Kruskal's uniqueness condition for Candecomp/Parafac

Alwin Stegeman
University of Groningen
The Netherlands

Nikos Sidiropoulos
Technical University of Crete
Greece

Candecomp/Parafac (CP)

- $\underline{\mathbf{X}}$ is a real-valued $I \times J \times K$ array with slices $\mathbf{X}_k$

- The CP model of $\underline{\mathbf{X}}$ with R factors is

$$\mathbf{X}_k = \mathbf{A}\ \mathbf{C}_k\ \mathbf{B}^{\mathrm{T}} + \mathbf{E}_k \qquad k = 1, \ldots, K$$

- Component matrices $\mathbf{A}$ $(I \times R)$, $\mathbf{B}$ $(J \times R)$ and $\mathbf{C}$ $(K \times R)$ with diagonals of $\mathbf{C}_k$ as rows

- CP is also written as $\underline{\mathbf{X}} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r + \underline{\mathbf{E}}$

Uniqueness in CP

- Uniqueness is studied for a fixed residual array
  ⇔ fixed fitted model array

- A CP solution can only be unique up to rescaling/counterscaling and jointly permuting columns of **A**, **B** and **C** (*essential uniqueness*)

- Kruskal's condition (1977) for essential uniqueness:

$$2R + 2 \leq k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}}$$

- k-rank of **A** = max number *k* such that every set of *k* columns of **A** is linearly independent

Kruskal's Uniqueness Theorem

Let (**A**,**B**,**C**) and (**D**,**E**,**F**) be two full CP decompositions of array $\underline{\mathbf{X}}$, both with R components. If

$$2R + 2 \leq k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \,, \qquad \text{(K)}$$

then there exists a unique permutation matrix $\mathbf{\Pi}$ and unique diagonal matrices $\mathbf{\Lambda}_a$, $\mathbf{\Lambda}_b$, $\mathbf{\Lambda}_c$ such that

$$\mathbf{D} = \mathbf{A}\,\mathbf{\Pi}\,\mathbf{\Lambda}_a \qquad \mathbf{E} = \mathbf{B}\,\mathbf{\Pi}\,\mathbf{\Lambda}_b \qquad \mathbf{F} = \mathbf{C}\,\mathbf{\Pi}\,\mathbf{\Lambda}_c$$

and $\quad \mathbf{\Lambda}_a\,\mathbf{\Lambda}_b\,\mathbf{\Lambda}_c = \mathbf{I}_R$

Kruskal's Permutation Lemma

Let $\mathbf{C}$ and $\mathbf{F}$ be $K \times R$ matrices and let $k_{\mathbf{C}} \geq 2$.

Suppose the following condition holds:

If a vector $\mathbf{y}$ is orthogonal to $h \geq \text{rank}(\mathbf{F}) - 1$ columns of $\mathbf{F}$, then $\mathbf{y}$ is orthogonal to at least h columns of $\mathbf{C}$.

Then there exists a unique permutation matrix $\mathbf{\Pi}$ and a unique diagonal matrix $\mathbf{\Lambda}$ such that

$$\mathbf{F} = \mathbf{C}\,\mathbf{\Pi}\,\mathbf{\Lambda}$$

Proof of Kruskal's Uniqueness Theorem

Two CP solutions (**A**,**B**,**C**) and (**D**,**E**,**F**), and (K) holds.

Step 1 (K) ➔ $k_{\mathbf{A}} \geq 2$ $k_{\mathbf{B}} \geq 2$ $k_{\mathbf{C}} \geq 2$

(K) ➔ $(\mathbf{A} \odot \mathbf{B})$ and $(\mathbf{C} \odot \mathbf{A})$ and $(\mathbf{B} \odot \mathbf{C})$ have full column rank

Step 2 (K) ➔ condition of Permutation Lemma for (**A**,**D**) and (**B**,**E**) and (**C**,**F**)

➔ $\mathbf{D} = \mathbf{A}\,\mathbf{\Pi}_a\,\mathbf{\Lambda}_a$ $\mathbf{E} = \mathbf{B}\,\mathbf{\Pi}_b\,\mathbf{\Lambda}_b$ $\mathbf{F} = \mathbf{C}\,\mathbf{\Pi}_c\,\mathbf{\Lambda}_c$

Step 3 $\mathbf{\Pi}_a = \mathbf{\Pi}_b = \mathbf{\Pi}_c$ and $\mathbf{\Lambda}_a\,\mathbf{\Lambda}_b\,\mathbf{\Lambda}_c = \mathbf{I}_R$

