Ann Inst Stat Math (2010) 62:807–822
DOI 10.1007/s10463-010-0294-5

About the maximal rank of 3-tensors over the real and the complex number field

Toshio Sumi · Mitsuhiro Miyazaki · Toshio Sakata

Received: 24 March 2009 / Revised: 1 December 2009 / Published online: 16 April 2010

Abstract Tensor data are becoming important recently in various application fields. In this paper, we consider the maximal rank problem of 3-tensors and extend Atkinson and Stephens' and Atkinson and Lloyd's results over the real number field. We also prove the assertion of Atkinson and Stephens: $\text{max.rank}_{\mathbb{R}}(m, n, p) \le m + \lfloor p/2 \rfloor n$, $\text{max.rank}_{\mathbb{R}}(n, n, p) \le (p+1)n/2$ if p is even, $\text{max.rank}_{\mathbb{F}}(n, n, 3) \le 2n-1$ if $\mathbb{F} = \mathbb{C}$ or n is odd, and $\text{max.rank}_{\mathbb{F}}(m, n, 3) \le m+n-1$ if $m < n$ where $\mathbb{F}$ stands for $\mathbb{R}$ or $\mathbb{C}$.

Keywords Tensor · Maximal rank

1 Introduction

High-dimensional arrays, that is, tensor data, are becoming important recently in various application fields. For example, Miwakeichi et al. (2004) applied parallel factor analysis (PARAFAC), originated by Harshman (1970), to the squared norm of the convolution of the time/space/frequency electroencephalogram (EEG) data and its

This work was supported in part by Grant-in-Aid for Scientific Research (B) (No. 20340021) of the Japan Society for the Promotion of Science.

T. Sumi (✉) · T. Sakata
Kyushu University, 4-9-1 Shiobaru, Minami-ku, Fukuoka 815-8540, Japan
e-mail: sumi@design.kyushu-u.ac.jp

T. Sakata
e-mail: sakata@design.kyushu-u.ac.jp

M. Miyazaki
Kyoto University of Education, 1 Fujinomoricho, Fukakusa,
Fushimi-ku, Kyoto 612-8522, Japan
e-mail: g53448@kyokyo-u.ac.jp

complex Morlet wavelet transformation and extracted the temporal signatures of theta and alpha atoms and also visualized the spatial signature of them. On the other hand, as an extension of independent component analysis (ICA), Vasilescu and Terzopoulos (2005) developed multilinear ICA (MICA) using the representation of n-mode product of tensor images, which can addresses the face recognition problem under the additional factors, such as illumination, viewpoint and expression. Further, the recent survey study by Muti and Bourennane (2007) reviews tensor filtering based on multimode principal component analysis (MPCA). In chemometric, there are vast literatures, and we can find them, for example, from Malinowski (2002), Harshman et al. (2003), Smilde et al. (2004), or Wu et al. (2009).

Our paper is mainly related to PARAFAC decomposition. PARAFAC decomposes a tensor into a sum of the simplest structured tensor, that is a rank 1 tensor. One important problem of PARAFAC decomposition is its uniqueness problem. Kruskal (1977, 1989) gave several essential results about the uniqueness of PARAFAC decomposition by showing a lower bound of tensor rank. Other important problem of PARAFAC is to determine an appropriate number of factors, and Bro and Kiers (2003) proposed core consistency diagnostic (CORCONDIA) for this problem. To know the maximal rank of a model is, in this context, also quite important. Thus, the upper bound or the maximal rank of tensors has also gathered the concern of many researchers (see, for the recent reference, Kolda and Bader 2009). However, no exact estimate of the maximal rank has been obtained yet even for $n \times n \times 3$ tensors, which is our target. Here, we start giving the rigid terminology needed to give our results.

A p-tensor is an element of $\mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \cdots \otimes \mathbb{F}^{n_p}$, where $\mathbb{F}$ is the real or complex number field and $n_1, n_2, \ldots, n_p$ are positive integers. It is known that every p-tensor can be expressed as a sum of p-tensors of the form $a_1 \otimes a_2 \otimes \cdots \otimes a_p$. The rank of a tensor x is, by definition, the smallest number such that x is expressed as a sum of the tensors of the above form. Since there is a canonical basis in $\mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_p}$, there is a one to one correspondence between the set of all p-tensors and the set of p-dimensional arrays of elements of $\mathbb{F}$. In particular, a 3-tensor can be identified to $A = (A_1; A_2; \ldots; A_{n_3})$, where each A_i is an $n_1 \times n_2$ matrix. The rank of a tensor may be naturally considered to express complexity of the tensor by definition and the maximal rank of a certain type of tensors are also considered as the model complexity. The factorization of a tensor to a sum of rank 1 tensors means that the datum is expressed by a sum of data with the simplest structure, and we may have better understanding of the datum. This is an essential attitude for data analysis and, therefore, the problem of tensor factorization is an essential one for applications. For modeling data, the maximal rank of "a set of tensors" (model) is also crucially important, because an excessive rank model is redundant and deficient rank model cannot describe the data fully.

In this paper, we consider the maximal rank problem of 3-tensors. In the following, we denote $\mathbb{F}^a \otimes \mathbb{F}^b \otimes \mathbb{F}^c$ by $T(a, b, c)$ or $\mathbb{F}^{a \times b \times c}$ for simplicity and call an element of $T(a, b, c)$ an $a \times b \times c$ tensor. We denote by $\text{max.rank}_{\mathbb{F}}(a, b, c)$ the maximal rank of all tensors in $T(a, b, c)$. Kruskal (1977) studied the rank of a p-tensor and mainly obtained its lower bound. Atkinson and Stephens (1979) and Atkinson and Lloyd (1980) developed a non-linear theory based on their own several lemmas. Basically, they estimated the upper bounds by adding two diagonal matrices, which enables the two matrices to be diagonalizable simultaneously. They did not solve the problem fully, and restricted the type of tensors for obtaining clear-cut results. They obtained that (1) $\max.\text{rank}_{\mathbb{C}}(m, n, p) \le m + \lfloor p/2 \rfloor n$ and that (2) if p is even, then $\max.\text{rank}_{\mathbb{C}}(n, n, p) \le (p+1)n/2$. Atkinson and Stephens further claimed that (3) $\max.\text{rank}_{\mathbb{C}}(n, n, 3) \le 2n - 1$ and $\max.\text{rank}_{\mathbb{C}}(n, n+1, 3) \le 2n$ without proof.

In this paper, we show that (1) and (2) are also valid over the real number field (see Theorems 1, 3) and prove the claim (3) in its generalized form. That is, $\max.\text{rank}_{\mathbb{C}}(n, n+k, 3) \le 2n+k-1$ for $k \ge 0$. We also show that $\max.\text{rank}_{\mathbb{R}}(n, n+k, 3) \le 2n+k-1$ for $k \ge 1$ and that under a mild condition, an $n \times n \times 3$ tensor over $\mathbb{R}$ has rank at most $2n - 1$ (see Theorems 5, 6). As an application of this result, we give upper bounds of the rank of relatively small tensors from $T(3, 3, 3)$ to $T(6, 6, 3)$.

2 Preliminaries

We first recall some basic facts and set terminology.

