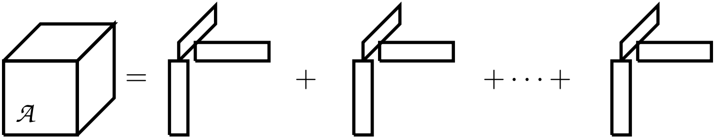
Pencil-based algorithms for the tensor rank decomposition are not stable

Paul Breiding (TU Berlin, MPI MiS)
Carlos Beltrán (Universidad de Cantabria)
Nick Vannieuwenhoven (KU Lueven)

personal-homepages.mis.mpg.de/breiding/
juliahomotopycontinuation.org

Hitchcock (1927) introduced the **tensor rank decomposition** or **CPD** for tensors $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$:

$$\mathcal{A} = \sum_{i=1}^{r} \mathbf{a}_i^1 \otimes \mathbf{a}_i^2 \cdots \otimes \mathbf{a}_i^d$$



The **rank** of a tensor is the minimum number of rank-1 tensors of which it is a linear combination.

Notation

$\mathcal{S} := \{\mathcal{A} \mid \mathrm{rank}(\mathcal{A}) = 1\}$ *are the rank-one tensors.*
$\sigma_r := \{\mathcal{A} \mid \mathrm{rank}(\mathcal{A}) \leq r\}$ *are the tensors of rank at most* r.

A direct algorithm for order-3 tensors

State-of-the art algorithms compute the CPD by applying optimization algorithms on the goal function $\frac{1}{2}\|\mathcal{A} - \sum_{i=1}^{r} \mathcal{A}_i\|^2$.

However, in some cases, the CPD of third-order tensors can be computed directly via a **generalized eigendecomposition**.

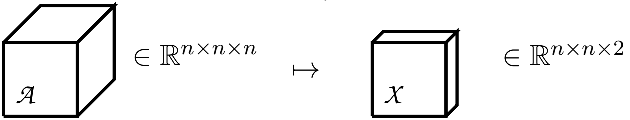
For simplicity, assume that $\mathcal{A} \in \mathbb{R}^{n\times n\times n}$ is of rank n. Say

$$\mathcal{A} = \sum_{i=1}^{n} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i.$$

The steps are as follows.

1. Choose a matrix $Q \in \mathbb{R}^{n\times 2}$ with orthonormal columns $\mathbf{q}_1, \mathbf{q}_2$.
2. Compute the **multilinear multiplication**

$$\mathcal{X} = (I, I, Q^T) \cdot \mathcal{A} := \sum_{i=1}^{n} \mathbf{a}_i \otimes \mathbf{b}_i \otimes (Q^T \mathbf{c}_i).$$



3. The two 3-**slices** X_1 and X_2 of $\mathcal{X}$ are

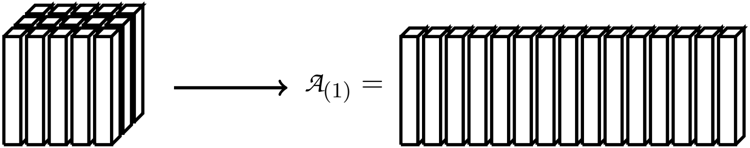
$$X_j = \sum_{i=1}^{n} \langle \mathbf{q}_j, \mathbf{c}_i \rangle \mathbf{a}_i \otimes \mathbf{b}_i = A \operatorname{diag}(\mathbf{q}_j^T C) B^T$$

where $A = [\mathbf{a}_i] \in \mathbb{R}^{n\times n}$ and likewise for B and C.
Hence, $X_1 X_2^{-1}$ has the following eigenvalue decomposition:

$$X_1 X_2^{-1} = A \operatorname{diag}(\mathbf{q}_1^T C) \operatorname{diag}(\mathbf{q}_2^T C)^{-1} A^{-1}$$

from which A can be found as the matrix of eigenvectors.

