Why do degenerate solutions occur in Candecomp/Parafac ?

Alwin Stegeman
University of Groningen
The Netherlands

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

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

- 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**

- Kruskal's (1977) uniqueness condition:

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

- Avoid scaling indeterminacy: norm the columns of two component matrices to length 1

Degenerate CP solutions

- *Two-factor degeneracy*

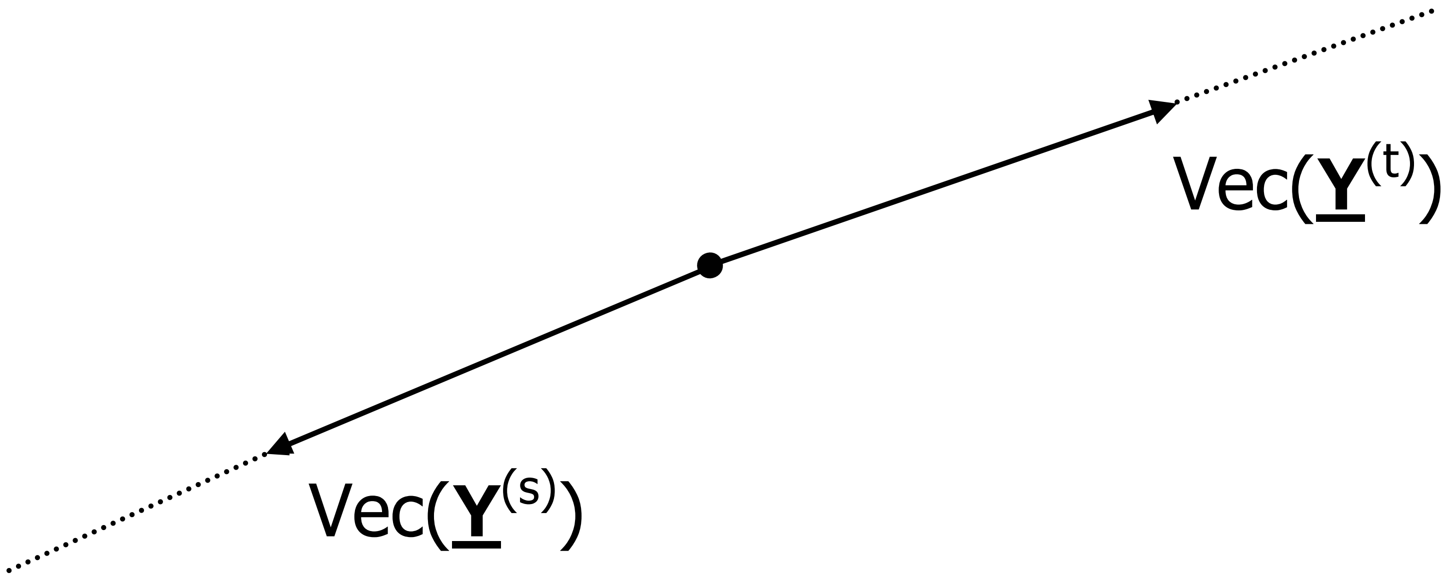
1. $\mathbf{a}_s \approx \pm\, \mathbf{a}_t \qquad \mathbf{b}_s \approx \pm\, \mathbf{b}_t \qquad \mathbf{c}_s \approx \pm\, \mathbf{c}_t$

2. product of three correlations tends to -1

3. magnitudes of $\mathbf{c}_s$ and $\mathbf{c}_t$ become arbitrarily large

Two-factor degeneracy

$\underline{\mathbf{Y}}^{(s)} = \mathbf{a}_s \circ \mathbf{b}_s \circ \mathbf{c}_s$ $\qquad$ $\underline{\mathbf{Y}}^{(t)} = \mathbf{a}_t \circ \mathbf{b}_t \circ \mathbf{c}_t$



$\underline{\mathbf{Y}}^{(s)} + \underline{\mathbf{Y}}^{(t)}$ remains “small” and contributes to a better CP fit