Notation 1. *By $\mathbb{F}$, we express the real number field $\mathbb{R}$ or the complex number field $\mathbb{C}$.*

2. *We denote $\mathbb{F}^m \otimes \mathbb{F}^n \otimes \mathbb{F}^p$ as $T(m, n, p)$ or $\mathbb{F}^{m \times n \times p}$.*
3. *For a tensor $x \in T(m, n, p)$ with $x = \sum_{ijk} a_{ijk} \boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k$, we identify x with $(A_1; \ldots; A_p)$, where $A_k = (a_{ijk})_{1 \le i \le m, 1 \le j \le n}$ for $k = 1, \ldots, p$ is an $m \times n$ matrix, and call $(A_1; \ldots; A_p)$ a tensor.*
4. *For an $m \times n \times p$ tensor $T = (A_1; \ldots; A_p)$, $l \times m$ matrix P and $n \times k$ matrix Q, we denote by PTQ the $l \times k \times p$ tensor $(PA_1Q; \ldots; PA_pQ)$.*
5. *For an $m \times n \times p$ tensor $T = (A_1; \ldots; A_p)$, we denote by T^T the $n \times m \times p$ tensor $(A_1^T; \ldots; A_p^T)$.*
6. *For p $m \times n$ matrices $A_1, \ldots, A_p$, we denote by $(A_1, \ldots, A_p)$ the $m \times np$ matrix obtained by aligning $A_1, \ldots, A_p$ horizontally.*
7. *For $m \times n$ matrices $A_1, \ldots, A_p$, we denote by $\langle A_1, \ldots, A_p \rangle$ the vector subspace spanned by $A_1, \ldots, A_p$ in the $\mathbb{F}$-vector space of all the $m \times n$ matrices with entries in $\mathbb{F}$.*
8. *For an $m \times n$ matrix M, we denote the $m \times j$ (resp. $m \times (n - j)$) matrix consisting of the first j (resp. last $n - j$) columns of M by $M_{\le j}$ (resp. ${}_{j<}M$). We denote the $i \times n$ (resp. $(m - i) \times n$) matrix consisting of the first i (resp. last $m - i$) rows of M by $M^{\le i}$ (resp. ${}^{i<}M$). For integers $i_1, \ldots, i_r$ and $j_1, \ldots, j_s$ with $1 \le i_1 < \cdots < i_r \le m$ and $1 \le j_1 < \cdots < j_s \le n$, we denote the $r \times s$ matrix consisting of i_1th, i_2th, $\ldots$, i_rth rows and j_1th, j_2th, $\ldots$, j_sth columns of M by $M^{=\{i_1,\ldots,i_r\}}_{=\{j_1,\ldots,j_s\}}$.*
9. *We denote by E_{ij} the matrix unit whose entry in (i, j) cell is 1 and 0 otherwise.*
10. *We set*

$$\text{Diag}(A_1, A_2, \ldots, A_t) = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_t \end{pmatrix}$$

for matrices $A_1, A_2, \ldots, A_t$.

Definition 1 Let x be an element of $T(m, n, p)$. We define the rank of x, denoted by $\operatorname{rank} x$, to be $\min\{r \mid \exists \boldsymbol{a}_i \in \mathbb{F}^m, \exists \boldsymbol{b}_i \in \mathbb{F}^n, \exists \boldsymbol{c}_i \in \mathbb{F}^p \text{ for } i = 1, \ldots, r \text{ such that } x = \sum_{i=1}^r \boldsymbol{a}_i \otimes \boldsymbol{b}_i \otimes \boldsymbol{c}_i\}$. $\max\{\operatorname{rank} x \mid x \in T(m, n, p)\}$ is denoted by $\max.\operatorname{rank}_{\mathbb{F}}(m, n, p)$.

It is clear from the definition that $\operatorname{rank}(x + y) \leq \operatorname{rank} x + \operatorname{rank} y$ for any $x, y \in T(m, n, p)$.

Definition 2 For a matrix $A = (a_{ij})$, we set $\operatorname{supp}(A) := \{(i, j) \mid a_{ij} \neq 0\}$ and call it the support of A.

The following lemmas are easily verified.

Lemma 1 *Let* $(A_1; \ldots; A_p)$ *be an* $m \times n \times p$ *tensor. Then* $\operatorname{rank}(A_1; \ldots; A_p) = \min\{r \mid \exists$ *rank* 1 *matrices* $C_1, \ldots, C_r$ *such that* $\langle A_1, \ldots, A_p\rangle \subset \langle C_1, \ldots, C_r\rangle\}$. *In particular,*

1. *if* $\langle A_1, \ldots, A_p\rangle = \langle B_1, \ldots, B_q\rangle$, *then* $\operatorname{rank}(A_1; \ldots; A_p) = \operatorname{rank}(B_1; \ldots; B_q)$,
2. *for any nonsingular matrices* P *and* Q *of size* m *and* n, *respectively,*

$$\operatorname{rank}(A_1; \ldots; A_p) = \operatorname{rank}(PA_1Q; \ldots; PA_pQ)$$

   and
3. $\operatorname{rank}(A_1^T; \ldots; A_p^T) = \operatorname{rank}(A_1; \ldots; A_p)$.

Lemma 2 $\operatorname{rank}(A_1; \ldots; A_p) \geq \operatorname{rank}(A_1, \ldots, A_p)$.

From now on, we denote by $\operatorname{rank}_{\mathbb{R}}$ or $\operatorname{rank}_{\mathbb{C}}$ instead of rank to specify over which field, $\mathbb{R}$ or $\mathbb{C}$, we are working. For the statements common to both fields, we use $\operatorname{rank}_{\mathbb{F}}$.

The following lemma is well known.

Lemma 3 *Let*

$$f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n$$

be a monic polynomial with a variable λ *and coefficients in* $\mathbb{F}$. *Suppose that* $f(\lambda) = 0$ *has* n *distinct roots in* $\mathbb{F}$. *Then there is a neighborhood* U *of* $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)^T$ *in* $\mathbb{F}^n$ *such that for any* $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T \in U$,

$$\lambda^n + x_1\lambda^{n-1} + \cdots + x_n = 0$$

has n *distinct roots in* $\mathbb{F}$ *and these roots are continuous functions of* $\boldsymbol{x}$.

3 Maximal rank of 3-tensors

In this section, we show the results in the real number field, which are obtained by Atkinson and Stephens (1979) and Atkinson and Lloyd (1980) in the complex number field. We show several results that corresponds with the results given by them, but few results are slightly different and some of them are new ones. Now we prepare several lemmas which is a real version of Lemma in Atkinson and Stephens (1979). First, we show the extended version of Lemma 3 in Atkinson and Stephens (1979).

