Canonical polyadic decomposition of third-order tensors: some results on uniqueness and algebraic algorithm

Ignat Domanov, Lieven De Lathauwer

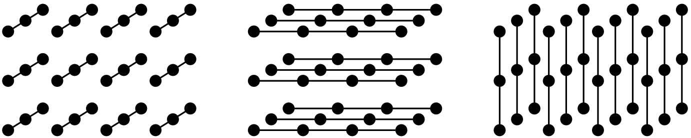
Group Science, Engineering and Technology, KU Leuven–Kulak
Department of Electrical Engineering (ESAT), SCD–STADIUS, KU Leuven

February 4, 2015

Vizualization of $\mathcal{T}$ and rank-1 tensor

$\mathcal{T} \equiv$ an $I \times J \times K$ array of numbers

rank-1 tensors

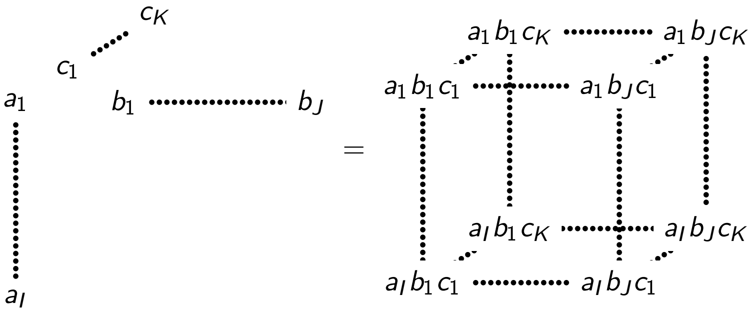


all fibers are proportional & all rows are proportional & all columns are proportional

rank-1 tensor: $\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} \in \mathbb{C}^{I \times J \times K}$, $\mathbf{a}, \mathbf{b}, \mathbf{c}$ - nonzero vectors

Decompositions into rank-1 terms: $I \times J \times K$ tensors

rank-1 tensor always equals the outer product of three vectors

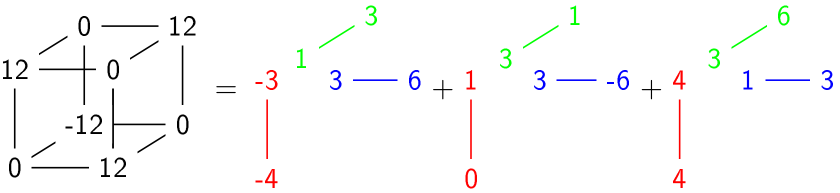


Parafac, CP decomposition,...

Polyadic Decomposition (PD): $\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r$

$$\left.\begin{array}{l}\text{Parafac or} \\ \text{Candecomp or} \\ \text{Candecomp/Parafac or} \\ \text{CP decomposition or} \\ \text{Canonical Polyadic Decomposition (CPD)}\end{array}\right\} \text{a PD with } \underbrace{\text{minimal } R}_{\text{rank of } \mathcal{T}}$$

Vizualization of PD of $\mathcal{T}$

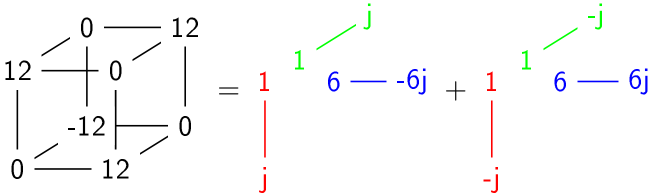


NOTATION:

$$\mathcal{T} = \left[\underbrace{\begin{bmatrix}-3 & 1 & 4\\ -4 & 0 & 4\end{bmatrix}}_{\mathbf{A}}, \underbrace{\begin{bmatrix}3 & 3 & 1\\ 6 & -6 & 3\end{bmatrix}}_{\mathbf{B}}, \underbrace{\begin{bmatrix}1 & 3 & 3\\ 3 & 1 & 6\end{bmatrix}}_{\mathbf{C}}\right]_3 = [\underbrace{\mathbf{A}, \mathbf{B}, \mathbf{C}}_{\text{factor matrices}}]_3$$

Can we use less than 3 terms? What is the rank of $\mathcal{T}$?

Can we use less than 3 terms? What is the rank of $\mathcal{T}$?