Degenerate CP solutions

- *Three-factor degeneracy*

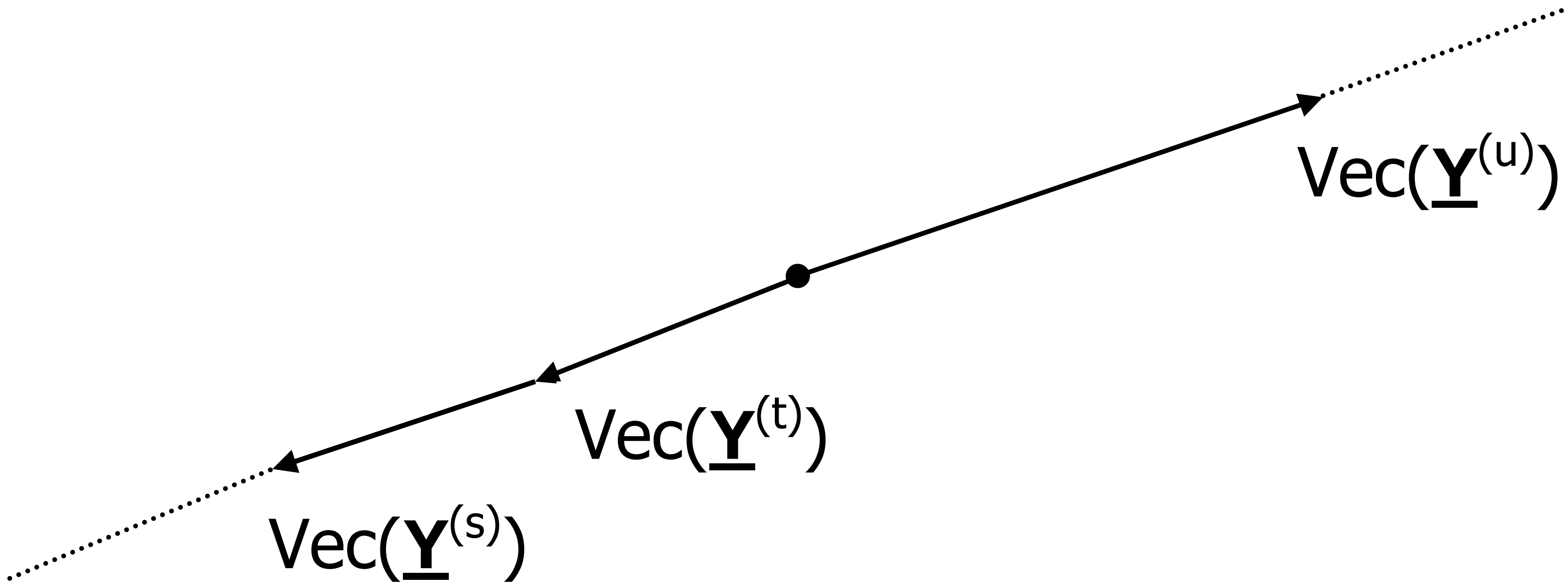
1. $\mathbf{a}_s \approx \pm\, \mathbf{a}_t \approx \pm\, \mathbf{a}_u$ $\qquad$ $\mathbf{b}_s \approx \pm\, \mathbf{b}_t \approx \pm\, \mathbf{b}_u$

2. $\mathbf{c}_s \approx \delta_1\, \mathbf{c}_t \approx \delta_2\, \mathbf{c}_u$ for constants δ_1 and δ_2

3. magnitudes of $\mathbf{c}_s$, $\mathbf{c}_t$ and $\mathbf{c}_u$ become arbitrarily large

4. $\mathbf{c}_s \pm \mathbf{c}_t \pm \mathbf{c}_u$ remains "small"

Three-factor degeneracy



$\underline{\mathbf{Y}}^{(s)} + \underline{\mathbf{Y}}^{(t)} + \underline{\mathbf{Y}}^{(u)}$ remains "small" and contributes to a better CP fit

Example for 3×3×2 with $R = 3$

$$\mathbf{A} = \begin{bmatrix} 0.48 & 0.46 & -0.47 \\ -0.66 & -0.65 & 0.66 \\ -0.57 & -0.60 & 0.58 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 0.72 & -0.69 & 0.71 \\ 0.61 & -0.65 & 0.63 \\ 0.33 & -0.31 & 0.32 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 653 & -625 & 1278 \\ 2162 & -2239 & 4398 \end{bmatrix}$$

factor 1 ≈ factor 2 ≈ − 2⋅factor 3

Idea of Kruskal, Harshman and Lundy (1989)

- "Degenerate CP solutions occur when the CP objective function does not have a minimum but an infimum"