Lemma 4 *Let $A = (a_{ij})$ and $B = (b_{ij})$ be $n \times n$ matrices with entries in $\mathbb{F}$. Then there exist diagonal matrices X, Y with entries in $\mathbb{F}$ satisfying the following.*

1. *$A + X$ is nonsingular.*
2. *$(A + X)^{-1}(B + Y)$ has n distinct eigenvalues in $\mathbb{F}$.*

Moreover, if $i_1, \dots, i_r$ are integers with $1 \le i_1 < \cdots < i_r \le n$, $A^{=\{i_1,\dots,i_r\}}_{=\{i_1,\dots,i_r\}}$ is nonsingular and $(A^{=\{i_1,\dots,i_r\}}_{=\{i_1,\dots,i_r\}})^{-1}(B^{=\{i_1,\dots,i_r\}}_{=\{i_1,\dots,i_r\}})$ has r distinct eigenvalues in $\mathbb{F}$, then we can take X and Y so that the entries of the (i_u, i_u) cell of X and Y are zero for $u = 1, \dots, r$. In particular,

(a) *if $(n, n) \in \operatorname{supp}(A)$, then we can take X and Y so that the entries of the (n, n) cell of X and Y are 0.*
(b) *if $\{(n-1, n), (n, n-1)\} \subset \operatorname{supp}(A)$, $(n, n) \notin \operatorname{supp}(A) \cup \operatorname{supp}(B)$ and $b_{n-1,n}/a_{n-1,n} \ne b_{n,n-1}/a_{n,n-1}$, then we can take X and Y so that the entries of the $(n-1, n-1)$ and (n, n) cells of X and Y are 0.*

Proof First, we prove the former half of the lemma. Take distinct elements $s_1, \dots, s_n$ of $\mathbb{F}$ and set $D = \operatorname{Diag}(s_1, \dots, s_n)$. Note that if the absolute values of all entries of A' are sufficiently small, then $A' + E_n$ is nonsingular and all entries of $(A' + E_n)^{-1}$ are continuous with respect to entries of A'. Thus, $(A' + E_n)^{-1}(B' + D)$ is a continuous function with respect to A' and B' if the absolute values of their entries are sufficiently small. Since

$$\det(\lambda E_n - (A' + E_n)^{-1}(B' + D)) = 0$$

has n distinct roots $s_1, s_2, \dots, s_n$ in $\mathbb{F}$ if $A' = B' = O$, we see by Lemma 3 that there is a neighborhood of O in $\mathbb{F}^{n^2}$ such that if A' and B' are both in it, then

$$\det(\lambda E_n - (A' + E_n)^{-1}(B' + D)) = 0$$

has n distinct roots in $\mathbb{F}$. Hence, for sufficiently small $\epsilon > 0$,

$$\det(\lambda E_n - (\epsilon A + E_n)^{-1}(\epsilon B + D)) = 0$$

has n distinct roots in $\mathbb{F}$ and, therefore,

$$\det(\lambda(A + (1/\epsilon)E_n) - (B + (1/\epsilon)D)) = 0$$

has n distinct roots in $\mathbb{F}$. So it is sufficient to set $X = (1/\epsilon)E_n$ and $Y = (1/\epsilon)D$.

Next, we prove the latter half of the lemma. By permuting the rows and columns simultaneously, we may assume that $i_1 = 1, \dots, i_r = r$. Set

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

where A_{11} and B_{11} are $r \times r$ matrices. Then, by assumption, A_{11} is nonsingular and $(A_{11})^{-1}B_{11}$ has r distinct eigenvalues, say $s_1, \ldots, s_r$, in $\mathbb{F}$. We take $n - r$ distinct elements $s_{r+1}, \ldots, s_n$ from $\mathbb{F}\backslash\{s_1, \ldots, s_r\}$ and set

$$D_1 = E_{n-r}, \quad D_2 = \mathrm{Diag}(s_{r+1}, \ldots, s_n).$$

Then by the same argument as the proof of the former half, we see that

$$\begin{pmatrix} A_{11} & \epsilon A_{12} \\ \epsilon A_{21} & \epsilon^2 A_{22} + D_1 \end{pmatrix}$$

is nonsingular and

$$\det\left(\lambda E_n - \begin{pmatrix} A_{11} & \epsilon A_{12} \\ \epsilon A_{21} & \epsilon^2 A_{22} + D_1 \end{pmatrix}^{-1} \begin{pmatrix} B_{11} & \epsilon B_{12} \\ \epsilon B_{21} & \epsilon^2 B_{22} + D_2 \end{pmatrix}\right) = 0$$

has m distinct roots for sufficiently small $\epsilon > 0$. Therefore,

$$\det\left(\lambda \begin{pmatrix} A_{11} & \epsilon A_{12} \\ \epsilon A_{21} & \epsilon^2 A_{22} + D_1 \end{pmatrix} - \begin{pmatrix} B_{11} & \epsilon B_{12} \\ \epsilon B_{21} & \epsilon^2 B_{22} + D_2 \end{pmatrix}\right) = 0$$

has m distinct roots for sufficiently small $\epsilon > 0$. Since

$$\begin{aligned} &\det\left(\lambda \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} + \epsilon^{-2} D_1 \end{pmatrix} - \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} + \epsilon^{-2} D_2 \end{pmatrix}\right) \\ &= \epsilon^{-2(n-r)} \det\left(\lambda \begin{pmatrix} A_{11} & \epsilon A_{12} \\ \epsilon A_{21} & \epsilon^2 A_{22} + D_1 \end{pmatrix} - \begin{pmatrix} B_{11} & \epsilon B_{12} \\ \epsilon B_{21} & \epsilon^2 B_{22} + D_1 \end{pmatrix}\right), \end{aligned}$$

we see that it is sufficient to set $X = \epsilon^{-2}\mathrm{Diag}(O, D_1)$ and $Y = \epsilon^{-2}\mathrm{Diag}(O, D_2)$. □

The following result is well known, but we write a proof for convenience.

Proposition 1 *If $p \geq mn$, it holds*

$$\mathrm{max.rank}_{\mathbb{F}}(m, n, p) = mn.$$

Proof It is clear from the definition that $\mathrm{max.rank}_{\mathbb{F}}(m, n, p) = \mathrm{max.rank}_{\mathbb{F}}(p, m, n)$. If $A = (A_1; A_2; \ldots; A_n)$ is an $p \times m \times n$ tensor, then it is also clear from the definition that $\mathrm{rank}_{\mathbb{F}} A \geq \mathrm{rank}_{\mathbb{F}}(A_1, A_2, \ldots, A_n)$. So we see that $\mathrm{max.rank}_{\mathbb{F}}(p, m, n) \geq mn$.

Next, let $A = (a_{ijk})$ be an arbitrary 3-tensor. Then

$$A = \sum_{i=1}^{a} \sum_{j=1}^{b} \boldsymbol{e}_i \otimes \boldsymbol{e}_j \otimes (a_{ij1}, a_{ij2}, \ldots, a_{ijn})^{\mathrm{T}}.$$

Therefore, $\mathrm{rank}_{\mathbb{F}} A \leq ab$. □

We can show the real case of Lemma 4 in Atkinson and Stephens (1979).

Lemma 5 (cf. Lemma 4 in Atkinson and Stephens 1979) *Let X and Y be an $m \times m$ matrix such that X is nonsingular and each root of* $\det(\lambda X - Y) = 0$ *is in* $\mathbb{F}$ *and not repeated. Then for any $m \times (n-m)$ matrices U and V, it holds that*

$$\mathrm{rank}_{\mathbb{F}}(X, U; Y, V) \leq n.$$

Proof We can apply the proof of Lemma 4 in Atkinson and Stephens (1979). □

The following theorem is a slight generalization of Theorem 1 in Atkinson and Stephens (1979).

Theorem 1 *Let $m \leq n$ and $\mathbb{F} = \mathbb{R}, \mathbb{C}$.*

1. *if p is odd, it holds that* $\mathrm{max.rank}_{\mathbb{F}}(m, n, p) \leq m + \frac{n(p-1)}{2}$.
2. *if p is even, it holds that* $\mathrm{max.rank}_{\mathbb{F}}(m, n, p) \leq 2m + \frac{n(p-2)}{2}$ *and in addition if $m = n$, it holds that* $\mathrm{max.rank}_{\mathbb{F}}(n, n, p) \leq \frac{n(p+2)}{2} - 1$.