4. By a **1-flattening**

$\mathcal{A} =$ 

we find

$$\mathcal{A}_{(1)} := \sum_{i=1}^{n} \mathbf{a}_i (\mathbf{b}_i \otimes \mathbf{c}_i)^T = A(B \odot C)^T,$$

where $B \odot C := [\mathbf{b}_i \otimes \mathbf{c}_i]_i \in \mathbb{R}^{n^2 \times n}$. Computing

$$A \odot (A^{-1}\mathcal{A}_{(1)})^T = A \odot (B \odot C) = [\mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i]_i,$$

solves the tensor decomposition problem.

Let's perform an experiment in Tensorlab v3.0 with this decomposition algorithm.

Create a rank-25 random tensor of size $25 \times 25 \times 25$:

```
% Ut{i} has as columns the i-th factors
>> Ut{1} = randn(25,25);
>> Ut{2} = randn(25,25);
>> Ut{3} = randn(25,25);
% generate the full tensor
>> A = cpdgen(Ut);
```

Compute $\mathcal{A}$'s decomposition and its distance to the input decomposition, relative to the machine precision $\epsilon \approx 2 \cdot 10^{-16}$:

```
>> Ur = cpd_gevd(A, 25);
>> E = kr(Ut) - kr(Ur);
>> norm( E(:), 2 ) / eps
ans =
     8.6249e+04
```

What happened?

Let us look more closely at the computational problem:

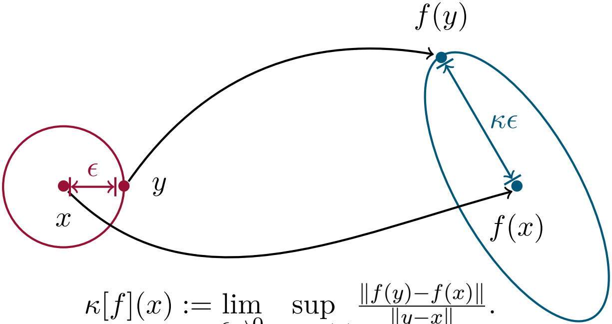
- The input is a tensor $\mathcal{A} \in \sigma_{25} \subset \mathbb{R}^{25\times 25\times 25}$ of rank 25.
- The output is the tuple $(\mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i)_{i=1}^{25} \in \mathcal{S}^{\times 25}$.

Let $f : \sigma_{25} \to \mathcal{S}^{\times 25}$ be the function that maps a tensor to its decomposition. Then, what we observed was

$$\frac{\|f(\mathcal{A}) - f(\mathcal{A}')\|}{\|\mathcal{A} - \mathcal{A}'\|} \approx 8 \cdot 10^4$$

with $\|\mathcal{A} - \mathcal{A}'\| \approx 2 \cdot 10^{-16}$.

The **condition number** quantifies the **worst-case sensitivity** of f to perturbations of the input.



$$\kappa[f](x) := \lim_{\epsilon \to 0} \sup_{y \in B_\epsilon(x)} \frac{\|f(y)-f(x)\|}{\|y-x\|}.$$

When f is differentiable: $\kappa[f](x) = \|\mathrm{d}_x f\|$.

Problem: There is no function $f : \sigma_{25} \to \mathcal{S}^{\times 25}$ that maps a tensors to its decomposition.

If the set of rank-1 tensors $\{\mathcal{A}_1, \ldots, \mathcal{A}_r\}$ is uniquely determined given the rank-r tensor $\mathcal{A} = \mathcal{A}_1 + \cdots + \mathcal{A}_r$, then we call $\mathcal{A}$ an r**-identifiable** tensor.

Note that matrices are never r-identifiable, because

$$M = \sum_{i=1}^{r} \mathbf{a}_i \otimes \mathbf{b}_i = AB^T = (AX^{-1})(BX^T)^T$$

for every invertible X. In general, these factorizations are different.

Kruskal (1977) gave a famous sufficient condition for proving the r-identifiability of third-order tensors.

