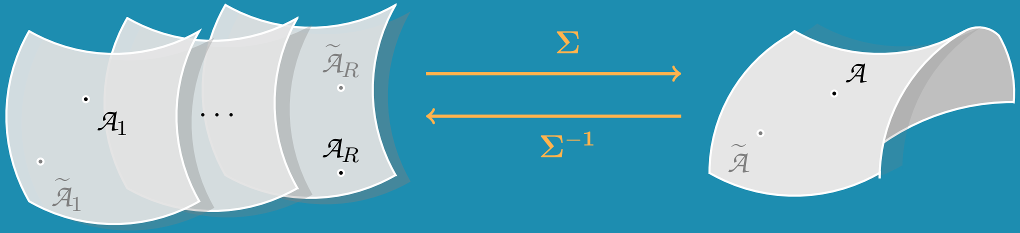
KU LEUVEN



Condition numbers of tensor decompositions

and their invariance properties

Nick Dewaele, Paul Breiding (U. Osnabrück), Nick Vannieuwenhoven

0 Motivating example

Suppose we want to decompose the $3 \times 3 \times 3$ tensor

$$\mathcal{A} = \left[\!\left[\begin{bmatrix} & 1 & \\ 1 & & \\ & & \end{bmatrix} \begin{bmatrix} 1 & & \varepsilon \\ & & \\ \varepsilon & & \varepsilon^2 \end{bmatrix} \begin{bmatrix} & \varepsilon & \\ & & \\ & & \end{bmatrix} \right]\!\right]$$

where $\varepsilon = 10^{-6}$. In practice, we only measure an approximation:

$$\tilde{\mathcal{A}} = \left[\!\left[\begin{bmatrix} & 1 & \\ 1 & \varepsilon & \\ & & \end{bmatrix} \begin{bmatrix} 1 & \varepsilon & \\ \varepsilon & \varepsilon^2 & \\ & & \end{bmatrix} \begin{bmatrix} & & \\ & & \\ & & \varepsilon \end{bmatrix} \right]\!\right].$$

Since $\left\| \mathcal{A} - \tilde{\mathcal{A}} \right\|_F \approx \sqrt{7}\varepsilon$, this is pretty accurate.

$\hookrightarrow$ Surely, we can recover the model accurately *...right?*

0 Motivating example

$\mathcal{A}$ splits uniquely into 3 tensors of rank 1:

$$\mathcal{A} = \underbrace{e_1 \otimes e_1 \otimes e_2}_{\mathbf{E_{112}}} + \underbrace{(e_1 + \varepsilon e_3) \otimes e_2 \otimes e_1}_{\approx \mathbf{E_{121}}} + \underbrace{e_2 \otimes (e_1 + \varepsilon e_3) \otimes (e_1 + \varepsilon e_3)}_{\approx \mathbf{E_{211}}}.$$

(where $\mathbf{E_{ijk}} \in \mathbb{R}^{3\times3\times3}$ has entry 1 at index ijk and 0 otherwise).

But $\widetilde{\mathcal{A}}$ decomposes uniquely as

$$\widetilde{\mathcal{A}} = \underbrace{\varepsilon^{-1}(e_1 + \varepsilon e_2)^{\otimes 3}}_{\approx \varepsilon^{-1}\mathbf{E_{111}}} - \underbrace{\varepsilon^{-1} e_1^{\otimes 3}}_{\varepsilon^{-1}\mathbf{E_{111}}} + \underbrace{\varepsilon e_3^{\otimes 3}}_{\varepsilon \mathbf{E_{333}}}.$$

The rank-1 terms are entirely different! The model is *uninterpretable!*

0 Motivating example

$\mathcal{A}$ splits uniquely into 3 tensors of rank 1:

$$\mathcal{A} = \underbrace{e_1 \otimes e_1 \otimes e_2}_{\mathbf{E_{112}}} + \underbrace{(e_1 + \varepsilon e_3) \otimes e_2 \otimes e_1}_{\approx \mathbf{E_{121}}} + \underbrace{e_2 \otimes (e_1 + \varepsilon e_3) \otimes (e_1 + \varepsilon e_3)}_{\approx \mathbf{E_{211}}}.$$

(where $\mathbf{E_{ijk}} \in \mathbb{R}^{3\times3\times3}$ has entry 1 at index ijk and 0 otherwise).