Proof Let $A = (A_1; \ldots; A_p) \in T(m, n, p)$. There is nonsingular matrices P and Q and integer $r \leq n$ such that $PA_pQ = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$. Then letting $B_j = PA_jQ$ for each $j = 1, \ldots, p$, we have

$$\mathrm{rank}_{\mathbb{F}}(A_1; \ldots; A_p) = \mathrm{rank}_{\mathbb{F}}(B_1; \ldots; B_p).$$

Let $D_p = B_p$ and $D_j = (D'_j, O)$ be $m \times n$ matrices with diagonal matrices D'_j for $1 \leq j < p$ such that $(B_{2i-1})_{\leq m} - D'_{2i-1}$ and $(B_{2i})_{\leq m} - D'_{2i}$ satisfy the conditions of 1 and 2 of Lemma 4 for $i = 1, \ldots, \lfloor (p-1)/2 \rfloor$. Then it holds

$$\mathrm{rank}_{\mathbb{F}}(A) \leq \mathrm{rank}_{\mathbb{F}}(D_1; \cdots; D_p) + \mathrm{rank}_{\mathbb{F}}(B_1 - D_1; \cdots; B_{p-1} - D_{p-1}; O).$$

Thus for odd integer $i = 1, 3, 5, \ldots$, we obtain $\mathrm{rank}_{\mathbb{F}}(B_i - D_i; B_{i+1} - D_{i+1}) \leq n$ by Lemma 5. Thus, if p is odd, we have

$$\begin{aligned} \mathrm{rank}_{\mathbb{F}}(A) &\leq m + \sum_{k=1}^{(p-1)/2} \mathrm{rank}_{\mathbb{F}}(B_{2k-1} - D_{2k-1}; B_{2k} - D_{2k}) \\ &\leq m + \frac{n(p-1)}{2} \end{aligned}$$

and otherwise

$$\begin{aligned} \mathrm{rank}_{\mathbb{F}}(A) &\leq m + \mathrm{rank}_{\mathbb{F}}(B_1 - D_1; B_2 - D_2) + \cdots + \mathrm{rank}_{\mathbb{F}}(B_{p-1} - D_{p-1}; O) \\ &\leq m + \frac{n(p-2)}{2} + m. \end{aligned}$$

Furthermore, if p is even and $m = n$, then

$$\mathrm{rank}_{\mathbb{F}}(A) \leq 2n + \frac{n(p-2)}{2} - 1 = \frac{n(p+2)}{2} - 1,$$

since we can take D_{p-1} so that $\mathrm{rank}(B_{p-1} - D_{p-1}) \leq n - 1$. □

Lemma 5 and Theorem 2 of Atkinson and Stephens (1979) are also true over the real number field, the proofs of which are quite similar.

Lemma 6 (Lemma 5 in Atkinson and Stephens 1979) *If $k \leq n$, then*

$$\mathrm{max.rank}_{\mathbb{F}}(m, n, mn-k) = m(n-k) + \mathrm{max.rank}_{\mathbb{F}}(m, k, mk-k).$$

Theorem 2 (Theorem 2 in Atkinson and Stephens 1979) *If $k \leq m \leq n$, then*

$$\mathrm{max.rank}_{\mathbb{F}}(m, n, mn-k) = mn - k^2 + \mathrm{max.rank}_{\mathbb{F}}(k, k, k^2-k).$$

Theorem 1 by Atkinson and Lloyd (1980) is also slightly generalized.

Theorem 3 *Let $m \leq n$. If p is even, it holds*

$$\mathrm{max.rank}_{\mathbb{F}}(m, n, p) \leq m + \frac{n(p-1)}{2}.$$

Proof Let $A = (A_1; \ldots; A_p) \in T(m, n, p)$. There are tensor T and nonsingular matrices P and Q so that $\mathrm{rank}_{\mathbb{F}}(T_1; T_2) \leq n/2$ and $P(A_p - T_1)Q$ and $P(A_{p-1} - T_2)Q$ are both of the form (D, O) with some diagonal matrix D (Sumi et al., 2009, Corollary 3.10). Set $B_j = PA_jQ$ for $j = 1, \ldots, p-2$, $D_{p-1} = P(A_{p-1} - T_2)Q$, and $D_p = P(A_p - T_1)Q$. For diagonal matrices D_j ($j = 1, \ldots, p-2$), we have

$$\begin{aligned}
\mathrm{rank}_{\mathbb{F}}(A) &\leq \mathrm{rank}_{\mathbb{F}}(B_1; \cdots; B_{p-2}; D_{p-1}; D_p) + \frac{n}{2} \\
&\leq \mathrm{rank}_{\mathbb{F}}(B_1 - D_1; \cdots; B_{p-2} - D_{p-2}; O; O) \\
&\quad + \mathrm{rank}_{\mathbb{F}}(D_1; \cdots; D_{p-2}; D_{p-1}; D_p) + \frac{n}{2} \\
&\leq \sum_{j=1}^{(p-2)/2} \mathrm{rank}_{\mathbb{F}}(B_{2j-1} - D_{2j-1}; B_{2j} - D_{2j}) + \frac{2m+n}{2}.
\end{aligned}$$

Thus, by Lemmas 4 and 5, we have

$$\mathrm{rank}_{\mathbb{F}}(A) \leq \frac{n(p-2)}{2} + \frac{2m+n}{2} = \frac{n(p-1)+2m}{2}$$

for some $D_1, \ldots, D_{p-2}$. □

4 Upper bound for the maximal rank of 3-tensors with three slices

In this section, we give a proof of the following statement (Theorem 4) asserted in Atkinson and Stephens (1979) without proof. In fact, we prove more general statements over $\mathbb{C}$ and, under mild condition, over $\mathbb{R}$ also. See Theorems 5 and 6.

Theorem 4 (Atkinson and Stephens (1979))

$$\text{max.rank}_{\mathbb{C}}(n, n, 3) \leq 2n - 1 \quad \textit{and} \quad \text{max.rank}_{\mathbb{C}}(n, n+1, 3) \leq 2n.$$

We begin with the following lemma.

Lemma 7 *Let m be an integer with $m \geq 2$. If $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s$, $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_t$ are m-dimensional non-zero vectors and $A_1, \ldots, A_u$, $B_1, \ldots, B_v$ are $m \times 2$ matrices of rank 2, then there is a nonsingular matrix P such that any entry of $P\boldsymbol{a}_i$ ($i = 1, \ldots, s$), $\boldsymbol{b}_i^T P^{-1}$ ($i = 1, \ldots, t$) and any 2-minor of PA_i ($i = 1, \ldots, u$) and $B_i^T P^{-1}$ ($i = 1, \ldots, v$) are not zero.*

Proof Let $X = (x_{ij})$ be an $m \times m$ matrix of indeterminates, i.e., $\{x_{ij}\}_{i,j=1}^m$ are independent indeterminates. None of the following polynomials of x_{ij} is zero, where $\text{Cof}(X)$ is the matrix of cofactors of X.

- $\det X$.
- jth entry of $X\boldsymbol{a}_i$.
- jth entry of $\boldsymbol{b}_i^T \text{Cof}(X)$.
- 2-minor of XA_i consisting of jth and kth rows with $1 \leq j < k \leq m$.
- 2-minor of $B_i^T \text{Cof}(X)$ consisting of jth and kth columns with $1 \leq j < k \leq m$.

