



Contents lists available at SciVerse ScienceDirect

# Linear Algebra and its Applications

journal homepage: [www.elsevier.com/locate/laa](http://www.elsevier.com/locate/laa)



## Kruskal's uniqueness inequality is sharp

Harm Derksen<sup>1</sup>

University of Michigan, Department of Mathematics, East Hall, 530 Church Street, Ann Arbor, MI 48109-1043, United States

### ARTICLE INFO

*Article history:*

Received 27 May 2010

Accepted 31 May 2011

Available online 19 August 2011

Submitted by V. Mehrmann

### ABSTRACT

Kruskal proved that a tensor in  $V_1 \otimes V_2 \otimes \cdots \otimes V_m$  of rank  $r$  has a unique decomposition as a sum of  $r$  pure tensors if a certain inequality is satisfied. We will show the uniqueness fails if the inequality is weakened. We give two different constructions for counter examples.

© 2011 Elsevier Inc. All rights reserved.

In this paper,  $\mathbb{F}$  will denote a field. Suppose that  $V$  is an  $\mathbb{F}$ -vector space. For a subset  $A \subseteq V$ , we define its Kruskal rank  $k(A)$  as the largest integer such that  $A$  has at least  $k(A)$  elements and every subset of  $A$  with  $k(A)$  elements is linearly independent.

Kruskal proved in [1,2] the following theorem for  $\mathbb{F} = \mathbb{C}$  and  $m = 3$ :

**Theorem 1** (Kruskal's theorem). *Suppose  $V_1, V_2, \dots, V_m$  are  $\mathbb{F}$ -vector spaces,*

$$z = \sum_{i=1}^r v_{1,i} \otimes v_{2,i} \otimes \cdots \otimes v_{m,i} \in V_1 \otimes V_2 \otimes \cdots \otimes V_m, \tag{1}$$

and  $k_1 + k_2 + \cdots + k_m \geq 2r + m - 1$ , where  $k_j := k(\{v_{j,1}, v_{j,2}, \dots, v_{j,r}\})$  for  $j = 1, 2, \dots, m$ . Then the decomposition (1) is unique in the following sense: If

$$z = \sum_{i=1}^q w_{1,i} \otimes w_{2,i} \otimes \cdots \otimes w_{m,i}$$

and  $q \leq r$ , then we have  $q = r$  and

$$\{v_{1,i} \otimes v_{2,i} \otimes \cdots \otimes v_{m,i} \mid 1 \leq i \leq r\} = \{w_{1,i} \otimes w_{2,i} \otimes \cdots \otimes w_{m,i} \mid 1 \leq i \leq r\}.$$

*E-mail address:* [hderksen@umich.edu](mailto:hderksen@umich.edu)

<sup>1</sup> The author was partially supported by NSF grant DMS 0901298.

It was shown in [5] that the case  $m > 3$  easily follows from the case  $m = 3$ . Many easier and shorter proof of this result have been given, see for example [6,4,3]. Kruskal’s theorem is usually formulated for  $\mathbb{F} = \mathbb{C}$ , but for the proofs in [3,4] are valid in arbitrary fields.

We will show that Kruskal’s theorem is sharp: The theorem is no longer true if  $k_1 + k_2 + \dots + k_m = 2r + m - 2$ .

**Theorem 2.** *Suppose that  $\mathbb{F}$  is a field with more than  $s$  elements, or that  $\mathbb{F}$  is a finite field of characteristic  $\geq s$ . If  $k_1, k_2, \dots, k_m$  are positive integers with  $k_1 + k_2 + \dots + k_m = s + m - 2$ , then there exist  $\mathbb{F}$ -vector spaces  $V_1, V_2, \dots, V_m$ , a positive integer  $q$  and vectors  $\{v_{i,j}\}$  such that  $q \leq s$ ,*

$$0 = \sum_{i=1}^q v_{1,i} \otimes v_{2,i} \otimes \dots \otimes v_{m,i} \tag{2}$$

and

$$k(\{v_{j,1}, v_{j,2}, \dots, v_{j,q}\}) \geq k_j$$

for all  $j$ .