For another small perturbation of $\mathcal{A}$, there isn't even a unique decomposition anymore:

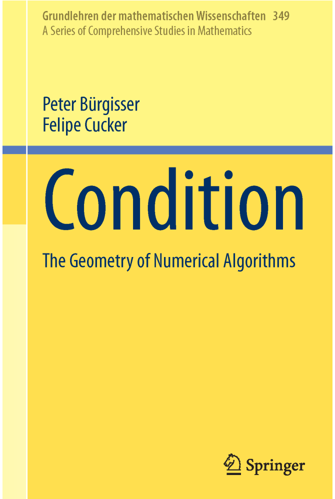
$$\begin{aligned}\tilde{\tilde{\mathcal{A}}} &:= e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_1 \\ &= e_1 \otimes \underbrace{(e_1 \otimes e_2 + e_2 \otimes e_1)}_{\infty \text{ many rank-2 decompositions}} + e_2 \otimes e_1 \otimes e_1.\end{aligned}$$

There's some numerical shenanigans here. We should measure the precise amount of it.

1 Outline

1. Condition: the main concept
2. Tensor manifolds and join sets
3. Condition of additive decompositions
4. Compression and invariance of the condition number

1 Condition numbers



Condition number = estimate of *numerical difficulty*.

I.e., how difficult it is to compute something precisely when working with inexact data or inexact arithmetic?

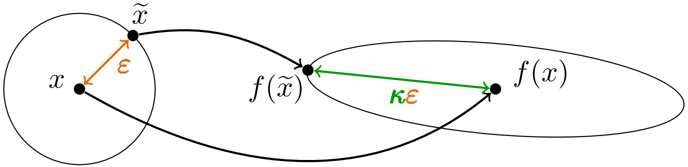
A staple in numerical analysis, it is used in

- perturbation theory
- analysis of (iterative) algorithms
- regression analysis
- ...

1 Definition of the condition number

Let $f : \mathcal{X} \to \mathcal{Y}$ be a map between metric spaces. The condition number of f is

$$\kappa[f](x) := \lim_{\varepsilon \to 0} \sup_{d(x,\widetilde{x}) \leqslant \varepsilon} \frac{d(f(x), f(\widetilde{x}))}{d(x, \widetilde{x})}.$$



If $\mathcal{X}$ and $\mathcal{Y}$ are Riemannian manifolds, then $\kappa[f](x)$ is the spectral norm of $Df(x)$ [Rice 1966].

1 Some celebrated theorems involving condition numbers

Theorem (distance to ill-posedness, Eckart & Young 1936; Schmidt 1907)

Let $A \in GL(n)$ *and let* $\Sigma := \mathbb{R}^{n \times n} \setminus GL(n)$. *Then*

$$d(A, \Sigma) := \inf_{X \in \Sigma} \frac{\|A - X\|_F}{\|A\|_F} = \frac{1}{\boldsymbol{\kappa}(\mathbf{A})}$$

where $\boldsymbol{\kappa}(\mathbf{A})$ *is the condition number of* $A \mapsto A^{-1}$.

Similar theorems exist for polynomial systems, eigenvalues, ...

1 Some celebrated theorems involving condition numbers

Theorem (condition numbers and convergence rate, see e.g. Nocedal & Wright 2006)

Let $h : \mathbb{R}^n \to \mathbb{R}, x \mapsto \frac{1}{2}x^T Ax - b^T x$ *be convex. If* $x_\star = \arg\min_x h(x)$ *is computed with the conjugate gradient algorithm, yielding iterates* $x_0, x_1, \ldots,$ *then*

$$\|x_\star - x_k\|_A \le 2 \left(\frac{\sqrt{\boldsymbol{\kappa}(\mathbf{A})} - 1}{\sqrt{\boldsymbol{\kappa}(\mathbf{A})} + 1} \right)^k \|x_\star - x_0\|_A.$$

1 Some celebrated theorems involving condition numbers

Theorem (Region of attraction, Blum et al. 1998)

Let $F(z) = 0$ *be a homogeneous polynomial system of degree* D *with simple roots* $\zeta_1, \ldots, \zeta_K \in \mathbb{CP}^n$. *For* $k = 1, \ldots, K$, *there exists a local solution map* $G_k : F \mapsto \zeta_k$. *Let* $z_0, z_1, z_2, \ldots$ *be the iterates from Newton's method applied to* $F(z) = 0$. *If*

$$d(z_0, \zeta_k) \leq \frac{3 - \sqrt{7}}{D^2 \boldsymbol{\kappa}[\mathbf{G_k}](\mathbf{F})}$$

then the iterates converge quadratically to ζ_k.

