (which is not empty) as a open and dense subset of $V_{m,n}$ is also irreducible. Contradiction. □

Theorem 3. $V_{4,n}$ is reducible for $n \geq 30$.

Proof. The proof of Theorem 3 is essentially the same except that here we applied the fact that $C(3, n)$ is reducible for $n \geq 30$, see Guralnick and Sethuraman [10]. □

Lemma 4. $V'_{3,n}$ is irreducible for any n.

Proof. Consider the map

$$\varphi : V'_{3,n} \longrightarrow C(2, n)$$

defined in the proof of Theorem 2. We apply the classical result of Motzkin-Taussky[18] and Gerstenhaber[7] that the variety of commuting pairs of matrices $C(2, n)$ is irreducible. Since φ is surjective, from algebraic geometry, it is suffices to show that each fiber is irreducible and has the same dimension. For any $(A, B) \in C(2, n)$,

$$\varphi^{-1}(A, B) = \{(X_1, X_2, X_2) \in V'_{3,n} | X_2X_1^{-1} = A, X_3X_1^{-1} = B\}.$$

In the fiber, X_1 determines X_2 and X_3, and X_1 can be any nonsingular matrix. So for any $(A, B) \in C(2, n)$,

$$\varphi^{-1}(A, B) \cong GL(n, \mathbb{C})$$

which is irreducible. □

Conjecture 1. $V_{3,n}$ is irreducible for any n.

Conjecture 1 is verified in the case $n = 3$ in [6] and [16], both using computational method. They also showed that $V_{3,3}$ is arithmetically Cohen-Macaulay.

Use the similar methods, we get the following lemmas together with corresponding conjectures.

Lemma 5. $V'_{4,n}$ is irreducible for $n \leq 8$.

In lemma 5, we apply the fact that $C(3, n)$ is irreducible for $n \leq 8$, see Guralnick and Sethuraman [10], Holbrook[?], Omladič[?], Han[?], Šivic[?]. From this lemma, I conjectured that

Conjecture 2. $V_{4,n}$ is irreducible for $n \leq 8$.

Also KirillovCNeretin [?] had proved that $C(m,2)$ and $C(m,3)$ are irreducible for any m. Along with this, we get

Lemma 6. $V'_{m,2}$ and $V'_{m,3}$ are irreducible for any m.

Conjecture 3. $V_{m,2}$ and $V_{m,3}$ are irreducible for any m.

An important consequence of irreducibility of $V'_{m,n}$ or $V_{m,n}$ is worth noting. Let

$$D = \{(X_1, ..., X_m) \in \mathbb{C}^{m\times n^2} | X_i ad(X_1) \text{ is diagonalizable}\}$$

and

$$E = \{(X_1, ..., X_m) \in \mathbb{C}^{m\times n^2} | X_i ad(X_1) \text{ has distinct eigenvalues}\}.$$

We have $E \subset D$ and E is open and dense in $\mathbb{C}^{m\times n^2}$, as it is defined as the non-vanishing of the discriminants of characteristic polynomials of $X_i ad(X_1)$. If $V'_{m,n}$ (or $V_{m,n}$) is irreducible, then $V'_{m,n} \cap E$ ($V_{m,n} \cap E$) is open and dense in $V'_{m,n}$ ($V_{m,n}$). That implies $V'_{m,n} \cap D$ ($V_{m,n} \cap D$) is dense in $V'_{m,n}$ ($V_{m,n}$).

References

1. R. Basili, On the irreducibility of varieties of commuting matrices, J. Pure Appl. Algebra, 149 (2000) 107-120.

2. R. Basili, On the number of irreducible components of commuting varieties, J. Pure appl. algebra 149 (2000) 121-126.

3. J. Berge, Kruskal's Polynomial for 2 x 2 x 2 Arrays and a Generalization to 2 x n x n Arrays, Psychometrika, v56 n4 p631-36, 1991.

4. W. Bruns and U. Vetter, Determinantal Rings, Springer Lecture Notes 1327, Springer-Verlag, Berlin, Heidelberg, New York, 1988.

5. D. Eisenbud, Commutative algebra. With a view toward algebraic geometry. Graduate Texts in Mathematics, 150. Springer-Verlag, New York, 1995.

6. L. Garcia, Michael Stillman, and Bernd Sturmfels, Algebraic geometry of Bayesian networks, J. Symbolic Comput. 39 (2005), no. 3-4, 331C355. MR MR2168286 (2006g:68242)

7. M. Gerstenhaber, On dominance and varieties of commuting matrices, Ann. Math. 73 (1961) 324-348.

8. R. Guralnick, A note on pairs of matrices with rank one commutator, Linear Multilinear Algebra 8 (1979) 97-99.

9. R. Guralnick, A note on commuting pairs of matrices, Linear and Multilinear Algebra 31 (1992) 71-75.

10. R. Guralnick and B. Sethuraman, Commuting pairs and triples of matrices and related varieties, Linear Algebra and its Applications 310 (2000) 139-148.

11. K. Hulek, A remark on certain matrix varieties, Linear Multilinear Algebra 10 (1981) 169-172.

12. J. Kruskal, Rank, decomposition, and uniqueness for 3-way and N-way arrays. In R. Coppi & S. Bolasco (Eds.), Multiway data analysis (pp. 7-18), 1989. Amsterdam: North-Holland.

13. J. Landsberg and L. Manivel, On the ideals of secant varieties of Segre varieties Found. Comput. Math. 4 (2004), no. 4, 397C422.

14. J. Landsberg and L. Manivel, Generalizations of strassens equations for secant varieties of segre varieties, preprint, math.AG/0601097.

15. J. Landsberg and J. Morton, The Geometry of Tensors: Application to complexity, statistics and engineering.

16. J. Landsberg and J. Morton, On the ideals and singularities of secant varieties of Segre varieties, preprint.

17. T. Lickteig, Typical tensorial rank, Linear Algebra Appl. 69 (1985), 95C120.

18. T. Motzkin, O. Taussky, Pairs of matrices with property L. II, Trans. Amer. Math. Soc. 80 (1955) 387-401.

19. M. Neubauer, The variety of pairs of matrices with rank[A, B]$\leq$1, Proc. Amer. Math. Soc. 105 (4) (1989) 787-792.

20. M. Neubauer and B. Sethuraman, Commuting Pairs in the Centralizers of 2-Regular Matrices, Journal of Algebra 214, 174-181(1999).

21. G. Ottaviani, Symplectic bundles on the plane, secant varieties and Lüroth quartics revisited, preprint math.AG/0702151.

22. Toshio Sakata, Toshio Sumi, Mitsuhiro Miyazaki, A simple estimation of the maximal rank of tensors with two slices by row and column operations, symmetrization and induction, arXiv:0808.2688, 2008.

23. V. Strassen, Rank and optimal computation of generic tensors, Linear Algebra Appl. 52/53 (1983), 645C685.

24. B. Sturmfels, Open problems in algebraic statistics, arXiv:0707.4558, 2008.