- Ten Berge, Kiers & De Leeuw (1988) proved this for a specific 2×2×2 array and $R = 2$

- Paatero (2000): degenerate sequences of 2×2×2 arrays of rank 2 approximating a rank-3 array

The CP problem

Minimize $\left\| \underline{\mathbf{X}} - \underline{\mathbf{Y}} \right\|^2$

subject to $\underline{\mathbf{Y}} \in D_R = \{ \underline{\mathbf{Y}}$ with rank $\leq R \}$

Lim (2004): the set D_R is not closed for $R \geq 2$

➔ CP may not have a minimum if $\underline{\mathbf{X}}$ lies outside D_R

Consider all $p \times q \times 2$ arrays

- Random $p \times q \times 2$ array $\underline{\mathbf{X}}$ with $p \geq q$
- For which $\{p, q, R\}$ do degenerate solutions occur?
- Typical rank (Ten Berge & Kiers, 1999) of $\underline{\mathbf{X}}$:

$$\begin{array}{ll} \min(p, 2q) & \text{if } p > q \\ \{p, p+1\} & \text{if } p = q \end{array}$$

- If $R \geq \text{rank}(\underline{\mathbf{X}})$, then no degenerate solutions occur

$\underline{\mathbf{X}}$	rank($\underline{\mathbf{X}}$)	R	Degeneracy?
$p = q$	p+1	$R = p$	always
$p = q$	p+1	$R < p$	sometimes
$p = q$	p	$R < p$	sometimes
$p > q$	min(p, 2q)	(p, 2q) $> R > q$	no
$p > q$	min(p, 2q)	$R = q$	sometimes
$p > q$	min(p, 2q)	$R < q$	sometimes

Stegeman (2006 b)

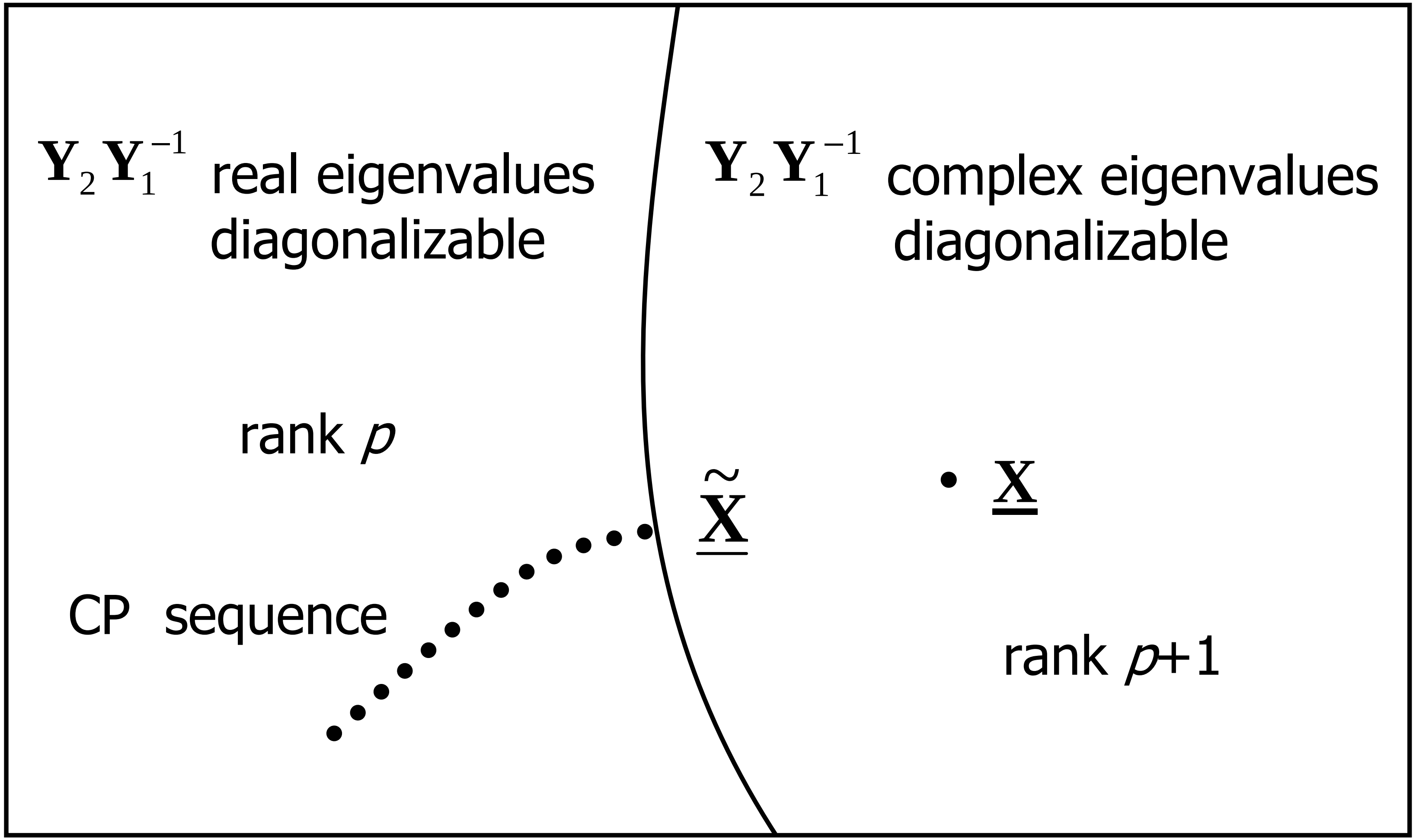
Consider $p \times p \times 2$ arrays with $R = p$