1 Ongoing research project

1. Characterise the condition number of any tensor decomposition
2. Compute it efficiently
3. Connect it to the previous theorems.

Today's talk: answer 1-2 for general additive decompositions.

2 Outline

1. Condition: the main concept
2. Tensor manifolds and join sets
3. Condition of additive decompositions
4. Compression and invariance of the condition number

2 Basic concepts

Tensors: $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_D}$. Rank 1 tensors: $\mathbf{a}_1 \otimes \cdots \otimes \mathbf{a}_D$.
Multilinear multiplication / Tucker product:

$$\mathcal{A} = (U_1 \otimes \cdots \otimes U_D)\mathcal{G} \tag{1}$$

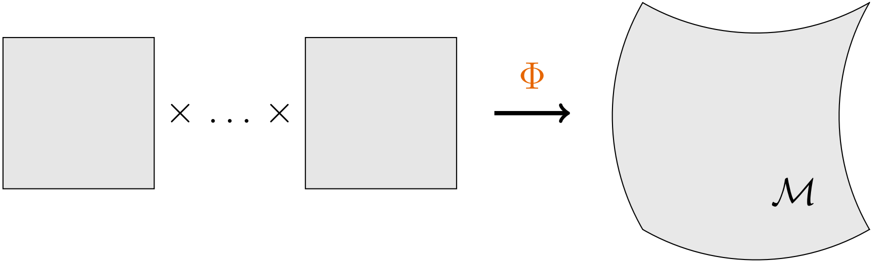
with $\mathcal{G} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$

- Coordinate definition: $\mathcal{A}_{ijk} = \sum_{l,m,n} (U_1)_{il} (U_2)_{jm} (U_3)_{kn} \mathcal{G}_{lmn}$
- Linear extension of $(U_1 \otimes \cdots \otimes U_D)(\mathbf{g}_1 \otimes \mathbf{g}_2 \otimes \mathbf{g}_3) = (U_1\mathbf{g}_1 \otimes U_2\mathbf{g}_2 \otimes U_3\mathbf{g}_3)$.

Multilinear rank of $\mathcal{A}$ = dimensions of the smallest $\mathcal{G}$ so that (1) holds. A dense subset $\mathbb{R}_\star^{n_1 \times \cdots \times n_D}$ has the maximal ML rank.

2 Tensor manifolds

Maps defined with multilinear products often have smooth manifolds as their image.



2 Examples of tensor manifolds

Segre manifold:

$$(\mathbb{R}^{n_1} \setminus \{0\}) \times (\mathbb{R}^{n_2} \setminus \{0\}) \times (\mathbb{R}^{n_3} \setminus \{0\}) \to \mathbb{R}^{n_1 \times n_2 \times n_3}$$
$$\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \mapsto \mathbf{a}_1 \otimes \mathbf{a}_2 \otimes \mathbf{v}_3$$

Tucker manifold:

$$\mathbb{R}_\star^{n_1 \times m_1} \times \mathbb{R}_\star^{n_2 \times m_2} \times \mathbb{R}_\star^{n_3 \times m_2} \times \mathbb{R}_\star^{m_1 \times m_2 \times m_3} \to \mathbb{R}^{n_1 \times n_2 \times n_3}$$
$$U_1, U_2, U_3, \mathcal{G} \mapsto (U_1 \otimes U_2 \otimes U_3)\mathcal{G}$$

Tensor train manifold:

$$A, \mathcal{B}, \mathcal{C}, D \mapsto \mathcal{X} \quad \text{where} \quad \mathcal{X}_{ijkl} = \sum_{\alpha,\beta,\gamma} A_{i,\alpha} \mathcal{B}_{\alpha,j,\beta} \mathcal{C}_{\beta,k,\gamma} D_{\gamma,k}$$

2 The structured Tucker manifold

Definition

A smooth manifold $\mathcal{N} \subseteq \mathbb{R}_\star^{m_1 \times \cdots \times m_D}$ is a *core structure* if it is invariant under arbitrary changes of basis.