More recently r-identifiability was studied in algebraic geometry. This is a natural framework because the set of rank-1 tensors

$$\mathcal{S}^{\mathbb{C}} := \left\{\mathbf{a}^1 \otimes \mathbf{a}^2 \otimes \cdots \otimes \mathbf{a}^d \mid \mathbf{a}^k \in \mathbb{C}^{n_k} \setminus \{0\}\right\}$$

is the smooth projective **Segre variety**.

The set of tensors of rank bounded by r,

$$\sigma_r^{\mathbb{C}} := \left\{\mathcal{A}_1 + \cdots + \mathcal{A}_r \mid \mathcal{A}_i \in \mathcal{S}\right\},$$

is the Zariski-open constructible part of the projective r**-secant variety** of the Segre variety.

In the words of algebraic geometry:

$\sigma_r^{\mathbb{C}}$ is generically r-identifiable, if the **addition map**:

$$\Phi_r : \mathcal{S}^{\mathbb{C}} \times \cdots \times \mathcal{S}^{\mathbb{C}} \to \sigma_r^{\mathbb{C}}$$
$$(\mathcal{A}_1, \ldots, \mathcal{A}_r) \mapsto \mathcal{A}_1 + \cdots + \mathcal{A}_r$$

is of *degree* $r!$.

Let $n_1 \geq \cdots \geq n_d$ and

$$r_{\mathsf{cr}} = \frac{n_1 \cdots n_d}{1 + \sum_{i=1}^{d} (n_i - 1)}, \quad r_{\mathsf{ub}} = n_2 \cdots n_d - \sum_{k=2}^{d} (n_k - 1).$$

Conjectured general rule:

$r \geq r_{\mathsf{cr}}$ or $d = 2$	$\to$	not gen. r-identifiable
$n_1 > r_{\mathsf{ub}}$ and $r \geq r_{\mathsf{ub}}$	$\to$	not gen. r-identifiable
none of foregoing and $r < r_{\mathsf{cr}}$	$\to$	gen. r-identifiable;

see Chiantini, Ottaviani, Vannieuwenhoven (2014) for a proof in the case $n_1 \cdots n_d \leq 15000$.

The real case is more involved because now

$$\sigma_r := \sigma_r^{\mathbb{R}} := \{\mathcal{A}_1 + \cdots + \mathcal{A}_r \mid \mathcal{A}_i \in \mathcal{S}^{\mathbb{R}}\},$$

is a **semi-algebraic set**.

Qi, Comon, and Lim (2016) showed that if $\sigma_r^{\mathbb{C}}$ is generically r-identifiable, then it follows that the set of real rank-r tensors with multiple complex CPDs is contained in a proper Zariski-closed subset of $\sigma_r^{\mathbb{R}}$.

In this sense, $\sigma_r^{\mathbb{R}}$ is also generically r-identifiable.

In the following we abbreviate $\mathcal{S} := \mathcal{S}^{\mathbb{R}}$, $\sigma_r := \sigma_r^{\mathbb{R}}$ and we will assume that $\sigma_r^{\mathbb{C}}$ is generically r-identifiable.

Defining the tensor decomposition map

To define the tensor decomposition map, we analyze the real addition map:

$$\Phi_r : \mathcal{S} \times \cdots \times \mathcal{S} \to \mathbb{R}^{n_1 \times \cdots \times n_d}$$
$$(\mathcal{A}_1, \ldots, \mathcal{A}_r) \mapsto \mathcal{A}_1 + \cdots + \mathcal{A}_r$$

Note that the domain and codomain are **smooth manifolds**.

The idea is to define Φ_r on the quotient of $\mathcal{S}^{\times r}$ by the symmetric group and restricting the domain of Φ_r to a Zariski-open smooth submanifold. Then, Φ_r restricts to a **diffeomorphism** onto its image.

Let $\mathbf{n} = (n_1, \ldots, n_d)$.