**Corollary 3.** *Suppose that  $r \geq 2$  and that  $\mathbb{F}$  is a field with more than  $2r$  elements. Then Theorem 1 is no longer true if we replace the condition  $k_1 + \dots + k_m \geq 2r + m - 1$  by  $k_1 + \dots + k_m \geq 2r + m - 2$ .*

**Proof.** We can apply Theorem 2 for  $s = 2r$ . By Theorem 1, we get  $2r + m - 2 = k_1 + k_2 + \dots + k_m \leq 2q + m - 2$ , so  $r \leq q \leq 2r$ . Then we have

$$\sum_{i=1}^r v_{1,i} \otimes v_{2,i} \otimes \dots \otimes v_{m,i} = \sum_{i=r+1}^q (-v_{1,i}) \otimes v_{2,i} \otimes \dots \otimes v_{m,i}.$$

Since  $k_1 + \dots + k_m > m$ , we have  $k_j \geq 2$  for some  $j$ . The vectors  $v_{j,1}, v_{j,2}, \dots, v_{j,q}$  are pairwise linearly independent. It follows that the tensors

$$v_{1,i} \otimes v_{2,i} \otimes \dots \otimes v_{m,i}, \quad 1 \leq i \leq q$$

are pairwise linearly independent.  $\square$

Suppose that  $\lambda_1, \lambda_2, \dots, \lambda_s \in \mathbb{F}$ . Then we define a Vandermonde matrix

$$V_k(\lambda_1, \lambda_2, \dots, \lambda_s) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_s \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_s^2 \\ \vdots & \vdots & & \vdots \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \dots & \lambda_s^{k-1} \end{pmatrix}.$$

The following is a well-known property of the Vandermonde matrix:

**Lemma 4.** *If  $\lambda_1, \lambda_2, \dots, \lambda_s$  are distinct, then every  $k$  columns of  $V_k(\lambda_1, \lambda_2, \dots, \lambda_s)$  are linearly independent.*

**Proposition 5.** *Suppose that  $\mathbb{F}$  is a field with more than  $s$  elements. Choose  $\lambda \in \mathbb{F} \setminus \{0\}$  whose multiplicative order is at least  $s$ . Define*

$$v_{j,i} := \begin{pmatrix} 1 \\ \lambda^i \\ \lambda^{2i} \\ \vdots \\ \lambda^{(k_j-1)i} \end{pmatrix}$$

for  $1 \leq j \leq m$  and  $0 \leq i < s$ . Define  $p(x) \in \mathbb{F}[x]$ ,  $p_0, \dots, p_{s-1} \in \mathbb{F}$  by

$$p(x) = (x - 1)(x - \lambda) \cdots (x - \lambda^{s-2}) = p_0 + p_1x + \cdots + p_{s-1}x^{s-1}.$$

Then we have

$$\sum_{i=0}^{s-1} p_i(v_{1,i} \otimes v_{2,i} \otimes \cdots \otimes v_{m,i}) = 0, \tag{3}$$

and

$$k(\{v_{j,0}, v_{j,1}, \dots, v_{j,s-1}\}) = k_j$$

for  $j = 1, 2, \dots, m$ .

**Proof.** For every  $k$  choose the basis

$$e_0 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_1 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_{k-1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

in  $\mathbb{F}^k$  and  $f_0, f_1, \dots, f_{k-1}$  be a dual basis. Suppose that  $0 \leq t_j \leq k_j - 1$  for all  $j$ . An element  $f_{t_1} \otimes f_{t_2} \otimes \cdots \otimes f_{t_m}$  can be viewed as linear function on  $\mathbb{F}^k \otimes \mathbb{F}^k \otimes \cdots \otimes \mathbb{F}^k$  via

$$\begin{aligned} f_{t_1} \otimes f_{t_2} \otimes \cdots \otimes f_{t_m}(e_{u_1} \otimes e_{u_2} \otimes \cdots \otimes e_{u_m}) &= f_{t_1}(e_{u_1})f_{t_2}(e_{u_2}) \cdots f_{t_m}(e_{u_m}) \\ &= \begin{cases} 1 & \text{if } t_i = u_i \text{ for } i = 1, 2, \dots, m; \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