Definition (D, Breiding & Vannieuwenhoven 2021)

A *structured Tucker manifold* is the image of

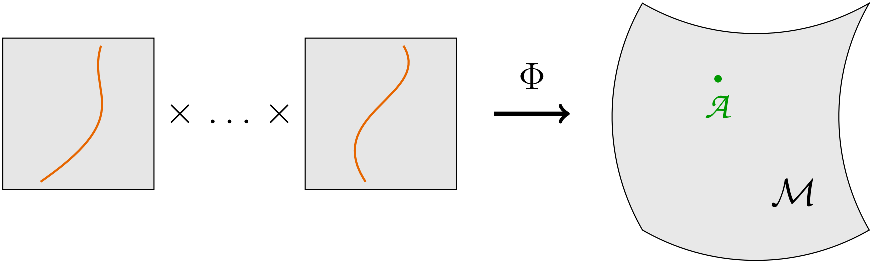
$$\begin{aligned} \Phi : \mathbb{R}_\star^{n_1 \times m_1} \times \cdots \times \mathbb{R}_\star^{n_D \times m_D} \times \mathcal{N} &\to \mathbb{R}^{n_1 \times \cdots \times n_D} \\ U_1, \ldots, U_D, \mathcal{G} &\mapsto (U_1 \otimes \cdots \otimes U_D)\mathcal{G} \end{aligned}$$

where $\mathcal{N}$ is a core structure.

This is also a smooth manifold.

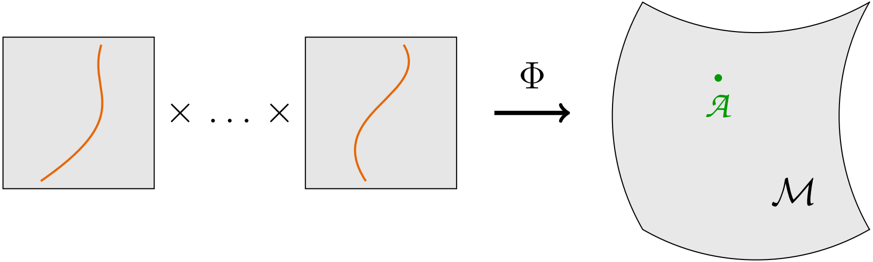
2 Intuition behind the manifold

Structured Tucker manifold $\approx$ smoothly factorisable tensors.



Given $\mathcal{A} \in \mathcal{M}$, the fibre $\Phi^{-1}(\mathcal{A})$ is a smooth submanifold of the domain. If you glue each fibre to one point, the above is a diffeomorphism.

2 Sidenote: condition of multiplicative decompositions



Because $\Phi^{-1}(\mathcal{A})$ is not single valued, it is not so easy to define how it changes under small perturbations of $\mathcal{A}$.
Ongoing work: we can compute a condition number that measures local changes to the fibres.

2 Additive decompositions

Given manifolds $\mathcal{M}_1, \dots, \mathcal{M}_R$, the *join set* is the image of

$$\begin{aligned}\Sigma : \mathcal{M}_1 \times \cdots \times \mathcal{M}_R &\mapsto \mathbb{R}^{n_1 \times \cdots \times n_D}\\ \mathcal{A}_1, \dots, \mathcal{A}_R &\mapsto \mathcal{A}_1 + \cdots + \mathcal{A}_R.\end{aligned}$$

Decomposition problem: given $\mathcal{A}$ in the join set, solve $\Sigma(\mathcal{A}_1, \dots, \mathcal{A}_R) = \mathcal{A}$.