So, the product $f(x_{ij})$ of all the above polynomials is not zero. Since $\mathbb{F}$ is an infinite field, we can take $p_{ij} \in \mathbb{F}$ so that $f(p_{ij}) \neq 0$. Then it is clear that we can take $P = (p_{ij})$, since $P^{-1} = (\det P)^{-1}\text{Cof}(P)$. □

In order to estimate the rank of $n \times n \times 3$ tensors, we prepare the following lemmas.

Lemma 8 *Let $(A_1; A_2; A_3)$ be an $m \times n \times 3$ tensor with $m \leq n$ such that $A_3 = (D, O)$ where D is a diagonal matrix with 0 entry in (m, m) cell and $(A_1)_{\leq m}$, $(A_2)_{\leq m}$ satisfy the condition of* (a) *or* (b) *of Lemma 4. Then it holds that* $\text{rank}_{\mathbb{F}}(A_1; A_2; A_3) \leq m + n - 1$.

Proof By Lemma 4, there are $m \times m$ diagonal matrices D_1 and D_2 with 0 entry in (m, m) cell such that $(A_1 + (D_1, O))_{\leq m}$ is nonsingular and $((A_1 + (D_1, O))_{\leq m})^{-1}$ $((A_2 + (D_2, O))_{\leq m})$ has m distinct eigenvalues. Therefore, by Lemma 5

$$\begin{aligned}
&\text{rank}_{\mathbb{F}}(A_1; A_2; A_3) \\
&\quad \leq \text{rank}_{\mathbb{F}}(A_1 + (D_1, O); A_2 + (D_2, O); O) \\
&\qquad + \text{rank}_{\mathbb{F}}(-(D_1, O); -(D_2, O); A_3) \leq n + m - 1.
\end{aligned}$$

□

Lemma 9 *Let n be an integer with $n \geq 3$ and A_1, A_2 $n \times n$ matrices with $(n, n) \notin$ $\mathrm{supp}(A_1) \cup \mathrm{supp}(A_2)$. Suppose that $(A_1)_{=\{n\}} \neq \mathbf{0}$ and $(A_1)^{=\{n\}} \neq \mathbf{0}^T$ and for any $t \in \mathbb{F}$, $(tA_1 + A_2)_{=\{n\}} \neq \mathbf{0}$ or $(tA_1 + A_2)^{=\{n\}} \neq \mathbf{0}^T$. Then there is a nonsingular $(n-1) \times (n-1)$ matrix P such that $A = \mathrm{Diag}(P, 1)A_1\mathrm{Diag}(P, 1)^{-1}$ and $B = \mathrm{Diag}(P, 1)A_2\mathrm{Diag}(P, 1)^{-1}$ satisfy the condition of* (b) *in Lemma 4.*

Proof Set

$$A_1 = \begin{pmatrix} (A_1)^{\leq n-1}_{\leq n-1} & \boldsymbol{a}_1 \\ \boldsymbol{b}_1^T & 0 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} (A_2)^{\leq n-1}_{\leq n-1} & \boldsymbol{a}_2 \\ \boldsymbol{b}_2^T & 0 \end{pmatrix}.$$

First, assume that $\mathrm{rank}(\boldsymbol{a}_1, \boldsymbol{a}_2) = 2$. Then by Lemma 7, we see that there is a nonsingular $(n-1) \times (n-1)$ matrix Q_1 such that any entry of $Q_1\boldsymbol{a}_1$ and $\boldsymbol{b}_1^T Q_1^{-1}$ and any 2-minor of $Q_1(\boldsymbol{a}_1, \boldsymbol{a}_2)$ are not zero. Set $Q_1(\boldsymbol{a}_1, \boldsymbol{a}_2) = (a_{ij})$ and $(\boldsymbol{b}_1, \boldsymbol{b}_2)^T Q^{-1} = (b_{ij})$. If $(a_{n-1,1}, a_{n-1,2})$ and $(b_{1,n-1}, b_{2,n-1})$ are linearly independent; then we can take $P = Q_1$ since

$$\mathrm{Diag}(Q_1, 1)A_i\mathrm{Diag}(Q_1, 1)^{-1} = \begin{pmatrix} Q_1(A_i)^{\leq n-1}_{\leq n-1}Q_1^{-1} & Q_1\boldsymbol{a}_i \\ \boldsymbol{b}_i^T Q_1^{-1} & 0 \end{pmatrix}.$$

If $(a_{n-1,1}, a_{n-1,2})$ and $(b_{1,n-1}, b_{2,n-1})$ are linearly dependent, then $(ta_{n-2,1} + a_{n-1,1}, ta_{n-2,2} + a_{n-1,2})$ and $(b_{1,n-1}, b_{2,n-1})$ are linearly independent for any $t \in \mathbb{F} \setminus \{0\}$, since $(a_{n-2,1}, a_{n-2,2})$ and $(a_{n-1,1}, a_{n-1,2})$ are linearly independent by the choice of Q_1. Choose $t \in \mathbb{F} \setminus \{0\}$ so that $ta_{n-2,1} + a_{n-1,1} \neq 0$ and set $Q_2 = E_{n-1} + tE_{n-1,n-2}$. Then we can take $P = Q_2Q_1$, since

$$\mathrm{Diag}(Q_2Q_1, 1)A_i\mathrm{Diag}(Q_2Q_1, 1)^{-1} = \begin{pmatrix} Q_2Q_1(A_i)^{\leq n-1}_{\leq n-1}Q_1^{-1}Q_2^{-1} & Q_2Q_1\boldsymbol{a}_i \\ \boldsymbol{b}_i^T Q_1^{-1}Q_2^{-1} & 0 \end{pmatrix},$$

$Q_2^{-1} = E_{n-1} - tE_{n-1,n-2}$, and the $(n-1, n)$ entry and $(n, n-1)$ entry of

$$\mathrm{Diag}(Q_2Q_1, 1)A_i\mathrm{Diag}(Q_2Q_1, 1)^{-1}$$

are $ta_{n-2,i} + a_{n-1,i}$ and $b_{i,n-1}$, respectively. Therefore, we have proved the case where $\mathrm{rank}(\boldsymbol{a}_1, \boldsymbol{a}_2) = 2$.

We can prove the case where $\mathrm{rank}(\boldsymbol{b}_1, \boldsymbol{b}_2) = 2$ in the same way.

Now assume that $\mathrm{rank}(\boldsymbol{a}_1, \boldsymbol{a}_2) = \mathrm{rank}(\boldsymbol{b}_1, \boldsymbol{b}_2) = 1$. Choose as before, a nonsingular $(n-1) \times (n-1)$ matrix Q_1 such that any entry of $Q_1\boldsymbol{a}_1$ and $\boldsymbol{b}_1^T Q_1^{-1}$ is not zero and set $Q_1(\boldsymbol{a}_1, \boldsymbol{a}_2) = (a_{ij})$, $(\boldsymbol{b}_1, \boldsymbol{b}_2)^T Q^{-1} = (b_{ij})$. Then, $a_{n-1,2}/a_{n-1,1} \neq b_{2,n-1}/b_{1,n-1}$, since otherwise $-(a_{n-1,2}/a_{n-1,1})\boldsymbol{a}_1 + \boldsymbol{a}_2 = -(b_{2,n-1}/b_{1,n-1})\boldsymbol{a}_1 + \boldsymbol{a}_2 = -(b_{2,n-1}/b_{1,n-1})\boldsymbol{b}_1 + \boldsymbol{b}_2 = \mathbf{0}$, which contradicts the assumption. Therefore, we can take $P = Q_1$. □

Now we state the following:

Theorem 5 *Let* $T = (A_1; A_2; A_3)$ *be an* $n \times n \times 3$ *tensor. If* $\langle A_1, A_2, A_3 \rangle$ *contains a non-zero singular matrix, then* $\mathrm{rank}_{\mathbb{F}} T \leq 2n - 1$. *In particular, if* $\mathbb{F} = \mathbb{C}$ *or* n *is odd, then* $\mathrm{rank}_{\mathbb{F}} T \leq 2n - 1$.