We have

$$\begin{aligned} (f_{t_1} \otimes f_{t_2} \otimes \cdots \otimes f_{t_m}) \left( \sum_{i=0}^{s-1} p_i(v_{1,i} \otimes v_{2,i} \otimes \cdots \otimes v_{m,i}) \right) \\ = \sum_{i=0}^{s-1} p_i \lambda^{(t_1+t_2+\cdots+t_m)i} = p(\lambda^{t_1+t_2+\cdots+t_m}) = 0 \end{aligned}$$

because  $t_1 + t_2 + \cdots + t_m \leq k_1 + k_2 + \cdots + k_m - m \leq s - 2$  and  $\lambda^{t_1+t_2+\cdots+t_m}$  is a root of  $p(x)$ .

The vectors  $v_{j,0}, v_{j,1}, \dots, v_{j,s-1}$  are the columns of  $V_{k_j}(1, \lambda, \lambda^2, \dots, \lambda^{s-1})$ . Since  $1, \lambda, \dots, \lambda^{s-1}$  are distinct, we have

$$k(\{v_{j,0}, v_{j,1}, \dots, v_{j,s-1}\}) = k_j. \quad \square$$

We will need the following well-known combinatorial identity:

**Lemma 6.** If  $0 \leq k \leq n - 1$ , then we have

$$\sum_{i=0}^n (-1)^i \binom{n}{i} i^k = 0$$

**Proof.** Define a  $\mathbb{Q}$ -linear operator  $S : \mathbb{Q}[x] \rightarrow \mathbb{Q}[x]$  by  $S(p(x)) = p(x + 1)$  and let  $I : \mathbb{Q}[x] \rightarrow \mathbb{Q}[x]$  be the identity operator. If  $p(x)$  is a polynomial of degree  $k$ , then  $(S - I)(p(x))$  is a polynomial of degree  $\leq k - 1$ . In particular, we have

$$0 = (S - I)^n p(x) = \sum_{i=0}^n (-1)^i \binom{n}{i} p(x + i) = 0.$$

Now the lemma follows from the case  $p(x) = x^k$ , after substituting  $x = 0$ .  $\square$

**Proposition 7.** Suppose that  $\mathbb{F}$  is a field of characteristic 0 or characteristic at least  $s$ . Define

$$v_{j,i} = \begin{pmatrix} 1 \\ i \\ i^2 \\ \vdots \\ i^{k_j-1} \end{pmatrix}$$

for  $0 \leq i \leq s - 1$  and  $1 \leq j \leq m$ . Then we have

$$\sum_{i=0}^{s-1} (-1)^i \binom{s-1}{i} v_{1,i} \otimes v_{2,i} \otimes \cdots \otimes v_{m,i} = 0,$$

and

$$k(\{v_{j,1}, v_{j,2}, \dots, v_{j,s}\}) = k_j$$

for  $j = 1, 2, \dots, m$ .

**Proof.** Suppose that  $0 \leq t_j \leq k_j - 1$  for all  $j$ . Then we have

$$(f_{t_1} \otimes f_{t_2} \otimes \cdots \otimes f_{t_m}) \left( \sum_{i=0}^{s-1} (-1)^i \binom{s-1}{i} (v_{1,i} \otimes v_{2,i} \otimes \cdots \otimes v_{m,i}) \right) = \sum_{i=0}^{s-1} (-1)^i \binom{s-1}{i} i^{t_1+t_2+\cdots+t_m} = 0$$

because of Lemma 6 and the inequality

$$t_1 + t_2 + \cdots + t_m \leq (k_1 - 1) + (k_2 - 1) + \cdots + (k_m - 1) \leq s - 2.$$