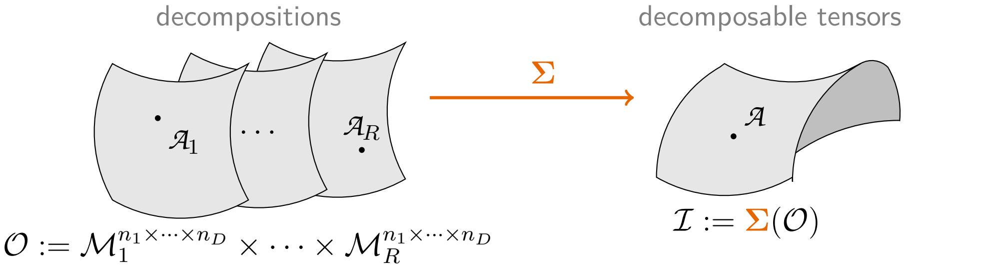
$\mathcal{M}_1, \dots, \mathcal{M}_R$	decomposition
Segre manifold	polyadic decomposition
Tucker manifold	block term decomposition
TT manifold	SoTT decomposition
structured Tucker manifold	*structured block term decomposition (SBTD)*

3 Outline

1. Condition: the main concept
2. Tensor manifolds and join sets
3. Condition of additive decompositions
4. Compression and invariance of the condition number

3 Finding the decomposition map geometrically

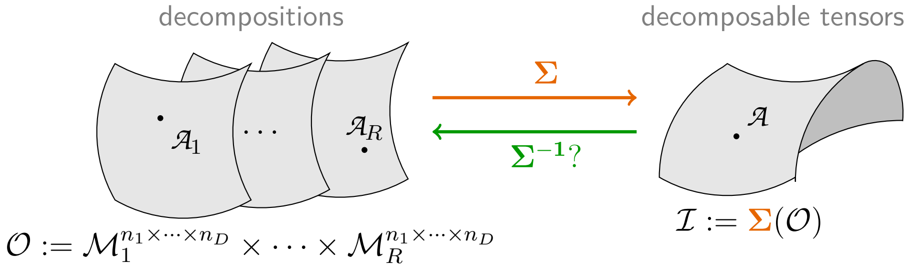
The easy version of the problem is *evaluating* the decomposition: $\mathcal{A} = \sum_{\mathbf{r=1}}^{\mathbf{R}} \mathcal{A}_r$.



KU LEUVEN

3 Finding the decomposition map geometrically

The easy version of the problem is *evaluating* the decomposition: $\mathcal{A} = \sum_{\mathbf{r}=\mathbf{1}}^{\mathbf{R}} \mathcal{A}_r$.



But what about *decomposing* $\mathcal{A}$?
By the inverse function theorem, $\exists$ smooth local map $\boldsymbol{\Sigma}^{-1}$ with closed-form derivative [Breiding & Vannieuwenhoven 2018], similar to [Bürgisser & Cucker 2013].

KU LEUVEN

3 Algorithm to compute the condition number

Given: $\mathcal{A}_1 \in \mathcal{M}_1^{n_1,\dots,n_D}, \dots, \mathcal{A}_R \in \mathcal{M}_R^{n_1,\dots,n_D}$ so that $\mathcal{A} = \sum_{r=1}^R \mathcal{A}_r$.

By [Breiding & Vannieuwenhoven 2018], the following computes the condition number of the decomposition map $\mathbf{\Sigma}^{-1}\colon \mathcal{A} \mapsto (\mathcal{A}_1, \dots, \mathcal{A}_R)$ at $\mathcal{A}$.

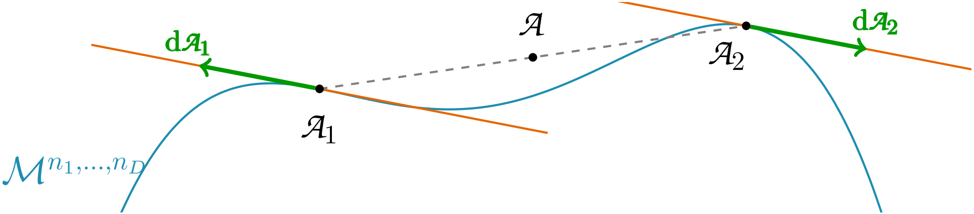
1. For $r = 1, \dots, R$, construct a matrix T_r containing an orthonormal basis of $T_{\mathcal{A}_r}\mathcal{M}_r^{n_1,\dots,n_D}$.
2. $\kappa = \sigma_{\min}\left(\begin{bmatrix} T_1 & \cdots & T_R \end{bmatrix}\right)^{-1}$.