$\mathbf{Y}_2 \mathbf{Y}_1^{-1}$	eigenvalues all real	some complex eigenvalues
diagonalizable	rank($\underline{\mathbf{Y}}$)=p positive volume	rank($\underline{\mathbf{Y}}$)$\geq p+1$ positive volume
not diag.	rank($\underline{\mathbf{Y}}$)$\geq p+1$ zero volume	rank($\underline{\mathbf{Y}}$)$\geq p+1$ zero volume

Ten Berge (1991), Ja′ Ja′ (1979)

Typical rank = $\{p, p+1\}$ (Ten Berge & Kiers, 1999)

$2p^2$ - dimensional space of $p \times p \times 2$ arrays



Boundary arrays

$\mathbf{Y}_2 \mathbf{Y}_1^{-1}$ has real eigenvalues, but not all distinct

$\mathbf{Y}_2 \mathbf{Y}_1^{-1}$ diagonalizable ➔ rank p (type 1)

$\mathbf{Y}_2 \mathbf{Y}_1^{-1}$ not diagonalizable ➔ rank ≥ p+1 (type 2)

dimensionality (type 1) < dimensionality (type 2)

In practice only boundary arrays $\underline{\tilde{\mathbf{X}}}$ of type 2 are encountered

Result for $p \times p \times 2$ arrays with $R = p$

Let random $\underline{\mathbf{X}}$ have rank $p+1$. There holds that:

(I) the CP objective function does not have a minimum, but an infimum, and

(II) any sequence of CP solutions of which the objective value approaches the infimum, will become degenerate.

Stegeman (2006 a)

$\underline{\mathbf{Y}}$ rank p ➔ eigendecomposition $\mathbf{Y}_2\mathbf{Y}_1^{-1} = \mathbf{K}\,\mathbf{\Lambda}\,\mathbf{K}^{-1}$

A rank-p decomposition of $\underline{\mathbf{Y}}$ is:

$$\mathbf{Y}_1 = \mathbf{K}\,\mathbf{I}_p\,(\mathbf{K}^{-1}\,\mathbf{Y}_1)$$

$$\mathbf{Y}_2 = \mathbf{K}\,\mathbf{\Lambda}\,(\mathbf{K}^{-1}\,\mathbf{Y}_1)$$

$$\mathbf{A} = \mathbf{K} \qquad \mathbf{B} = (\mathbf{K}^{-1}\,\mathbf{Y}_1)^T \qquad \mathbf{C} = \begin{bmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_p \end{bmatrix}$$

Kruskal holds: $2p + 2 \leq k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} = p + p + 2$

➔ <u>uniqueness</u> (if all eigenvalues are different)

Sequence of solutions $\underline{\mathbf{Y}}$ converges to $\underline{\tilde{\mathbf{X}}}$

$\mathbf{Y}_2\mathbf{Y}_1^{-1} = \mathbf{K}\ \mathbf{\Lambda}\ \mathbf{K}^{-1}$ converges to $\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_1^{-1}$