The vectors  $v_{j,0}, v_{j,1}, \dots, v_{j,s-1}$  are the columns of the matrix  $V_{k_j}(0, 1, 2, \dots, s - 1)$ . Since  $0, 1, \dots, s - 1 \in \mathbb{F}$  are distinct, we have

$$k(\{v_{j,0}, v_{j,1}, \dots, v_{j,s-1}\}) = k_j. \quad \square$$

**Proof of Theorem 2.** The theorem follows from Proposition 5 and Proposition 7. Note that in Proposition 5 we can replace  $v_{1,i}$  by  $p_i v_{1,i}$  so that (3) becomes (2) after relabelling the  $v_{j,i}$ 's. For some  $i$  we may have  $p_i = 0$ . In that case  $q$  will be strictly less than  $s$ . Clearly we have  $p_{s-1} = 1$ , so  $q \geq 1$ . In fact,  $q$  is more than  $\max\{k_1, \dots, k_m\}$  because otherwise, the vectors on the right-hand side of (2) would be linearly independent.  $\square$

**Example 8.** The construction in Proposition 7 for  $\mathbb{F} = \mathbb{C}$ ,  $r = 3$ ,  $s = 2r = 6$ ,  $m = 3$ ,  $k_1 = k_2 = 2$ ,  $k_3 = 3$  yields:

$$\begin{aligned} & \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - 5 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 10 \begin{pmatrix} 1 \\ 2 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 2 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} \\ &= 10 \begin{pmatrix} 1 \\ 3 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 3 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 3 \\ 9 \end{pmatrix} - 5 \begin{pmatrix} 1 \\ 4 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 4 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 4 \\ 16 \end{pmatrix} + \begin{pmatrix} 1 \\ 5 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 5 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 5 \\ 25 \end{pmatrix}. \end{aligned}$$

**Example 9.** We use the construction in Proposition 5 for  $\mathbb{F} = \mathbb{R}$ ,  $r = 3$ ,  $s = 2r = 6$ ,  $m = 3$ ,  $k_1 = k_2 = 2$ ,  $k_3 = 3$ ,  $\lambda = 2$ . We expand

$$(x - 1)(x - 2)(x - 4)(x - 8)(x - 16) = -1024 + 1984x - 1240x^2 + 310x^3 - 31x^4 + x^5.$$

We have

$$\begin{aligned} & 1024 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 1240 \begin{pmatrix} 1 \\ 4 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 4 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 4 \\ 16 \end{pmatrix} + 31 \begin{pmatrix} 1 \\ 16 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 16 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 16 \\ 256 \end{pmatrix} \\ &= 1984 \begin{pmatrix} 1 \\ 2 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 2 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} + 310 \begin{pmatrix} 1 \\ 8 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 8 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 8 \\ 64 \end{pmatrix} + \begin{pmatrix} 1 \\ 32 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 32 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 32 \\ 1024 \end{pmatrix}. \end{aligned}$$

Note that in this example the tensors have nonnegative entries when they are viewed as multi-arrays. Whenever one chooses  $\lambda > 0$  in Proposition 5 one obtains counterexamples with nonnegative entries, because exactly half of the coefficients of  $p(x)$  are positive and half of them are negative when  $s = 2r$  is even.

## Acknowledgement

The ideas for this paper were inspired by the AIM workshop *Geometry and representation theory of tensors for computer science, statistics and other areas*, July 21–25, 2008. The author would like to thank AIM, the organizers of the workshop, and Lek-Heng Lim in particular for encouragement for writing this paper. The author also thanks the anonymous referees for useful comments and suggestions.

## References

- [1] J.B. Kruskal, Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics, *Linear Algebra Appl.* 18 (2) (1977) 95–138.
- [2] J.B. Kruskal, Rank, decomposition, and uniqueness for 3-way and  $N$ -way arrays, in: *Multway Data Analysis* (Rome, 1988), North-Holland, Amsterdam, 1989, pp. 7–18.
- [3] J.M. Landsberg, Kruskal's theorem <arXiv:0902.0543>.
- [4] J.A. Rhodes, A concise proof of Kruskal's theorem on tensor decomposition, *Linear Algebra Appl.* 432 (7) (2010) 1818–1824.
- [5] N.D. Sidiropoulos, R. Bro, On the uniqueness of multilinear decomposition of  $N$ -way arrays, *J. Chemometrics* 14 (2000) 229–239.
- [6] A. Stegeman, N.D. Sidiropoulos, On Kruskal's uniqueness condition for the Candecomp/Parafac decomposition, *Linear Algebra Appl.* 420 (2007) 540–552.