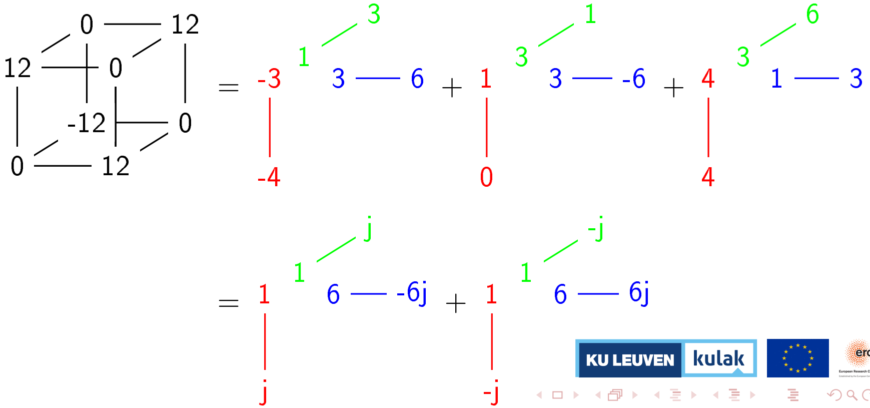
$$= \left[\!\!\left[\begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}, \begin{bmatrix} 6 & 6 \\ -6j & 6j \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix} \right]\!\!\right]_2$$

Uniqueness of the CPD

CPD of $\mathcal{T}$ is unique

$\Updownarrow$

all CPDs of $\mathcal{T}$ have the same rank one terms



Kruskal's results on uniqueness of CPD

Definition: The k-rank of a matrix $\mathbf{A}$

$$k_{\mathbf{A}} = \max\{k : \text{ any } k \text{ columns of } \mathbf{A} \text{ are linearly independent}\}.$$

Example:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \qquad r_{\mathbf{A}} = 2, \qquad k_{\mathbf{A}} = 2$$

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \qquad r_{\mathbf{B}} = 2, \qquad k_{\mathbf{B}} = 1$$

Example: $\mathbf{A}$ is generic $I \times R$ matrix $\Rightarrow \quad k_{\mathbf{A}} = \min(I, R)$

Kruskal's results on uniqueness of CPD

Theorem

Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. *If*

$$\begin{cases} k_{\mathbf{A}} + r_{\mathbf{B}} + r_{\mathbf{C}} & \geq 2R + 2, \\ \min(r_{\mathbf{C}} + k_{\mathbf{B}}, k_{\mathbf{C}} + r_{\mathbf{B}}) & \geq R + 2, \end{cases} \tag{1}$$

or "permute **A**, **B**, **C** *in* (1)*" or*

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2, \tag{2}$$

then $r_{\mathcal{T}} = R$ *and the CPD of* $\mathcal{T}$ *is unique.*

Obviously: $k_{\mathbf{A}} \leq r_{\mathbf{A}}$, $k_{\mathbf{B}} \leq \dots \Rightarrow$ (2) is more restrictive than (1).

J.B. Kruskal. Three-way arrays: Rank and uniqueness of trilinear decompositions. *Lin. Alg. Appl.*, 18(2):95—138, 1977.

Conditions for uniqueness: Lim–Comon

Computation of k-rank is NP hard.
it is known that $\frac{1}{\mu_{\mathbf{A}}} \geq k_{\mathbf{A}}$, where

coherence of **A**: $\mu_{\mathbf{A}} := \max_{1 \leq i < j \leq R} \frac{|\mathbf{a}_i^T \mathbf{a}_j|}{\|\mathbf{a}_i\| \|\mathbf{a}_j\|}$

Idea: replace in (2) the k-ranks by coherences^{-1}

Theorem

Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. If $\frac{1}{\mu_{\mathbf{A}}} + \frac{1}{\mu_{\mathbf{B}}} + \frac{1}{\mu_{\mathbf{C}}} \geq 2R + 2$, then $r_{\mathcal{T}} = R$ and the CPD of tensor $\mathcal{T}$ is unique.

L.-H. Lim and P. Comon. Blind multilinear identification. *IEEE Transactions on Information Theory* 60(2):1260–1280, 2014

We can also replace the k-ranks by coherences^{-1} in (1)

Theorem

Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. *If*

$$\begin{cases} \frac{1}{\mu_{\mathbf{A}}} + r_{\mathbf{B}} + r_{\mathbf{C}} & \geq 2R + 2, \\ \min(r_{\mathbf{C}} + \frac{1}{\mu_{\mathbf{B}}}, \frac{1}{\mu_{\mathbf{C}}} + r_{\mathbf{B}}) & \geq \ R + 2, \end{cases} \tag{3}$$

then $r_{\mathcal{T}} = R$ *and the CPD of* $\mathcal{T}$ *is unique.*

REMARK: "permute **A**, **B**, **C** in (3)" – two other conditions

Problem 1:(Uniqueness)

Given CPD $\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r$. Is the CPD unique?