Proof We prove by induction on n.

Since $\mathrm{max.rank}_{\mathbb{F}}(1, 1, 3) = 1$ and $\mathrm{max.rank}_{\mathbb{F}}(2, 2, 3) = 3$, we may assume that $n \geq 3$. By Lemma 1 and the assumption, we may assume that $A_3 = \mathrm{Diag}(E_r, O)$ with $r < n$ and $\mathrm{supp}(A_1) \supset \mathrm{supp}(A_2)$.

If $(i, j) \in \mathrm{supp}(A_1)$ for some (i, j) with $i > r$ and $j > r$, by permuting rows and columns within $(r+1)$th, ..., nth one, if necessary, we can apply Lemma 8. Therefore, $\mathrm{rank}_{\mathbb{F}} T \leq 2n - 1$.

Now assume that $(i, j) \notin \mathrm{supp}(A_1)$ for any i, j with $i > r$ and $j > r$. Set ${}_{r<}(A_i)^{\leq r} = A_{12i}$ and ${}^{r<}(A_i)_{\leq r} = A_{21i}$. If there is a column vector of A_{121}, which is $\mathbf{0}$, then $\mathrm{rank}_{\mathbb{F}} T \leq n+n-1$ by Lemma 5, since $\mathrm{supp}(A_1) \supset \mathrm{supp}(A_2)$ and, therefore, T is essentially an $n \times (n-1) \times 3$ tensor in this case. Therefore, we may assume that no column vector of A_{121} is $\mathbf{0}$. We may also assume that no row vector of A_{211} is $\mathbf{0}^T$.

Set $A_{12i} = (\boldsymbol{a}_{i,r+1}, \ldots, \boldsymbol{a}_{in})$ and $A_{21i}^T = (\boldsymbol{b}_{i,r+1}, \ldots, \boldsymbol{b}_{in})$. Assume first that there is $j > r$ such that $\boldsymbol{a}_{1j}, \boldsymbol{a}_{2j}$ are linearly independent. Then by exchanging the $(r+1)$th and the jth columns, we may assume that $(A_1)_{\leq r+1}^{\leq r+1}$ and $(A_2)_{\leq r+1}^{\leq r+1}$ satisfy the condition of Lemma 9. So we take the nonsingular $r \times r$ matrix P of the conclusion of Lemma 9 and set

$$\mathrm{Diag}(P, E_{n-r}) A_k \mathrm{Diag}(P, E_{n-r})^{-1} = (a_{ijk}).$$

Then, $a_{r+1,r+1,k} = 0$ for any k and $a_{r,r+1,2}/a_{r,r+1,1} \neq a_{r+1,r,2}/a_{r+1,r,1}$. Therefore, by exchanging the $(r+1)$th and the nth rows and columns, and exchanging the rth and the $(n-1)$th rows and columns, if necessary, we may transform

$$\mathrm{Diag}(P, E_{n-r})(A_1; A_2; A_3)\mathrm{Diag}(P, E_{n-r})^{-1}$$

to a tensor that satisfies the condition of Lemma 8 (we do not need the permutation if $r = n - 1$). So the conclusion follows by Lemma 8. The case that there is $j > r$ such that $\boldsymbol{b}_{1j}, \boldsymbol{b}_{2j}$ are linearly independent is proved in the same way.

Next, assume that $\boldsymbol{a}_{1j}, \boldsymbol{a}_{2j}$ are linearly dependent and $\boldsymbol{b}_{1j}, \boldsymbol{b}_{2j}$ are linearly dependent for any $j > r$.

Since the vector space spanned by the column vectors of $(A_1)_{\leq r+1}^{\leq r}$ is at most r and the last column of $(A_1)_{\leq r+1}^{\leq r}$ is not zero, we see that there is j with $1 \leq j \leq r$ such that jth column of $(A_1)_{\leq r+1}^{\leq r}$ is a linear combination of the columns of ${}_{j<}(A_1)_{\leq r+1}^{\leq r}$. Therefore, we see that there is an $(r+1) \times (r+1)$ lower triangular unipotent matrix V such that $((A_1)_{\leq r+1} V)_{\leq r}^{\leq r} = ((A_1)_{\leq r+1}^{\leq r} V)_{\leq r}$ has a column vector, which is $\mathbf{0}$.

So by the induction hypothesis,

$$\begin{aligned}
&\mathrm{rank}_{\mathbb{F}} T \\
&\quad = \mathrm{rank}_{\mathbb{F}}(A_1; A_2; A_3) \\
&\quad = \mathrm{rank}_{\mathbb{F}}(A_1 \mathrm{Diag}(V, E_{n-r-1}); A_2 \mathrm{Diag}(V, E_{n-r-1}); A_3 \mathrm{Diag}(V, E_{n-r-1}))
\end{aligned}$$

$$
\begin{aligned}
&\leq \operatorname{rank}_{\mathbb{F}}(((A_1)_{\leq r+1}V)^{\leq r}_{\leq r}; ((A_2)_{\leq r+1}V)^{\leq r}_{\leq r}; ((A_3)_{\leq r+1}V)^{\leq r}_{\leq r}) \\
&\quad + \sum_{j=r+1}^{n} \operatorname{rank}_{\mathbb{F}}(\boldsymbol{a}_{1j}; \boldsymbol{a}_{2j}; \boldsymbol{0}) + \sum_{j=r+1}^{n} \operatorname{rank}_{\mathbb{F}}((\boldsymbol{b}_{1j}^T, 0)V; (\boldsymbol{b}_{2j}, 0)^T V; \boldsymbol{0}^T) \\
&\leq 2r - 1 + (n-r) + (n-r) \\
&= 2n - 1,
\end{aligned}
$$

since $\operatorname{rank}_{\mathbb{F}}(\boldsymbol{a}_{1j}; \boldsymbol{a}_{2j}; \boldsymbol{0}) \leq 1$ and $\operatorname{rank}_{\mathbb{F}}(\boldsymbol{b}_{1j}^T; \boldsymbol{b}_{2j}^T; \boldsymbol{0}^T) \leq 1$ for any j with $j > r$. □

It is possible that there is no non-zero singular matrix in $\langle A_1, A_2, A_3 \rangle$ over the real number field. For example, let

$$
A_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}
$$

and $A_3 = E_4$. Since the determinant of $xA_1 + yA_2 + zA_3$ is $(x^2 + y^2 + z^2)^2$, $xA_1 + yA_2 + zA_3$ is singular only when $x = y = z = 0$.

Next, we consider the non-square case. First, we prepare the following lemmas.

Lemma 10 *Let A and B be $n \times n$ matrices, $\boldsymbol{a} = (a_1, \ldots, a_n)^{\mathrm{T}}$ and $\boldsymbol{b} = (b_1, \ldots, b_n)^{\mathrm{T}}$ be n-dimensional vectors. Suppose $a_i \neq 0$ for any $i = 1, \ldots, n$. Then there are diagonal matrices X and Y and a vector $\boldsymbol{p}$ such that*

1. *$A + X$ is nonsingular,*
2. *$(A + X)\boldsymbol{p} = \boldsymbol{a}$ and $(B + Y)\boldsymbol{p} = \boldsymbol{b}$.*

Moreover, if $b_1/a_1, \ldots, b_n/a_n$ are distinct from each other, then we can take X and Y so that $(A + X)^{-1}(B + Y)$ has n distinct eigenvalues in $\mathbb{F}$.