Step 1 $\quad 2R + 2 \leq k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}}$

$$k_{\mathbf{C}} \leq R \quad \text{and} \quad k_{\mathbf{B}} \leq R \qquad \Rightarrow \qquad k_{\mathbf{A}} \geq 2$$

$$\text{rank}(\mathbf{A} \odot \mathbf{B}) \geq k_{(\mathbf{A} \odot \mathbf{B})} \geq \min(k_{\mathbf{A}} + k_{\mathbf{B}} - 1, R) = R$$

Sidiropoulos & Bro (2000), Ten Berge (2000)

Suppose $k_{\mathbf{A}} = 1$ and $\mathbf{a}_1 = 2\,\mathbf{a}_2$

$$\mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1 + \mathbf{a}_2 \circ \mathbf{b}_2 \circ \mathbf{c}_2 =$$

$$2\,\mathbf{a}_2 \circ \mathbf{b}_1 \circ \mathbf{c}_1 + \mathbf{a}_2 \circ \mathbf{b}_2 \circ \mathbf{c}_2 =$$

$$\mathbf{a}_2 \circ 2\,\mathbf{b}_1 \circ (\mathbf{c}_1 - \mathbf{c}_2) + \mathbf{a}_2 \circ (2\,\mathbf{b}_1 + \mathbf{b}_2) \circ \mathbf{c}_2$$

➔ $k_{\mathbf{A}} \geq 2$ $k_{\mathbf{B}} \geq 2$ $k_{\mathbf{C}} \geq 2$ is necessary for uniqueness

Suppose rank($\mathbf{A} \odot \mathbf{B}$) < R and ($\mathbf{A} \odot \mathbf{B}$) $\mathbf{n} = \mathbf{0}$

$\mathbf{X}^{(JI \times K)} = (\mathbf{A} \odot \mathbf{B})\, \mathbf{C}^{\mathsf{T}} = (\mathbf{A} \odot \mathbf{B})\, (\mathbf{C} + \mathbf{z}\mathbf{n}^{\mathsf{T}})^{\mathsf{T}}$

for any vector $\mathbf{z}$

$\mathbf{z}$ can be chosen such that a column of $\mathbf{C} + \mathbf{z}\mathbf{n}^{\mathsf{T}}$

becomes $\mathbf{0}$ ➔ $\underline{\mathbf{X}}$ satisfies CP with $R - 1$ factors

➔ rank($\mathbf{A} \odot \mathbf{B}$) = rank($\mathbf{C} \odot \mathbf{A}$) = rank($\mathbf{B} \odot \mathbf{C}$) = R

is necessary for uniqueness

Step 2 $\mathbf{C}$ and $\mathbf{F}$ are $K \times R$ matrices and $k_{\mathbf{C}} \geq 2$

$q(\mathbf{C})$ = the number of columns of $\mathbf{C}$ not orthogonal to $\mathbf{y}$
$q(\mathbf{F})$ = the number of columns of $\mathbf{F}$ not orthogonal to $\mathbf{y}$

To show: $q(\mathbf{F}) \leq R - \text{rank}(\mathbf{F}) + 1 \Rightarrow q(\mathbf{C}) \leq q(\mathbf{F})$

Proof (Sidiropoulos & Bro, 2000)

Construct upper bound and lower bound for $q(\mathbf{F})$

$$\mathbf{X}^{(JI \times K)}\mathbf{y} = (\mathbf{A} \odot \mathbf{B})\,\mathbf{C}^{\mathrm{T}}\mathbf{y} = (\mathbf{D} \odot \mathbf{E})\,\mathbf{F}^{\mathrm{T}}\mathbf{y}$$

$(\mathbf{A} \odot \mathbf{B})$ has full column rank

Hence $q(\mathbf{F}) = 0 \Rightarrow q(\mathbf{C}) = 0$

$$\text{span}^{\perp}(\mathbf{F}) \subseteq \text{span}^{\perp}(\mathbf{C}) \Rightarrow \text{span}(\mathbf{C}) \subseteq \text{span}(\mathbf{F})$$

$$\Rightarrow \text{rank}(\mathbf{C}) \leq \text{rank}(\mathbf{F})$$

$$\begin{aligned} q(\mathbf{F}) &\leq R - \text{rank}(\mathbf{F}) + 1 \\ &\leq R - \text{rank}(\mathbf{C}) + 1 \\ &\leq R - k_{\mathbf{C}} + 1 \\ &\leq k_{\mathbf{A}} + k_{\mathbf{B}} - (R+1) \end{aligned}$$

$$\sum_{k=1}^{K} y_k \, \mathbf{X}_k = \mathbf{A} \, \mathrm{diag}(\mathbf{C}^\top \mathbf{y}) \, \mathbf{B}^\top = \mathbf{D} \, \mathrm{diag}(\mathbf{F}^\top \mathbf{y}) \, \mathbf{E}^\top$$