Problem 2:("Algebraic" algorithm)

Compute the CPD $\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r$ by means of Linear Algebra:

- computation of ranks
- computation of the null space
- eigenvalue decompositions

An $R \times R \times 2$ tensor of rank R

Given PD $\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r \in \mathbb{C}^R \times \mathbb{C}^R \times \mathbb{C}^2$,

that is, $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, in which

$$\begin{cases} \mathbf{A} := [\mathbf{a}_1 \ \ldots \mathbf{a}_R] \text{ and } \mathbf{B} := [\mathbf{b}_1 \ \ldots \mathbf{b}_R], & \text{are } R \times R \text{ matrices,} \\ \mathbf{C} := [\mathbf{c}_1 \ \ldots \ \mathbf{c}_R] := \begin{bmatrix} \mathbf{c}^1 \\ \mathbf{c}^2 \end{bmatrix}, & \text{is an } 2 \times R \text{ matrix.} \end{cases}$$

An $R \times R \times 2$ tensor of rank R

Assume that

- $\mathbf{A} := [\mathbf{a}_1 \ \ldots \mathbf{a}_R]$ and $\mathbf{B} := [\mathbf{b}_1 \ \ldots \mathbf{b}_R]$ are nonsingular and
- $\mathbf{C}$ has no proportional columns.

Then

P1: $rk(\mathcal{T}) = R$ and the CPD of $\mathcal{T}$ is unique,

P2: CPD of $\mathcal{T}$ can be found algebraically.

An $R \times R \times 2$ tensor of rank R: reduction to ED

STEP 1: Rewrite $\mathcal{T} = \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r$ in a matrix form:

$$\mathbf{T}_1 = \mathbf{A}\mathrm{Diag}(\mathbf{c}^1)\mathbf{B}^T, \quad \mathbf{T}_2 = \mathbf{A}\mathrm{Diag}(\mathbf{c}^2)\mathbf{B}^T, \qquad (*)$$

where $\mathbf{T}_1$ and $\mathbf{T}_2$ denote the frontal slices of $\mathcal{T}$.

STEP 2: eliminate $\mathbf{A}$ and $\mathbf{B}$ in $(*)$

$$\mathbf{A}\mathrm{Diag}(\mathbf{d})\mathbf{A}^{-1} = \mathbf{T}_1\mathbf{T}_2^{-1}, \qquad \mathbf{B}\mathrm{Diag}(\mathbf{d})\mathbf{B}^{-1} = (\mathbf{T}_2^{-1}\mathbf{T}_1)^T,$$

where $\mathrm{Diag}(\mathbf{d}) := \mathrm{Diag}(\mathbf{c}^1)\mathrm{Diag}(\mathbf{c}^2)^{-1}$ has distinct diagonal entries.

STEP 3: ,

$\mathbf{a}_1, \dots, \mathbf{a}_R$ are the eigenvectors of $\mathbf{T}_1\mathbf{T}_2^{-1}$

$\mathbf{b}_1, \dots, \mathbf{b}_R$ are the eigenvectors of $(\mathbf{T}_2^{-1}\mathbf{T}_1)^T$

$\mathbf{c}^1, \mathbf{c}^2$ are easily found from $(*)$.

Definition

Let $\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B})$ be a matrix whose columns are

$$\Lambda^m(\mathbf{a}_{i_1} \otimes \cdots \otimes \mathbf{a}_{i_m}) \otimes \Lambda^m(\mathbf{b}_{i_1} \otimes \cdots \otimes \mathbf{b}_{i_m}), \quad 1 \leq i_1 < i_2 < \cdots < i_m \leq R.$$

Theorem (ID&LDL,2014)

Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$, $k_\mathbf{C} = r_\mathbf{C}$, $m = R - r_\mathbf{C} + 2$ *and let*

$$\mathcal{C}_m(\mathbf{A}) \odot \mathcal{C}_m(\mathbf{B}) \text{ have full column rank.}$$

Then

(i) $r_\mathcal{T} = R$ *and the CPD of* $\mathcal{T}$ *is unique;*

(ii) *the CPD of* $\mathcal{T}$ *can be found algebraically.*

I. Domanov and L. De Lathauwer. Canonical polyadic decomposition of third-order tensors: reduction to generalized eigenvalue decomposition. *SIAM J. Matrix Anal. Appl.*, 35(2):636-660, 2014.