Let $\mathcal{M}_{r;\mathbf{n}} \subset \mathcal{S}^{\times r}$ be the set of tuples of $n_1 \times \cdots \times n_d$ rank-1 tensors $\mathfrak{a} = (\mathcal{A}_1, \ldots, \mathcal{A}_r)$ that satisfy:

1. $\Phi_r(\mathfrak{a})$ is a **smooth point** of the semi-algebraic set σ_r;
2. $\Phi_r(\mathfrak{a})$ is complex r**-identifiable**;
3. the **derivative** $\mathrm{d}_{\mathfrak{a}}\Phi_r$ **is injective**

Definition

The set of r**-nice tensors** is

$$\mathcal{N}_{r;\mathbf{n}} := \Phi_r(\mathcal{M}_{r;\mathbf{n}}).$$

One can prove the following results:

Proposition

Let $\sigma_r^{\mathbb{C}}$ be generically r-identifiable. Then, the following holds.

1. *$\mathcal{N}_{r,\mathbf{n}}$ is an open dense submanifold of σ_r.*
2. *$\widehat{\mathcal{M}}_{r;\mathbf{n}} := \mathcal{M}_{r;\mathbf{n}}/\mathfrak{S}_r$ is a manifold and the projection is a local diffeomorphism.*
3. *The addition map*

$$\Phi_r : \widehat{\mathcal{M}}_{r;\mathbf{n}} \to \mathcal{N}_{r;\mathbf{n}}, \{\mathcal{A}_1, \ldots, \mathcal{A}_r\} \to \mathcal{A}_1 + \cdots + \mathcal{A}_r$$

is a diffeomorphism.

The inverse of $\Phi_r : \widehat{\mathcal{M}}_{r;\mathbf{n}} \to \mathcal{N}_{r;\mathbf{n}}$ is

$$\tau_{r;\mathbf{n}} : \mathcal{N}_{r;\mathbf{n}} \to \widehat{\mathcal{M}}_{r;\mathbf{n}},\ \mathcal{A}_1 + \cdots + \mathcal{A}_r \to \{\mathcal{A}_1, \ldots, \mathcal{A}_r\}.$$

We call it the **tensor rank decomposition map**.

Definition

The condition number of the tensor rank decomposition for a tensor $\mathcal{A} \in \mathcal{N}_{r;\mathbf{n}}$ is

$$\kappa[\tau_{r;\mathbf{n}}](\mathcal{A}) = \|\mathrm{d}_{\mathcal{A}}\tau_{r;\mathbf{n}}\|_2 = \|(\mathrm{d}_{\mathfrak{a}}\Phi_r)^{-1}\|_2.$$

If

$$\mathcal{A} = \mathcal{A}_1 + \cdots + \mathcal{A}_r = \sum_{i=1}^{r} \mathbf{a}_i^1 \otimes \cdots \otimes \mathbf{a}_i^d$$

$$\mathcal{B} = \mathcal{B}_1 + \cdots + \mathcal{B}_r = \sum_{i=1}^{r} \mathbf{b}_i^1 \otimes \cdots \otimes \mathbf{b}_i^d$$

are tensors in $\mathbb{R}^{n_1 \times \cdots \times n_d}$, then for $\|\mathcal{A} - \mathcal{B}\|_F \approx 0$ we have the **asymptotically sharp bound**

$$\underbrace{\min_{\pi \in \mathfrak{S}_r} \sqrt{\sum_{i=1}^{r} \|\mathcal{A}_i - \mathcal{B}_{\pi_i}\|_F^2}}_{\text{forward error}} \lesssim \underbrace{\kappa[\mathcal{T}_{r;\mathbf{n}}](\mathcal{A})}_{\text{condition number}} \cdot \underbrace{\|\mathcal{A} - \mathcal{B}\|_F}_{\text{backward error}}$$

Back to our example

```
>> Ut{1} = randn(25,25);
>> Ut{2} = randn(25,25);
>> Ut{3} = randn(25,25);
>> A = cpdgen(Ut);
>> Ur = cpd_gevd(A, 25);
>> E = kr(Ut) - kr(Ur);
>> norm( E(:), 2 ) / eps
ans =
     8.6249e+04
```

We understand now that this can happen, because of a high condition number. However,

```
>> kappa = condition_number( Ut )
ans =
     2.134
```

The only explanation is that **there is something wrong with the algorithm**.