$$\begin{aligned}
q(\mathbf{F}) &= \mathrm{rank}(\mathrm{diag}(\mathbf{F}^\top \mathbf{y})) \\
&\geq \mathrm{rank}(\mathbf{D} \, \mathrm{diag}(\mathbf{F}^\top \mathbf{y}) \, \mathbf{E}^\top) \\
&= \mathrm{rank}(\mathbf{A} \, \mathrm{diag}(\mathbf{C}^\top \mathbf{y}) \, \mathbf{B}^\top) \\
&= \mathrm{rank}(\mathbf{A}^* \, \mathrm{diag}(\mathbf{t}) \, \mathbf{B}^{*\top}) \\
&\geq \mathrm{rank}(\mathbf{A}^*) + \mathrm{rank}(\mathbf{B}^* \, \mathrm{diag}(\mathbf{t})) - q(\mathbf{C}) \\
&= \mathrm{rank}(\mathbf{A}^*) + \mathrm{rank}(\mathbf{B}^*) - q(\mathbf{C})
\end{aligned}$$

$$\text{rank}(\mathbf{A}^*) \geq \min(q(\mathbf{C}), k_{\mathbf{A}})$$

$$\text{rank}(\mathbf{B}^*) \geq \min(q(\mathbf{C}), k_{\mathbf{B}})$$

$$q(\mathbf{F}) \geq \min(q(\mathbf{C}), k_{\mathbf{A}}) + \min(q(\mathbf{C}), k_{\mathbf{B}}) - q(\mathbf{C}) \qquad (2)$$

$$k_{\mathbf{A}} + k_{\mathbf{B}} - (R+1) \geq q(\mathbf{F}) \qquad (1)$$

(1) and (2) ➔ $\min(q(\mathbf{C}), k_{\mathbf{A}}) = \min(q(\mathbf{C}), k_{\mathbf{B}}) = q(\mathbf{C})$

(2) ➔ $q(\mathbf{C}) \leq q(\mathbf{F})$

Step 3 $\mathbf{D} = \mathbf{A}\,\mathbf{\Pi}_a\,\mathbf{\Lambda}_a$ $\mathbf{E} = \mathbf{B}\,\mathbf{\Pi}_b\,\mathbf{\Lambda}_b$ $\mathbf{F} = \mathbf{C}\,\mathbf{\Pi}_c\,\mathbf{\Lambda}_c$

to show: $\mathbf{\Pi}_a = \mathbf{\Pi}_b = \mathbf{\Pi}_c$ and $\mathbf{\Lambda}_a\,\mathbf{\Lambda}_b\,\mathbf{\Lambda}_c = \mathbf{I}_R$

If $\mathbf{\Pi}_a = \mathbf{\Pi}_b$ then we are done.

Proof (Stegeman & Sidiropoulos, 2005)

$$\begin{aligned}
\mathbf{X}^{(JI \times K)} &= (\mathbf{A} \odot \mathbf{B})\,\mathbf{C}^T \\
&= (\mathbf{A}\,\mathbf{\Pi}\,\mathbf{\Lambda}_a \odot \mathbf{B}\,\mathbf{\Pi}\,\mathbf{\Lambda}_b)\,(\mathbf{C}\,\mathbf{\Pi}_c\,\mathbf{\Lambda}_c)^T \\
&= (\mathbf{A} \odot \mathbf{B})\,(\mathbf{C}\,\mathbf{\Pi}_c\,\mathbf{\Lambda}_a\,\mathbf{\Lambda}_b\,\mathbf{\Lambda}_c\,\mathbf{\Pi}^T)^T
\end{aligned}$$

$(\mathbf{A} \odot \mathbf{B})$ full column rank ➔ $\mathbf{C} = \mathbf{C}\,\mathbf{\Pi}_c\,\mathbf{\Lambda}_a\,\mathbf{\Lambda}_b\,\mathbf{\Lambda}_c\,\mathbf{\Pi}^T$

$k_{\mathbf{C}} \geq 2$ ➔ $\mathbf{\Pi}_c = \mathbf{\Pi}$ and $\mathbf{\Lambda}_a\,\mathbf{\Lambda}_b\,\mathbf{\Lambda}_c = \mathbf{I}_R$

To show: if (K) holds and

$$\mathbf{D} = \mathbf{A}\,\mathbf{\Pi}_a\,\mathbf{\Lambda}_a \qquad \mathbf{E} = \mathbf{B}\,\mathbf{\Pi}_b\,\mathbf{\Lambda}_b \qquad \mathbf{F} = \mathbf{C}\,\mathbf{\Pi}_c\,\mathbf{\Lambda}_c\,,$$

then $\mathbf{\Pi}_a = \mathbf{\Pi}_b$

Proof Stegeman & Sidiropoulos (2005)
Kruskal (1977)

This completes the proof of

Kruskal's Uniqueness Theorem !!