When the CP algorithm terminates:

- $\mathbf{A} = \mathbf{K}$ has some columns close to linear dependence

- $\mathbf{B} = (\mathbf{K}^{-1}\ \mathbf{Y}_1)^T$ has large magnitudes in these columns

- $\mathbf{C} = \begin{bmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_p \end{bmatrix}$ has these columns nearly identical

$\underline{\mathbf{X}}$	rank($\underline{\mathbf{X}}$)	R	Degeneracy?	
$p = q$	p+1	$R = p$	always	OK
$p = q$	p+1	$R < p$	sometimes	
$p = q$	p	$R < p$	sometimes	
$p > q$	min(p, 2q)	(p, 2q) > R $R > q$	no	next
$p > q$	min(p, 2q)	$R = q$	sometimes	
$p > q$	min(p, 2q)	$R < q$	sometimes	

Stegeman (2006 b)

Consider $p \times q \times 2$ arrays with $(p, 2q) > R > q$

Let $W_R = \{ \underline{\mathbf{Y}}$ with rank $[\mathbf{Y}_1 | \mathbf{Y}_2] \leq R \}$

$D_R = \{ \underline{\mathbf{Y}}$ with rank$(\underline{\mathbf{Y}}) \leq R \} \subset W_R$

- W_R is a closed set
- rank$(\underline{\mathbf{X}})$ = min$(p,2q)$ ➔ $\underline{\mathbf{X}}$ does not lie in W_R
- approximation of $\underline{\mathbf{X}}$ from W_R always has an optimal solution $\underline{\tilde{\mathbf{X}}}$

$\underline{\mathbf{Y}}$ lies in W_R ⇔ there exists a non-singular **S** with

$$\mathbf{S}\,[\mathbf{Y}_1|\mathbf{Y}_2] = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$$