We show that algorithms based on a reduction to tensors in $\mathbb{R}^{n_1 \times n_2 \times 2}$ are **numerically unstable**.

The forward error produced by the algorithm divided by the backward error is "much" larger than the condition number, for some inputs.

Pencil-based algorithms

A **pencil-based algorithm** (PBA) is an algorithm that computes the CPD of

$$\mathcal{A} = \sum_{i=1}^{r} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i \in \mathcal{N}_{r;n_1,n_2,n_3} \subset \mathbb{R}^{n_1 \times n_2 \times n_d}$$

in the following way:

- S1. Choose a fixed $Q \in \mathbb{R}^{n_3 \times 2}$ with orthonormal columns.
- S2. $\mathcal{B} \leftarrow (I, I, Q^T) \cdot \mathcal{A}$;
- S3. $\{\mathbf{a}_1, \ldots, \mathbf{a}_r\} \leftarrow \widehat{\theta}(\mathcal{B})$;
- S4.a Choose an order $A := (\mathbf{a}_1, \ldots, \mathbf{a}_r)$;
- S4.b $(\mathbf{b}_1 \otimes \mathbf{c}_1, \ldots, \mathbf{b}_r \otimes \mathbf{c}_r) \leftarrow (A^\dagger \mathcal{A}_{(1)})^T$;
- S5. $\texttt{output} \leftarrow \pi(\odot((\mathbf{a}_1, \ldots, \mathbf{a}_r), (\mathbf{b}_1 \otimes \mathbf{c}_1, \ldots, \mathbf{b}_r \otimes \mathbf{c}_r)))$.

where $\pi : \mathcal{S}^{\times r} \to (\mathcal{S}^{\times r}/\mathfrak{S}_r)$ and $\odot$ is the Khatri–Rao product: $\odot(A, B) := (\mathbf{a}_i \otimes \mathbf{b}_i)_i$.

Pencil-based algorithms

A **pencil-based algorithm** (PBA) is an algorithm that computes the CPD of

$$\mathcal{A} = \sum_{i=1}^{r} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i \in \mathcal{N}_{r;n_1,n_2,n_3} \subset \mathbb{R}^{n_1 \times n_2 \times n_d}$$

in the following way:

OK	Choose a fixed $Q \in \mathbb{R}^{n_3 \times 2}$ with orthonormal columns.
OK	$\mathcal{B} \leftarrow (I, I, Q^T) \cdot \mathcal{A}$;
BAD	$\{\mathbf{a}_1, \ldots, \mathbf{a}_r\} \leftarrow \widehat{\theta}(\mathcal{B})$;
OK	Choose an order $A := (\mathbf{a}_1, \ldots, \mathbf{a}_r)$;
OK	$(\mathbf{b}_1 \otimes \mathbf{c}_1, \ldots, \mathbf{b}_r \otimes \mathbf{c}_r) \leftarrow (A^\dagger \mathcal{A}_{(1)})^T$;
OK	$\mathtt{output} \leftarrow \pi(\odot((\mathbf{a}_1, \ldots, \mathbf{a}_r), (\mathbf{b}_1 \otimes \mathbf{c}_1, \ldots, \mathbf{b}_r \otimes \mathbf{c}_r)))$.

where $\pi : \mathcal{S}^{\times r} \to (\mathcal{S}^{\times r}/\mathfrak{S}_r)$ and $\odot$ is the Khatri–Rao product: $\odot(A, B) := (\mathbf{a}_i \otimes \mathbf{b}_i)_i$.