Illustration for $K = R$: uniqueness

We re-explain results from

L. De Lathauwer. A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization. *SIAM J. Matrix Anal. Appl.*, 28(3): 642-666, 2006.

T. Jiang and N.D. Sidiropoulos. Kruskal's permutation lemma and the identification of CANDECOMP/PARAFAC and bilinear models with constant modulus constraints. *IEEE Trans. Signal Process.*, 52(9):2625-2636, 2004.

$K = R \;\Rightarrow m = R - r_{\mathbf{C}} + 2 = 2$

$\mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B})$ has $\frac{R(R-1)}{2}$ columns

$$(\mathbf{a}_i \otimes \mathbf{a}_j - \mathbf{a}_j \otimes \mathbf{a}_i) \otimes (\mathbf{b}_i \otimes \mathbf{b}_j - \mathbf{b}_j \otimes \mathbf{b}_i), \quad 1 \leq i < j \leq R$$

If $\mathcal{C}_2(\mathbf{A}) \odot \mathcal{C}_2(\mathbf{B})$ has f.c.r $\Rightarrow$ CPD is unique

OLD: Illustration for $K = R$: algebraic algorithm

Main construction:

INPUT: $\mathcal{T} \in \mathbb{C}^{I \times J \times K}$
OUTPUT: $\mathbf{R}_2(\mathcal{T}) \in \mathbb{C}^{I^2 J^2 \times K^2}$

STEP 1: $\mathcal{T}^{\otimes^2} \leftarrow \mathcal{T} \otimes \mathcal{T} \in \mathbb{C}^{I \times J \times K \times I \times J \times K}$
STEP 2: $\mathcal{T}_I^{\otimes^2} \leftarrow$ partially skew-symmetrize $\mathcal{T}^{\otimes^2}$ along "I" dim-s
STEP 3: $\mathcal{T}_{IJ}^{\otimes^2} \leftarrow$ partially skew-symmetrize $\mathcal{T}_I^{\otimes^2}$ along "J" dim-s
STEP 4: $\mathbf{R}_2(\mathcal{T}) \leftarrow$ reshape $\mathcal{T}_{IJ}^{\otimes^2}$ into $I^2 J^2 \times K^2$ matrix

OLD: Illustration for $K = R$: algebraic algorithm

Properties of $\mathbf{R}_2(\mathcal{T})$

- rows of $\mathbf{R}_2(\mathcal{T})$ are vectorized $R \times R$ symmetric matrices
- $\dim \underbrace{\left(\ker(\mathbf{R}_2(\mathcal{T})) \bigcap (\text{ vectorized symmetric matrices})\right)}_{W} = R$
- $\mathbf{R}_2(\mathcal{T})(\mathbf{x} \otimes \mathbf{x}) = \mathbf{0} \Leftrightarrow \mathbf{x}$ is proportional to a column of $\mathbf{C}^{-T}$

$\Rightarrow$

if $\mathbf{w}_1, \ldots, \mathbf{w}_R$ is a basis of W,
then there exists a unique nonsingular $R \times R$ matrix $\mathbf{M}$:

$[\mathbf{w}_1 \ \ldots \ \mathbf{w}_R] = (\mathbf{C}^{-T} \odot \mathbf{C}^{-T})\mathbf{M} \sim$ CPD of $R \times R \times R$ tensor of rank R

$$\mathcal{W} = [\mathbf{C}^{-T}, \mathbf{C}^{-T}, \mathbf{M}^T]_R$$

NEW: Algorithm for $K = R$

I. Domanov, L. De Lathauwer. Canonical polyadic decomposition of third-order tensors: relaxed uniqueness conditions and algebraic algorithm, arXiv:1501.07251, 2015.

INPUT: $\mathcal{T} \in \mathbb{C}^{I\times J\times R}$ and $l \geq 0$
OUTPUT: Matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ such that $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$

1: Construct the $I^{2+l}J^{2+l} \times R^{2+l}$ matrix $\mathbf{R}_{2,l}(\mathcal{T})$
2: Find a basis $\mathbf{w}_1, \ldots, \mathbf{w}_R$ of $\ker(\mathbf{R}_{2,l}(\mathcal{T})) \bigcap S^{2+l}(\mathbb{R}^{R^{2+l}})$
3: $\mathbf{W} \leftarrow [\mathbf{w}_1 \ \ldots \ \mathbf{w}_R]$
4: Reshape the $R^{2+l} \times R$ matrix $\mathbf{W}$ into an $R \times R^{1+l} \times R$ tensor $\mathcal{W}$
5: Compute the CPD $\mathcal{W} = [\mathbf{C}^{-T}, \mathbf{C}^{-T} \odot \cdots \odot \mathbf{C}^{-T}, \mathbf{M}]_R$
6: Find the columns of $\mathbf{A}$ and $\mathbf{B}$ (trivial)