An expression for step 1 was provided by [D, Breiding & Vannieuwenhoven 2021].

KU LEUVEN

3 Geometric intuition behind the condition number

κ measures the collinearity between the tangent spaces. E.g., let $\mathcal{A} = \frac{1}{2}\mathcal{A}_1 + \frac{1}{2}\mathcal{A}_2$:



Suppose there are almost parallel tangent vectors $\mathrm{d}\mathcal{A}_1 \approx -\mathrm{d}\mathcal{A}_2$. Then

1. A tiny nudge to $\mathcal{A}$ can push $\mathcal{A}_1$ and $\mathcal{A}_2$ far away in opposite directions.
2. The derivative of $(\mathcal{A}_1, \mathcal{A}_2) \mapsto \frac{1}{2}\mathcal{A}_1 + \frac{1}{2}\mathcal{A}_2 - \mathcal{A}$ is singular at the solution.

KU LEUVEN

3 Revisiting the bad tensor

Take our example from the beginning:

$$\mathcal{A} = \left[\!\left[\begin{bmatrix} & 1 & \\ 1 & & \\ & & \end{bmatrix} \begin{bmatrix} 1 & & \varepsilon \\ & & \\ \varepsilon & & \varepsilon^2 \end{bmatrix} \begin{bmatrix} & \varepsilon & \\ & & \\ & & \end{bmatrix} \right]\!\right].$$

For $\varepsilon = 10^{-6}$ the unique tensor rank decomposition $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3$ has $\kappa \approx 10^{12}$.

Suppose $\widetilde{\mathcal{A}} = \widetilde{\mathcal{A}}_1 + \widetilde{\mathcal{A}}_2 + \widetilde{\mathcal{A}}_3$ is an approximation of $\mathcal{A}$.

To get $\left\|(\widetilde{\mathcal{A}}_1, \widetilde{\mathcal{A}}_2, \widetilde{\mathcal{A}}_3) - (\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3)\right\| \le 10^{-p}$, we may need $\left\|\widetilde{\mathcal{A}} - \mathcal{A}\right\| \lesssim 10^{-12-p}$.

4 Outline

1. Condition: the main concept
2. Tensor manifolds and join sets
3. Condition of additive decompositions
4. Compression and invariance of the condition number

4 Compression

Many tensors in applications live in the tensor product of low-dimensional subspaces $\mathbb{V}_1, \ldots, \mathbb{V}_D$ with $\dim(\mathbb{V}_i) = m_i$. In coordinates, $\exists \mathcal{G} \in \mathbb{R}^{m_1 \times \cdots \times m_D}$ so that

$$\mathcal{A} = (\mathbf{Q_1} \otimes \cdots \otimes \mathbf{Q_D})\mathcal{G} \quad \text{where} \quad \operatorname{span}(\mathbf{Q_i}) = \mathbb{V}_i.$$

The SBTD is defined precisely so that the decompositions of $\mathcal{A}$ and $\mathcal{G}$ can be converted into each other:

$$\mathcal{G} = \sum_{r=1}^{R} (U_1 \otimes \cdots \otimes U_D)\mathcal{C}_r \quad \Rightarrow \quad \mathcal{A} = \sum_{r=1}^{R} (\mathbf{Q_1} U_1 \otimes \cdots \otimes \mathbf{Q_D} U_D)\mathcal{C}_r.$$

Conversely, if $\mathcal{A}$ has an SBTD, so does $\mathcal{G}$.

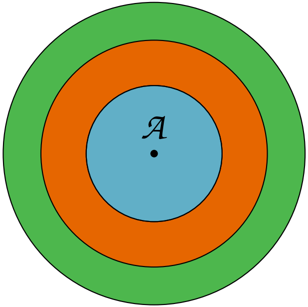
4 Algorithm to decompose tensors by compression

1. *(Compress)* Identify the lowest dimensional subspaces $\mathbb{V}_1, \ldots, \mathbb{V}_D$ so that $\mathcal{A} \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_D$. In coordinates: $\mathcal{A} = (\mathbf{Q_1} \otimes \cdots \otimes \mathbf{Q_D})\mathcal{G}$
2. *(Decompose)* Decompose $\mathcal{G} = \sum_{r=1}^{R} (U_1 \otimes \cdots \otimes U_D)\mathcal{C}_r$.
3. *(Expand)* Convert step 2 to $\mathcal{A} = \sum_{r=1}^{R} (\mathbf{Q_1} U_1 \otimes \cdots \otimes \mathbf{Q_D} U_D)\mathcal{C}_r$.