The magic map $\widehat{\theta}$ needs to recover the vectors from the first factor matrix **when restricted to** $\mathcal{N}_{r;n_1,n_2,2}$:

$$\begin{aligned} \widehat{\theta}|_{\mathcal{N}_{r;n_1,n_2,2}} : \qquad\qquad \mathcal{N}_{r;n_1,n_2,2} &\longrightarrow (\mathbb{S}^{n_1-1})^{\times r}/\mathfrak{S}_r \\ \mathcal{B} = \sum_{i=1}^{r} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{z}_i &\longmapsto \{\mathbf{a}_1, \ldots, \mathbf{a}_r\} \end{aligned}$$

Since the input to $\widehat{\theta}$ will be the result of a previous numerical computation, the domain of definition of $\widehat{\theta}$ should also encompass a sufficiently large neighborhood of $\mathcal{N}_{r;n_1,n_2,2}$!

For proving instability, it does not matter what $\widehat{\theta}$ computes outside of $\mathcal{N}_{r;n_1,n_2,2}$.

For a valid input $\mathcal{A} \in \mathcal{N}_{r;n_1,n_2,n_3}$, let $\{\widetilde{\mathcal{A}}_1, \ldots, \widetilde{\mathcal{A}}_r\}$ be the CPD (in floating-point representation) returned by the PBA.

Our proof strategy consists of showing that for every $\epsilon > 0$ there exists an open neighborhood $\mathcal{O}_\epsilon \subset \mathcal{N}_{r;n_1,n_2,n_3}$ of r-nice tensors such that the **excess factor**

$$\omega(\mathcal{A}) = \frac{\text{observed forward error due to algorithm}}{\text{maximum forward error due to problem}}$$

$$:= \frac{\min_{\pi \in \mathfrak{S}_r} \sqrt{\sum_{i=1}^{r} \|\mathcal{A}_i - \widetilde{\mathcal{A}}_i\|^2}}{\kappa[\tau_{r;n_1,n_2,n_3}](\mathcal{A}) \cdot \|\mathcal{A} - \mathrm{fl}(\mathcal{A})\|_F}$$

behaves like a constant times ϵ^{-1}.

Formally, we showed the following result:

Theorem (Beltrán, Breiding, Vannieuwenhoven (2018))

There exist a constant $k > 0$ and a tensor $\mathcal{O} \in \mathcal{N}_{r;n_1,n_2,n_3}$ with the following properties: For all sufficiently small $\epsilon > 0$, there exists an open neighborhood $\mathcal{O}_\epsilon$ of $\mathcal{O}$, such that for all tensors $\mathcal{A} \in \mathcal{O}_\epsilon$ we have

1. $\mathcal{A} \in \mathcal{N}_{r;n_1,n_2,n_3}$ *is a valid input for a PBA, and*
2. $\omega(\mathcal{A}) \geq k\epsilon^{-1}$.

In other words, the forward error produced by a PBA can be larger than the maximum forward error expected from the tensor decomposition problem by an arbitrarily large factor.

Distribution of the condition number

We conceived the existence of this problem after seeing the distribution of the condition number of random rank-1 tensors

$$\mathcal{A}_i = \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i \in \mathbb{R}^{n_1 \times n_2 \times n_3}$$

where

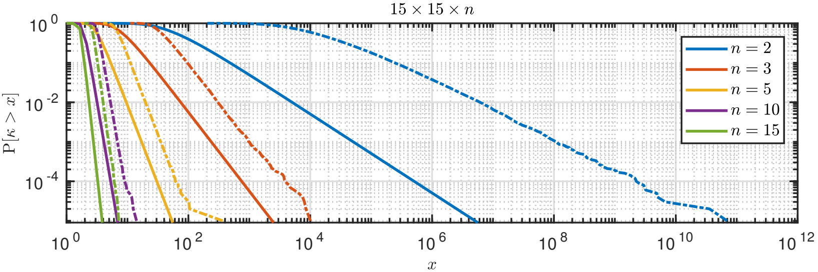
- $\mathbf{a}_i \in \mathbb{R}^{n_1}$ and $\mathbf{b}_i \in \mathbb{R}^{n_2}$ are **arbitrary**, and
- the $\mathbf{c}_i \in \mathbb{R}^{n_3}$ are i.i.d. random Gaussian vectors.