Proof Set $A = (a_{ij})$ and $B = (b_{ij})$. For $0 < \epsilon \in \mathbb{R}$, we set

$$
\begin{aligned}
a_i(\epsilon) &= a_i - \epsilon \sum_{j=1}^{n} a_{ij}, \\
b_i(\epsilon) &= b_i - \epsilon \sum_{j=1}^{n} b_{ij}, \\
D_1(\epsilon) &= \operatorname{Diag}(a_1(\epsilon), \ldots, a_n(\epsilon)), \quad \text{and} \\
D_2(\epsilon) &= \operatorname{Diag}(b_1(\epsilon), \ldots, b_n(\epsilon)).
\end{aligned}
$$

Then,

$$
(\epsilon A + D_1(\epsilon))\boldsymbol{1} = \boldsymbol{a} \quad \text{and} \quad (\epsilon B + D_2(\epsilon))\boldsymbol{1} = \boldsymbol{b}
$$

where $\boldsymbol{1} = (1, \ldots, 1)^{\mathrm{T}}$.

By the same argument as the proof of Lemma 4, we see that $\epsilon A + D_1(\epsilon)$ is nonsingular if $\epsilon > 0$ is sufficiently small, and if $b_1/a_1, \ldots, b_n/a_n$ are distinct from each other, we can take ϵ so that $(\epsilon A + D_1(\epsilon))^{-1}(\epsilon B + D_2(\epsilon))$ has n distinct eigenvalues in $\mathbb{F}$.

Therefore, it is sufficient to set $X = (1/\epsilon)D_1(\epsilon)$, $Y = (1/\epsilon)D_2(\epsilon)$ and $\boldsymbol{p} = \epsilon\mathbf{1}$. □

Lemma 11 *Let* $(A_1; A_2)$ *be an* $m \times n \times 2$ *tensor with* $m < n$. *Set* $A_i = (\boldsymbol{a}_{i1}, \ldots, \boldsymbol{a}_{in})$ *for* $i = 1, 2$. *Suppose* $(A_1)_{\leq m}$ *is nonsingular and*

$$((A_1)_{\leq m})^{-1}(A_2)_{\leq m}$$

has m *distinct eigenvalues. Suppose also that there are integers* $j_1, \ldots, j_s$ *with* $m < j_1 < \cdots < j_s \leq n$ *and* m*-dimensional vectors* $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_s$ *such that*

$$(A_i)_{\leq m}\boldsymbol{p}_t = \boldsymbol{a}_{ij_t} \quad \text{for } i = 1, 2,\ t = 1, 2, \ldots, s.$$

Then $\operatorname{rank}_{\mathbb{F}}(A_1; A_2) \leq n - s$.

Proof Let V be the $n \times n$ upper triangular unipotent matrix, the jth column of which is $\begin{pmatrix} -\boldsymbol{p}_t \\ \mathbf{0} \end{pmatrix} + \boldsymbol{e}_{j_t}$, if $j = j_t$ for some t and $\boldsymbol{e}_j$ otherwise. Then $j_1, j_2, \ldots, j_s$th columns of A_iV are zero by the assumption and, therefore, we see by Lemma 5 that

$$\operatorname{rank}_{\mathbb{F}}(A_1; A_2) = \operatorname{rank}_{\mathbb{F}}(A_1V; A_2V) \leq n - s,$$

since $(A_1V; A_2V)$ is essentially an $m \times (n-s) \times 2$ tensor. □

Now we state the following

Theorem 6 *If* $m < n$, *then* $\operatorname{max.rank}_{\mathbb{F}}(m, n, 3) \leq m + n - 1$.

Proof We prove for an arbitrary $m \times n \times 3$ tensor $T = (A_1; A_2; A_3)$, $\operatorname{rank}_{\mathbb{F}}T \leq m + n - 1$.

Set $r = \max\{\operatorname{rank}A \mid A \in \langle A_1, A_2, A_3\rangle\}$. Then by Lemma 1, we may assume that $A_3 = (\operatorname{Diag}(E_r, O), O)$ and $\operatorname{supp}(A_1) \supset \operatorname{supp}(A_2)$.

Set $A_i = (\boldsymbol{a}_{i1}, \ldots, \boldsymbol{a}_{in})$ for $i = 1, 2$. If there is $j > m$ such that $\boldsymbol{a}_{1j} = \mathbf{0}$, then, since we are assuming that $\operatorname{supp}(A_1) \supset \operatorname{supp}(A_2)$, T is essentially an $m \times (n-1) \times 3$ tensor. So $\operatorname{rank}_{\mathbb{F}}T \leq m + n - 1$ by Theorem 1.

Now assume that $\boldsymbol{a}_{1j} \neq \mathbf{0}$ for any $j > m$.

We first consider the case where $\boldsymbol{a}_{1j}, \boldsymbol{a}_{2j}$ are linearly dependent for any j with $j > m$. Since the vector space spanned by the column vectors of $(A_1)_{\leq m+1}$ is at most m and the last column of $(A_1)_{\leq m+1}$ is not zero, we see that there is j with $1 \leq j \leq m$ such that jth column vector of A_1 is a linear combination of the column vectors of ${}_{j<}(A_1)_{\leq m+1}$. Therefore, we see that there is an $(m+1) \times (m+1)$ lower triangular unipotent matrix V such that $(((A_1)_{\leq m+1})V)_{\leq m}$ has a column vector, which is $\mathbf{0}$.

So we see by Theorem 5

$$\begin{aligned}
&\operatorname{rank}_{\mathbb{F}} T \\
&\quad = \operatorname{rank}_{\mathbb{F}}(A_1 \operatorname{Diag}(V, E_{n-m-1}); A_2 \operatorname{Diag}(V, E_{n-m-1}); A_3 \operatorname{Diag}(V, E_{n-m-1})) \\
&\quad \leq \operatorname{rank}_{\mathbb{F}}(((A_1)_{\leq m+1} V)_{\leq m}; ((A_2)_{\leq m+1} V)_{\leq m}; ((A_3)_{\leq m+1} V)_{\leq m}) \\
&\qquad + \sum_{j=m+1}^{n} \operatorname{rank}_{\mathbb{F}}(\boldsymbol{a}_{1j}; \boldsymbol{a}_{2j}; \boldsymbol{0}) \\
&\quad \leq 2m - 1 + n - m \\
&\quad = m + n - 1,
\end{aligned}$$

since $\boldsymbol{a}_{1j}, \boldsymbol{a}_{2j}$ are linearly dependent for $j > m$.

From now on, we assume that there is j with $j > m$ such that $\boldsymbol{a}_{1j}, \boldsymbol{a}_{2j}$ are linearly independent.