Proof of Kruskal's Permutation Lemma

$\mathbf{C}$ and $\mathbf{F}$ are $K \times R$ matrices and $k_{\mathbf{C}} \geq 2$

For any vector $\mathbf{y}$

$$q(\mathbf{F}) \leq R - \text{rank}(\mathbf{F}) + 1 \quad \Rightarrow \quad q(\mathbf{C}) \leq q(\mathbf{F})$$

$q(\mathbf{C})$ = the number of columns of $\mathbf{C}$ not orthogonal to $\mathbf{y}$
$q(\mathbf{F})$ = the number of columns of $\mathbf{F}$ not orthogonal to $\mathbf{y}$

To show: $\mathbf{F} = \mathbf{C}\,\boldsymbol{\Pi}\,\boldsymbol{\Lambda}$

<u>Proof</u> $q(\mathbf{F}) = 0 \;\Rightarrow\; q(\mathbf{C}) = 0$

$\Rightarrow$ span(**C**) $\subseteq$ span(**F**)

$\Rightarrow$ rank(**F**) $\geq$ rank(**C**) $\geq$ $k_{\mathbf{C}}$ $\geq$ 2

Partition the columns of **F** into the sets

G_0 = { the all-zero columns of **F** }

G_m = { a column **f** of **F** and all nonzero columns of **F** which are proportional to **f** }

$m = 1, \ldots, M$

Definition a subset H_k of columns of **F** is called a k-dimensional column set if

(i) $\text{rank}(H_k) = k$

(ii) H_k contains all columns of **F** in $\text{span}(H_k)$

$H_0 = G_0$ $\quad$ $H_1 = G_0 \cup G_m$ $\quad$ $H_{\text{rank}(\mathbf{F})} = \mathbf{F}$

$\mathbf{y} \perp \mathbf{f}$ and $\mathbf{g}$ $\quad \Rightarrow \quad$ $\mathbf{y} \perp \text{span}(\mathbf{f}, \mathbf{g}) = \text{span}(H_2)$

$\text{rank}(\mathbf{f}, \mathbf{g}) = 2$

together with $\text{span}(\mathbf{C}) \subseteq \text{span}(\mathbf{F})$, this yields the following result

Lemma Under conditions of Permutation Lemma

columns **C** in span(H_k) ≥ # columns **F** in span(H_k)

$k = 0, 1, \ldots, \text{rank}(\mathbf{F})$.

Proof Stegeman & Sidiropoulos (2005)
Kruskal (1977)

$k = 0$: $k_{\mathbf{C}} \geq 2$ ➔ $H_0 = G_0$ is empty
$k = 1$:

columns **C** in span(G_m) ≥ # columns **F** in span(G_m)

$m = 1, \ldots, M$

$k_{\mathbf{C}} \geq 2$ ➔ # columns **C** in span(G_m) $\leq$ 1

1 $\geq$ # columns **C** in span(G_m) $\geq$
columns **F** in span(G_m) $\geq$ 1

hence: # columns **C** in span(G_m) = 1
columns **F** in span(G_m) = 1

➔ every column of **F** has
a proportional column in **C**

➔ $\mathbf{F} = \mathbf{C}\,\mathbf{\Pi}\,\mathbf{\Lambda}$

permutation matrix $\mathbf{\Pi}$ and diagonal matrix $\mathbf{\Lambda}$
are unique

This completes the proof of

Kruskal's Permutation Lemma !!

proof above: Stegeman & Sidiropoulos (2005)
Kruskal (1977)

alternative proof: Jiang & Sidiropoulos (2004)

References

- Jiang, T. & Sidiropoulos, N.D. (2004). Kruskal's permutation lemma and the identification of Candecomp/Parafac and bilinear models with constant modulus constraints. *IEEE Transactions on Signal Processing*, **52**, 2625-2636.
- Kruskal, J.B. (1977). Three-way arrays: rank and uniqueness of trilinear decompositions, with applications to arithmetic complexity and statistics. *Linear Algebra and its Applications*, **18**, 95-138.
- Sidiropoulos, N.D. & Bro, R. (2000). On the uniqueness of multilinear decomposition of *N*-way arrays. *Journal of Chemometrics*, **14**, 229-239.
- Stegeman, A. & Sidiropoulos, N.D. (2005). On Kruskal's uniqueness condition for the Candecomp/Parafac decomposition. Submitted.
- Ten Berge, J.M.F. (2000). The k-rank of a Khatri-Rao product. *Unpublished Note*, Heymans Institute of Psychological Research, University of Groningen, The Netherlands.