We showed, based on Cai, Fan, and Jiang (2013), that

$$\mathrm{P}[\kappa \geq \alpha] \geq \mathrm{P}\Big[\max_{1 \leq i \neq j \leq r} \frac{1}{\sqrt{1 - \langle \mathbf{c}_i, \mathbf{c}_j \rangle}} \geq \alpha\Big] \to 1 - e^{-Kr^2\alpha^{1-n_3}},$$

as $r \to \infty$; herein, K is a constant depending only on n_3.

Empirical distribution of $\kappa(\mathcal{A})$



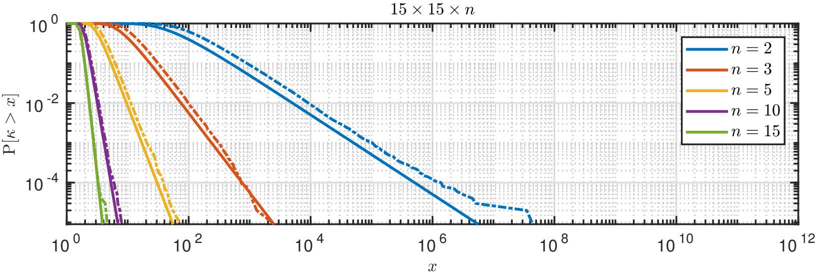
10^5 random rank-15 tensors $\sum_{i=1}^{15} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i \in \mathbb{R}^{15\times15\times n}$, where

$$A = [\mathbf{a}_1, \ldots, \mathbf{a}_{15}] \text{ and } B = [\mathbf{b}_1, \ldots, \mathbf{b}_{15}] \text{ and } C = [\mathbf{c}_1, \ldots, \mathbf{c}_{15}]$$

are Gaussian matrices.

dashed lines = empirical distribution.

Empirical distribution of $\kappa(\mathcal{A})$



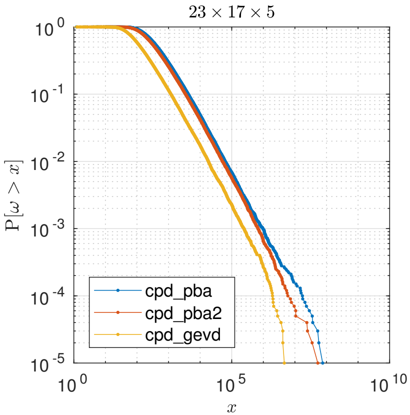
10^5 random rank-15 tensors $\sum_{i=1}^{15} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i \in \mathbb{R}^{15\times 15\times n}$, where

$$A = [\mathbf{a}_1, \ldots, \mathbf{a}_{15}] \text{ and } B = [\mathbf{b}_1, \ldots, \mathbf{b}_{15}]$$

are random orthogonal matrices and $C = [\mathbf{c}_1, \ldots, \mathbf{c}_{15}]$ is a Gaussian matrix.

dashed lines = empirical distribution.

Empirical distribution of the excess factor

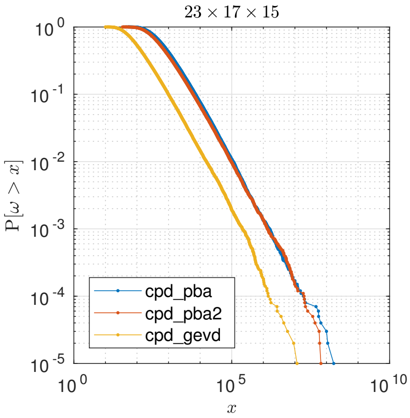


10^5 random rank-17 tensors $\sum_{i=1}^{17} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i \in \mathbb{R}^{23\times17\times5}$:

$A = [\mathbf{a}_1, \ldots, \mathbf{a}_{17}]$ and $B = [\mathbf{b}_1, \ldots, \mathbf{b}_{17}]$ and $C = [\mathbf{c}_1, \ldots, \mathbf{c}_{17}]$

are Gaussian matrices.