Example: $I \times J \times K$ tensor of rank $R = K = (I-1)(J-1)$

$I \times J \times (I-1)(J-1)$	l	C_{R+l+1}^{2+l}	t_1(sec)	t_2(sec)
$3 \times 3 \times 4$	0	10	0.012	0.008
$3 \times 4 \times 6$	0	21	0.022	0.013
$3 \times 5 \times 8$	0	36	0.038	0.013
$3 \times 6 \times 10$	0	55	0.060	0.014
$3 \times 7 \times 12$	1	364	0.368	0.035
$3 \times 8 \times 14$	1	560	0.725	0.071
$3 \times 9 \times 16$	1	816	1.342	0.156
$3 \times 10 \times 18$	1	1140	2.333	0.284
$3 \times 11 \times 20$	1	1540	4.259	0.773
$3 \times 12 \times 22$	1	2024	6.119	0.970
$3 \times 13 \times 24$	1	2600	9.386	1.698
$4 \times 4 \times 9$	0	45	0.047	0,013
$4 \times 5 \times 12$	1	364	0.367	0.034
$4 \times 6 \times 15$	1	680	0.988	0.108
$4 \times 7 \times 18$	2	5985	22.375	8.566
$4 \times 8 \times 21$	2	10626	56.758	36.272
$4 \times 9 \times 24$	2	17550	150.261	210.018
$5 \times 5 \times 16$	1	816	1.321	0.152
$5 \times 6 \times 20$	2	8855	41.213	22.903
$5 \times 7 \times 24$	2	17550	139.622	212.346
$6 \times 6 \times 25$	2	20475	771.171	443.346

Comparision with Tensorlab: $3 \times 7 \times 12$ tensor of rank 12

$\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_{12}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 5 & 7 & 0 & 6 & 6 & 7 & 9 & 0 & 8 \\ 2 & 3 & 5 & 7 & 0 & 6 & 6 & 7 & 9 & 0 & 8 & 2 \\ 3 & 5 & 7 & 0 & 6 & 6 & 7 & 9 & 0 & 8 & 2 & 1 \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 3 & 4 & 5 & 6 & 7 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 4 & 5 & 6 & 7 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 5 & 6 & 7 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 6 & 7 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 7 & 0 & 1 & 2 & 3 \end{bmatrix}, \qquad \mathbf{C} = \mathbf{I}_{12}$$

Our algorithm: takes less than 1 second
Tensorlab: cannot find with 500 random initializations
(Gauss-Newton dogleg trust region method, max_iter =10000)

Conjecture

Our algorithm can compute the CPD of a generic $I \times J \times K$ tensor of rank $R \leq K \leq (I-1)(J-1)$.

confirmed numerically for $R \leq 27$.

Case $K \leq R$: similar results on uniqueness and algorithm

$\mathbf{R}_{m,l}(\mathcal{T})$ is constructed for $m = R - K + 2 \geq 2$

	$I \times J \times K$	R	m	l	C_{R+l+m}^{m+l-1}	t_1(sec)	t_2(sec)
1	$4 \times 5 \times 6$	7	3	1	126	0.139	0.043
2	$5 \times 7 \times 7$	9	4	1	462	1.352	0.246
3	$6 \times 9 \times 8$	11	5	1	1716	18.969	9.647
4	$7 \times 7 \times 7$	10	5	1	924	6.227	1.965
5	$4 \times 6 \times 8$	9	3	1	330	0.553	0.077
6	$4 \times 7 \times 10$	11	3	1	715	1. 996	0.356
7	$5 \times 6 \times 6$	8	4	2	462	1.091	0.165
8	$5 \times 7 \times 8$	10	4	2	1716	11.396	3.156
9	$6 \times 7 \times 6$	9	5	3	1287	29.426	2.054

Corollary

Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$. *Suppose that*

$$k_{\mathbf{A}} + r_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2, \quad \text{and} \quad k_{\mathbf{B}} + k_{\mathbf{C}} \geq R + 2.$$

Then $r_{\mathcal{T}} = R$ *and the CPD of tensor* $\mathcal{T}$ *is unique and can be found algebraically.*

Corollary

Let $\mathcal{T} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]_R$ *and let* $k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2$.
Then the CPD of $\mathcal{T}$ *is unique and can be found algebraically.*

Future work

- Higher-order tensor with a long dimension
- Perturbation theory

Thank you for your attention.