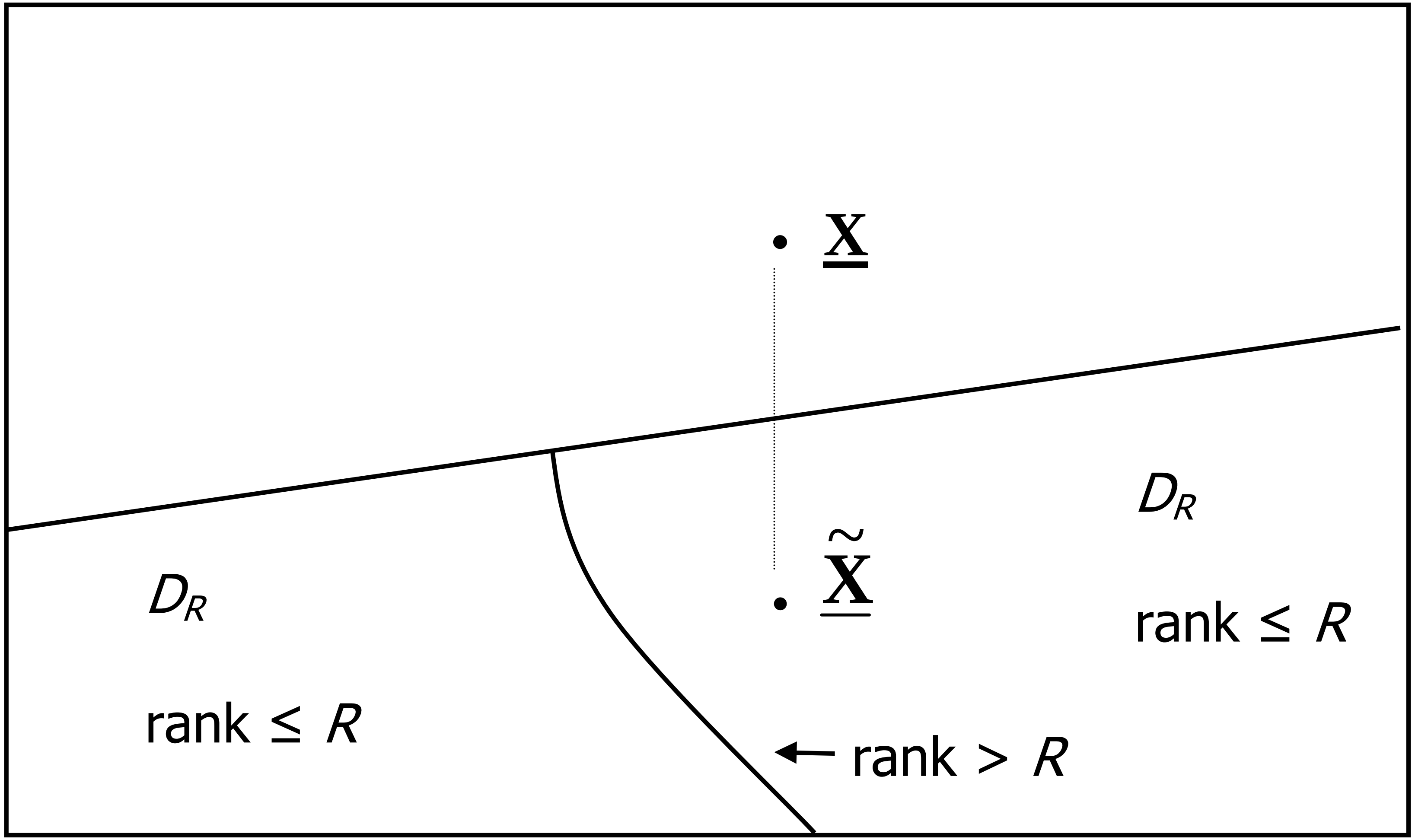
where $\mathbf{H}_i$ are $R \times q$

rank($\underline{\mathbf{Y}}$) = rank($\underline{\mathbf{H}}$) and typical rank $\underline{\mathbf{H}}$ = R

→ dimensionality D_R = dimensionality W_R

→ arrays in W_R of rank higher than R lie in a set of lower dimensionality

$p \times q \times 2$ arrays with $(p,2q) > R > q$



$\underline{\mathbf{X}}$	rank($\underline{\mathbf{X}}$)	R	Degeneracy?	
$p = q$	$p+1$	$R = p$	always	OK
$p = q$	$p+1$	$R < p$	sometimes	
$p = q$	p	$R < p$	sometimes	
$p > q$	min(p, 2q)	(p, 2q) > R $R > q$	no	OK
$p > q$	min(p, 2q)	$R = q$	sometimes	next
$p > q$	min(p, 2q)	$R < q$	sometimes	

Stegeman (2006 b)

Consider $p \times q \times 2$ arrays with $p > R = q$

$W_R = \{ \underline{\mathbf{Y}}$ with rank $[\mathbf{Y}_1|\mathbf{Y}_2] \leq R \}$