Empirical distribution of the excess factor

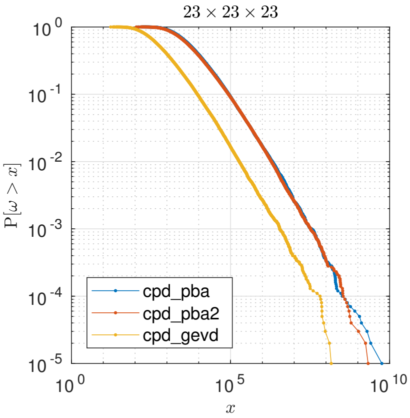


10^5 random rank-17 tensors $\sum_{i=1}^{17} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i \in \mathbb{R}^{23\times17\times15}$:

$A = [\mathbf{a}_1, \ldots, \mathbf{a}_{17}]$ and $B = [\mathbf{b}_1, \ldots, \mathbf{b}_{17}]$ and $C = [\mathbf{c}_1, \ldots, \mathbf{c}_{17}]$

are Gaussian matrices.

Empirical distribution of the excess factor



10^5 random rank-23 tensors $\sum_{i=1}^{23} \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i \in \mathbb{R}^{23\times 23\times 23}$:

$A = [\mathbf{a}_1, \ldots, \mathbf{a}_{23}]$ and $B = [\mathbf{b}_1, \ldots, \mathbf{b}_{23}]$ and $C = [\mathbf{c}_1, \ldots, \mathbf{c}_{23}]$

are Gaussian matrices.

Conclusions

Take-away story:

1. Tensors are conjectured to be generically r-identifiable for almost all low ranks r.
2. The condition number of the CPD measures the stability of the unique rank-1 tensors.
3. Reduction to a matrix pencil yields numerically unstable algorithms for computing CPDs.

Further reading

- Beltrán, Breiding, and Vannieuwenhoven, *Pencil-based algorithms for tensor rank decomposition are not stable*, SIAM J. Matrix Anal. and Appl.
- Beltrán, Breiding, and Vannieuwenhoven, *The average condition number of most tensor rank decomposition problems is infinite*, arXiv1903.05527.
- Breiding and Vannieuwenhoven, *The condition number of join decompositions*, SIAM J. Matrix Anal. and Appl., 2018.
- Breiding and Vannieuwenhoven, *On the average condition number of tensor rank decompositions*, arXiv:1801.01673.
- Breiding and Vannieuwenhoven, *A Riemannian trust region method for the canonical tensor rank approximation problem*, SIAM J. Optim, 2018.



References

- Kruskal, *Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics*, Lin. Alg. Appl., 1977.
- Abo, Ottaviani and Peterson, *Induction for secant varieties of Segre varieties*, Trans. AMS, 2009.
- Chiantini and Ottaviani, *On Generic Identifiability of 3-Tensors of Small Rank*, SIMAX, 2012.
- Bocci and Chiantini, *On the identifiability of binary Segre products*, J. Algebraic Geom., 2013.
- Bocci, Chiantini, and Ottaviani, *Refined methods for the identifiability of tensors*, Ann. Mat. Pura Appl., 2013.
- Chiantini, Mella and Ottaviani, *One example of general unidentifiable tensors*, J. Alg. Stat., 2014.
- Chiantini, Ottaviani, and Vannieuwenhoven, *An algorithm for generic and low-rank specific identifiability of complex tensors*, SIMAX, 2014.
- Qi, Comon, and Lim, *Semialgebraic geometry of nonnegative tensor rank*, SIMAX, 2016.