We often use orthonormal bases $Q_1, \ldots, Q_D$.

4 The importance of the domain

Tensors $\mathcal{A}$ for which this algorithm works are special in several ways:



- $\mathcal{I}_1$: decomposable
- $\mathcal{I}_2$: decomposition in an $(m_1, \ldots, m_D)$-dimensional tensor subspace
- $\mathcal{I}_3$: decomposition in $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_D$

$\mathcal{I}_3 \subseteq \mathcal{I}_2 \subseteq \mathcal{I}_1$

Recall the definition of the condition number of a map $f : \boldsymbol{\mathcal{I}} \to \mathcal{O}$:

$$\kappa[f](\mathcal{A}) := \limsup_{\boldsymbol{\mathcal{I}} \ni \widetilde{\mathcal{A}} \to \mathcal{A}} \frac{d(f(\mathcal{A}), f(\widetilde{\mathcal{A}}))}{d(\mathcal{A}, \widetilde{\mathcal{A}})}.$$

4 The importance of the domain

- $\mathcal{I}_1$: decomposable
- $\mathcal{I}_2$: decomposition in an $(m_1, \ldots, m_D)$-dimensional tensor subspace
- $\mathcal{I}_3$: decomposition in $\mathbb{V}_\mathbf{1} \otimes \cdots \otimes \mathbb{V}_{\boldsymbol{D}}$

Let f be the SBTD decomposition map:

$$\kappa[f](\mathcal{A}) := \limsup_{\boldsymbol{\mathcal{I}} \ni \widetilde{\mathcal{A}} \to \mathcal{A}} \frac{d(f(\mathcal{A}), f(\widetilde{\mathcal{A}}))}{d(\mathcal{A}, \widetilde{\mathcal{A}})}.$$

If $\mathcal{A} = (Q_1 \otimes \cdots \otimes Q_D)\mathcal{G}$ and $\mathcal{G}$ has an SBTD, then $\kappa[f](\mathcal{G})$ corresponds to the case $\boldsymbol{\mathcal{I}} = \boldsymbol{\mathcal{I}_3}$.
However, $\kappa[f](\mathcal{A})$ corresponds to $\boldsymbol{\mathcal{I}} = \boldsymbol{\mathcal{I}_1} \supseteq \boldsymbol{\mathcal{I}_3}$.

4 Invariance of the condition number

Theorem (D, Breiding & Vannieuwenhoven 2021)

Suppose $\mathcal{A} = \sum_{r=1}^{R} \mathcal{A}_r$ *and* $\mathcal{G} = \sum_{r=1}^{R} \mathcal{G}_r$ *are two SBTDs related by* $\mathcal{A}_r = (Q_1, \ldots, Q_D) \cdot \mathcal{G}_r$ *for some matrices* $Q_1, \ldots, Q_D$ *with orthonormal columns. Then these two SBTDs have the same condition number.*

Bad news: making the problem smaller does not make it numerically easier.
Good news: properties of a large SBTD can be revealed through small-scale calculations.

4 Fast computation of the condition number: example

Let's take a tensor rank decomposition from chemometrics [Bro & Andersson 1998]:

$$\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3 \in \mathbb{R}^{265\times371\times7}.$$

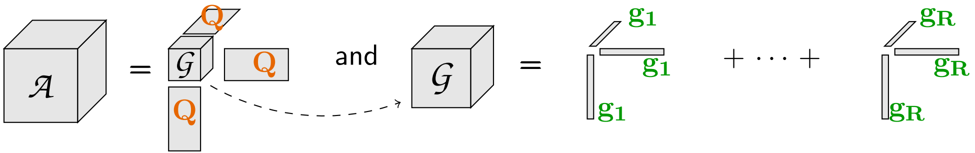
Computing the condition number takes 110 seconds. However, since $\mathcal{A}$ has rank 3, it can be compressed to

$$\mathcal{G} = \mathcal{G}_1 + \mathcal{G}_2 + \mathcal{G}_3 \in \mathbb{R}^{3\times3\times3}.$$

Compressing and computing the condition number of $(\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3)$ takes 6.9 milliseconds (combined)!