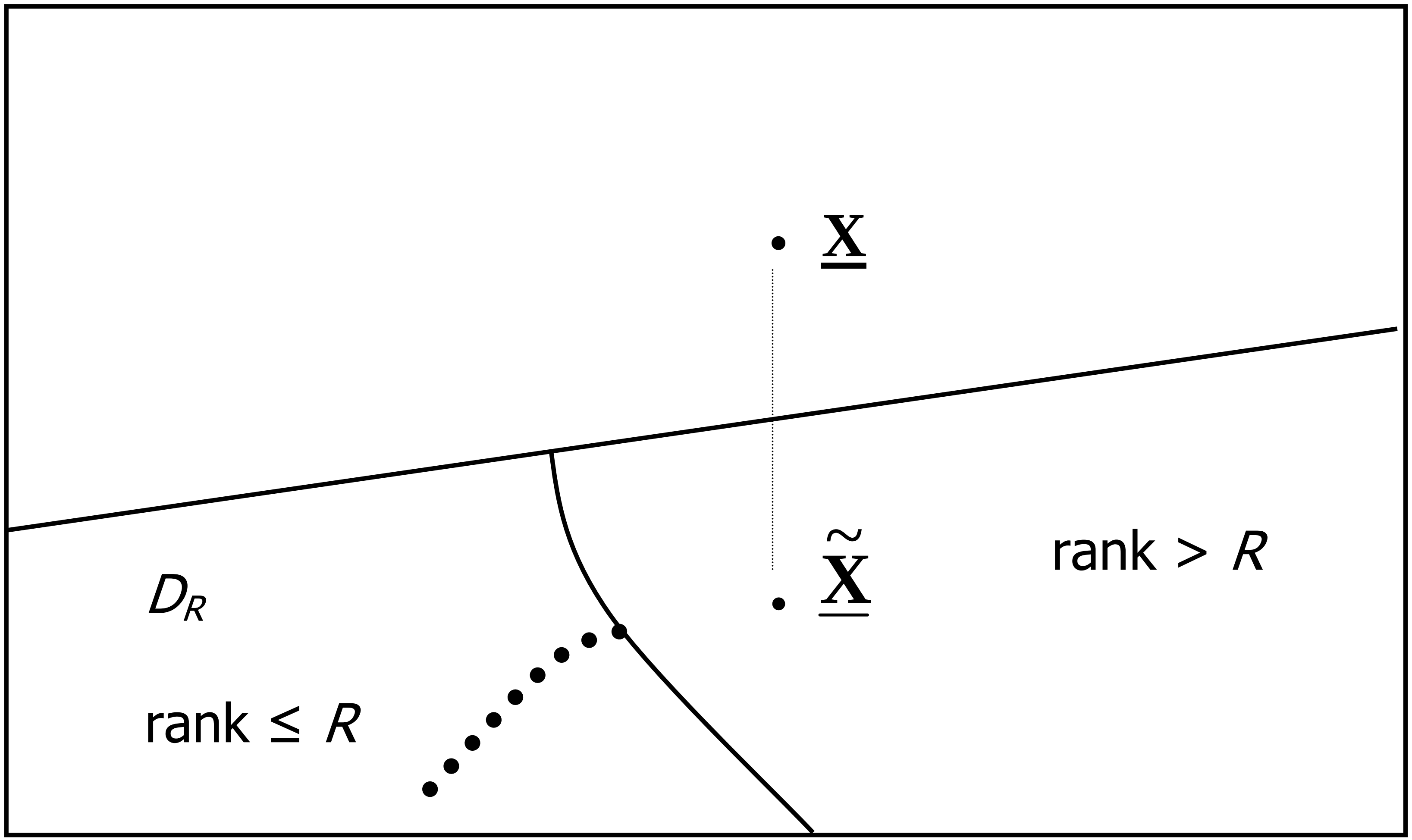
$\underline{\mathbf{Y}}$ lies in W_R ⟺ $\mathbf{S}\,[\mathbf{Y}_1|\mathbf{Y}_2] = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$, with $\mathbf{H}_i$ $q \times q$

rank($\underline{\mathbf{Y}}$) = rank($\underline{\mathbf{H}}$) and typical rank $\underline{\mathbf{H}} = \{q, q+1\}$

➔ dimensionality D_R = dimensionality W_R

➔ arrays in W_R of rank higher than $R = q$ lie in a set of equal dimensionality

$p \times q \times 2$ arrays with $p > R = q$



$\underline{\mathbf{X}}$	rank($\underline{\mathbf{X}}$)	R	Degeneracy?	
$p = q$	p+1	$R = p$	always	OK
$p = q$	p+1	$R < p$	sometimes	next
$p = q$	p	$R < p$	sometimes	next
$p > q$	min(p, 2q)	(p, 2q) > R $R > q$	no	OK
$p > q$	min(p, 2q)	$R = q$	sometimes	OK
$p > q$	min(p, 2q)	$R < q$	sometimes	next

Stegeman (2006 b)

Consider $p \times q \times 2$ arrays with $p \geq q > R$

$$W_R = \{ \underline{\mathbf{Y}} \text{ with rank } [\mathbf{Y}_1 | \mathbf{Y}_2] \leq R \text{ and rank } \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} \leq R \}$$

$$D_R = \{ \underline{\mathbf{Y}} \text{ with rank}(\underline{\mathbf{Y}}) \leq R \} \subset W_R$$

- W_R is a closed set
- $\underline{\mathbf{X}}$ does not lie in W_R
- approximation of $\underline{\mathbf{X}}$ from W_R always has an optimal solution $\underline{\tilde{\mathbf{X}}}$

$\underline{\mathbf{Y}}$ lies in W_R ⇔ $\mathbf{S}\,\mathbf{Y}_i\mathbf{T} = \begin{bmatrix} \mathbf{G}_i & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$, with $\mathbf{G}_i$ $R \times R$

rank($\underline{\mathbf{Y}}$) = rank($\underline{\mathbf{G}}$) and typical rank $\underline{\mathbf{G}}$ = $\{R, R+1\}$