We first consider the case where $r = m$. By Lemma 7, we see that there is a nonsingular $m \times m$ matrix P such that any entry of $P\boldsymbol{a}_{1j}$ and any 2-minor of $P(\boldsymbol{a}_{1j}, \boldsymbol{a}_{2j})$ is not zero. Set $B_i = PA_i \operatorname{Diag}(P, E_{n-m})^{-1}$ and $B_i = (\boldsymbol{b}_{i1}, \ldots, \boldsymbol{b}_{in})$ for $i = 1, 2, 3$. Then $B_3 = (E_m, O)$ and every entry of $\boldsymbol{b}_{1j}$ and every 2-minor of $(\boldsymbol{b}_{1j}, \boldsymbol{b}_{2j})$ are not zero. So by Lemma 10, we see that there are $m \times m$ diagonal matrices D_1 and D_2 and an m-dimensional vector $\boldsymbol{p}$ such that

$((B_i)_{\leq m} + D_i)\boldsymbol{p} = \boldsymbol{b}_{ij}$ for $i = 1, 2$,
$(B_1)_{\leq m} + D_1$ is nonsingular and
$((B_1)_{\leq m} + D_1)^{-1}((B_2)_{\leq m} + D_2)$ has m distinct eigenvalues.

Therefore, by Lemma 11, we see that

$$\operatorname{rank}_{\mathbb{F}}(B_1 + (D_1, O); B_2 + (D_2, O)) \leq n - 1.$$

So

$$\begin{aligned}
&\operatorname{rank}_{\mathbb{F}} T \\
&\quad = \operatorname{rank}_{\mathbb{F}}(B_1; B_2; B_3) \\
&\quad \leq \operatorname{rank}_{\mathbb{F}}(B_1 + (D_1, O); B_2 + (D_2, O)) + \operatorname{rank}_{\mathbb{F}}(-(D_1, O); -(D_2, O); (E_m, O)) \\
&\quad \leq n - 1 + m.
\end{aligned}$$

Finally, we consider the case where $r < m$. Since $A_3 = (\operatorname{Diag}(E_r, O), O)$ and $\operatorname{rank}(tA_3 + A_1) \leq r$ for any $t \in \mathbb{F}$ by the definition of r, we see that $(i, j) \notin \operatorname{supp}(A_1)$ if $i > r$ and $j > r$.

If the $(r+1)$th row of A_1 is zero, then $(A_1; A_2; A_3)$ is essentially an $(m-1) \times n \times 3$ tensor. So

$$\operatorname{rank}_{\mathbb{F}}(A_1; A_2; A_3) \leq m - 1 + n$$

by Lemma 5. Therefore, we may assume that $(r+1)$th row of A_1 is not zero. Take j with $j > m$ such that $\boldsymbol{a}_{1j}, \boldsymbol{a}_{2j}$ are linearly independent. Exchanging the $(r+1)$th and the jth columns of A_i, we may assume that $\boldsymbol{a}_{1,r+1}, \boldsymbol{a}_{2,r+1}$ are linearly independent. By applying Lemma 9 to $(A_1)^{\le r+1}_{\le r+1}$ and $(A_2)^{\le r+1}_{\le r+1}$, we see that there is a nonsingular $r \times r$ matrix P such that $\mathrm{Diag}(P, 1)(A_1)^{\le r+1}_{\le r+1}\mathrm{Diag}(P, 1)^{-1}$ and $\mathrm{Diag}(P, 1)(A_2)^{\le r+1}_{\le r+1}\mathrm{Diag}(P, 1)^{-1}$ satisfy the condition of (b) in Lemma 4. Set $B_i = \mathrm{Diag}(P, E_{m-r})A_i\mathrm{Diag}(P, E_{n-r})^{-1}$ for $i = 1, 2, 3$. Then $B_3 = (\mathrm{Diag}(E_r, O), O)$, $(B_1)^{\le r+1}_{\le r+1}$ and $(B_2)^{\le r+1}_{\le r+1}$ satisfy the condition (b) in Lemma 4.

Let C_i be the $m \times n$ matrix obtained by exchanging the $(r+1)$th and mth rows and columns and rth and $(m-1)$th rows and columns of B_i, respectively, for $i = 1, 2, 3$. Then $(C_1)_{\le m}$ and $(C_2)_{\le m}$ satisfy the condition of (b) in Lemma 4 and $C_3 = (\mathrm{Diag}(E_{r-1}, O, 1, 0), O)$. Therefore, we see that

$$\mathrm{rank}_{\mathbb{F}} T = \mathrm{rank}_{\mathbb{F}}(C_1; C_2; C_3) \le m + n - 1$$

by Lemma 8. □

Finally, we state some upper bounds of the maximal rank for small tensors, which are direct consequences of Theorem 5.

Proposition 2 *The followings are true.*

1. $\mathrm{max.rank}_{\mathbb{F}}(3, 3, 3) \le 5$
2. $\mathrm{max.rank}_{\mathbb{C}}(4, 4, 3) \le 7$
3. $\mathrm{max.rank}_{\mathbb{F}}(5, 5, 3) \le 9$
4. $\mathrm{max.rank}_{\mathbb{C}}(6, 6, 3) \le 11$

References

Atkinson, M. D., Lloyd, S. (1980). Bounds on the ranks of some 3-tensors. *Linear Algebra and its Applications, 31*, 19–31.

Atkinson, M. D., Stephens, N. M. (1979). On the maximal multiplicative complexity of a family of bilinear forms. *Linear Algebra and its Applications, 27*, 1–8.

Bro, R., Kiers, H. A. L. (2003). A new efficient method for determining the number of components in PARAFAC models. *Journal of Chemometrics, 17*, 274–286.

Harshman, R. A. (1970). Foundations of the PARAFAC procedure: Models and conditions for an “explanatory” multi-modal factor analysis. *UCLA Working Papers in Phonetics, 16*, 1–84 (University Microfilms, Ann Arbor, No. 10,085).

Harshman, R. A., Hong, S., Lundy, M. E. (2003). Shifted factor analysis—part I: Models and properties. *Journal of Chemometrics, 17*, 363–378.

Kolda, T. G., Bader, B. W. (2009). Tensor decompositions and applications. *SIAM Review, 51*(3), 455–500.

Kruskal, J. B. (1977). Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra and its Applications, 18*, 95–138.

Kruskal, J. B. (1989). Rank, decomposition, and uniqueness for 3-way and N-way arrays. In R. Coppi, S. Bolasco (Eds.), *Multiway data analysis* (pp. 7–18). Amsterdam: Elsevier.

Malinowski, E. R. (2002). *Factor analysis in chemistry* (3rd ed.). New York: Wiley.

Miwakeichi, F., Martinez-Montes, E., Valdés-Sosa, P. A., Nishiyama, N., Mizuhara, H., Yamaguchi, Y. (2004). Decomposing EEG data into space–time–frequency components using parallel factor analysis. *NeuroImage, 22*, 1035–1045.

Muti, D., Bourennane, S. (2007). Survey on tensor signal algebraic filtering. *Signal Processing Archive 87(2)*, 237–249.

Smilde, A., Bro, R., Geladi, P. (2004). *Multi-way analysis: Applications in the chemical sciences*. New York: Wiley.

Sumi, T., Miyazaki, M., Sakata, T. (2009). Rank of 3-tensors with 2 slices and Kronecker canonical forms. *Linear Algebra and its Applications, 431*, 1858–1868.

Vasilescu, M., Terzopoulos, D. (2005). Multilinear independent components analysis. In *IEEE computer society conference on computer vision and pattern recognition, 2005, CVPR 2005* (Vol. 1, pp. 547–553).

Wu, H.-L., Nie, J.-F., Yu, Y.-J., Yu, R.-Q. (2009). Multi-way chemometric methodologies and applications: A central summary of our research work. *Analytica Chimica Acta, 650*, 131–142.