➔ dimensionality D_R = dimensionality W_R

➔ arrays in W_R of rank higher than R lie in a set of equal dimensionality

Conclusions for $p \times q \times 2$ arrays

- the CP objective function may not have a minimum if the boundary of the set D_R consists (partly) of arrays with rank larger than R

- this is “caused by” the two-valued typical rank of $p \times p \times 2$ arrays

- this type of degeneracy does not occur in the complex-valued CP model

- a sequence of degenerate CP solutions is unique at every point

All known arrays with typical rank $\{k,k+1\}$

asymmetric slices		symmetric slices	
$p \times p \times 2$	$\{p,p+1\}$	$p \times p \times 2$	$\{p,p+1\}$
$3\times3\times5$	$\{5,6\}$	$3\times3\times5$	$\{5,6\}$
$8\times4\times3$	$\{8,9\}$	$3\times3\times4$	$\{4,5\}$

Ten Berge & Kiers (1999), Ten Berge (2000,2004)
Ten Berge, Sidiropoulos & Rocci (2004)

Take $\underline{\mathbf{X}}$ random with rank($\underline{\mathbf{X}}$)$=k+1$ and $R=k$

Symmetric slices: CP solution has **A** = **B** (Indscal)

Array size	Typical rank	Degeneracy?	Unique?
$p \times p \times 2$	$\{p, p+1\}$	always	yes
$p \times p \times 2$ (s)	$\{p, p+1\}$	always	yes
$3 \times 3 \times 5$	{5,6}	sometimes	partially
$3 \times 3 \times 5$ (s)	{5,6}	always	no
$3 \times 3 \times 4$ (s)	{4,5}	always	yes
$8 \times 4 \times 3$	{8,9}	always (?)	no

Stegeman (2006 c)

References

- Ja' Ja', J. (1979). Optimal evaluation of pairs of bilinear forms. *SIAM Journal on Computing*, **8**, 443-462.
- 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.
- Kruskal, J.B., Harshman, R.A. & Lundy, M.E. (1989). How 3-MFA data can cause degenerate Parafac solutions, among other relationships. In: *Multiway Data Analysis*, Coppi R. & Bolasco, S. (editors), North-Holland, 115-121.
- Lim, L-H. (2004). What's possible and what's impossible in tensor decompositions \ approximation. Talk at *ARCC Tensor Decompositions Workshop*. Pdf available at: http://csmr.ca.sandia.gov/~tgkolda/tdw2004/
- Paatero, P. (2000). Construction and analysis of degenerate Parafac models. *Journal of Chemometrics*, **14**, 285-299.
- Stegeman, A. (2006 a). Degeneracy in Candecomp/Parafac explained for $p \times p \times 2$ arrays of rank $p+1$ or higher. *Psychometrika*, to appear.
- Stegeman, A. (2006 b). Degeneracy in Candecomp/Parafac explained for random $p \times q \times 2$ arrays. Submitted.
- Stegeman, A. (2006 c). Degeneracy in Candecomp/Parafac and Indscal explained for several three-sliced arrays with a two-valued typical rank. Submitted.
- Ten Berge, J.M.F., Kiers, H.A.L. & De Leeuw, J. (1988). Explicit Candecomp/Parafac solutions for a contrived 2×2×2 array of rank three. *Psychometrika*, **53**, 579-584.
- Ten Berge, J.M.F. (1991). Kruskal's polynomial for 2×2×2 arrays and a generalization to $2 \times n \times n$ arrays. *Psychometrika*, **56**, 631-636.

- Ten Berge, J.M.F. & Kiers, H.A.L. (1999). Simplicity of core arrays in three-way principal component analysis and the typical rank of $p \times q \times 2$ arrays. *Linear Algebra and its Applications*, **294**, 169-179.
- Ten Berge, J.M.F. (2000). The typical rank of tall three-way arrays. *Psychometrika*, **65**, 525-532.
- Ten Berge, J.M.F. (2004). Partial uniqueness in Candecomp/Parafac. *Journal of Chemometrics*, **18**, 12-16.
- Ten Berge, J.M.F., Sidiropoulos, N.D. & Rocci, R. (2004). Typical rank and Indscal dimensionality for symmetric three-way arrays of order $I \times 2 \times 2$ or $I \times 3 \times 3$. *Linear Algebra and its Applications*, **388**, 363-377.