4 Compressed symmetric decompositions

The symmetric decomposition expresses $\mathcal{A} = \mathbf{a}_1^{\otimes D} + \cdots + \mathbf{a}_R^{\otimes D}$ where $\mathbf{a}^{\otimes D} = \mathbf{a} \otimes \cdots \otimes \mathbf{a}$.
Similar compression methods exist in which



$$\mathcal{A} = \mathbf{Q}^{\otimes \mathbf{D}} \sum_{r=1}^{R} \mathbf{g_r}^{\otimes \mathbf{D}} = \sum_{r=1}^{R} (\mathbf{Q}\mathbf{g_r})^{\otimes D}$$

so that $\mathcal{A}$ and $\mathcal{G}$ have equivalent decompositions.

4 Several ways to quantify the condition number

Let $\mathcal{A} = Q^{\otimes D}\mathcal{G}$ for some orthonormal basis Q.
There are 4 condition numbers, each of which accounts for some perturbation $\widetilde{\mathcal{A}}$ or $\widetilde{\mathcal{G}}$.

$\kappa_{\mathcal{A}}^{\text{sym}}$:	$\widetilde{\mathcal{A}} = \sum_{r=1}^{R} \tilde{\mathbf{a}}_r^{\otimes D}$	$\kappa_{\mathcal{G}}^{\text{sym}}$:	$\widetilde{\mathcal{G}} = \sum_{r=1}^{R} \tilde{\mathbf{g}}_r^{\otimes D}$
$\kappa_{\mathcal{A}}^{\text{cpd}}$:	$\widetilde{\mathcal{A}} = \sum_{r=1}^{R} \tilde{\mathbf{a}}_r^{1} \otimes \cdots \otimes \tilde{\mathbf{a}}_r^{D}$	$\kappa_{\mathcal{G}}^{\text{cpd}}$:	$\widetilde{\mathcal{G}} = \sum_{r=1}^{R} \tilde{\mathbf{g}}_r^{1} \otimes \cdots \otimes \tilde{\mathbf{g}}_r^{D}$

Theorem 1: $\kappa_{\mathcal{G}}^{\text{sym}} = \kappa_{\mathcal{A}}^{\text{sym}}$ if $\mathcal{G}$ is not minimal.
Theorem 2: $\kappa_{\mathcal{G}}^{\text{sym}} \leq \kappa_{\mathcal{A}}^{\text{sym}} \leq \sqrt{D}\kappa_{\mathcal{G}}^{\text{sym}}$.
Theorem 3: $\kappa_{\mathcal{A}}^{\text{sym}} = \kappa_{\mathcal{A}}^{\text{cpd}}$ if $R \leq 2$.
Conjecture: all equal!

5 References I

Blum, L., F. Cucker, M. Shub & S. Smale (1998). *Complexity and Real Computation*. en. New York, NY: Springer New York. ISBN: 978-1-4612-6873-4 978-1-4612-0701-6. DOI: 10.1007/978-1-4612-0701-6. URL: http://link.springer.com/10.1007/978-1-4612-0701-6 (visited on 08/23/2022).

Breiding, P. & N. Vannieuwenhoven (Jan. 2018). "The Condition Number of Join Decompositions". In: *SIAM Journal on Matrix Analysis and Applications* 39.1, pp. 287–309. ISSN: 0895-4798. DOI: 10.1137/17M1142880. URL: https://epubs.siam.org/doi/10.1137/17M1142880.

Bro, R. & C. A. Andersson (1998). "Improving the speed of multiway algorithms part II: Compression". In: *Chemometrics and Intelligent Laboratory Systems* 42.1-2, pp. 105–113. ISSN: 01697439. DOI: 10.1016/S0169-7439(98)00011-2.

Bürgisser, P. & F. Cucker (2013). *Condition*. Vol. 349. Series Title: Grundlehren der mathematischen Wissenschaften Publication Title: Media. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN: 978-3-642-38895-8.

KU LEUVEN

5 References II

- D, P. Breiding & N. Vannieuwenhoven (2021). "The condition number of many tensor decompositions is invariant under Tucker compression". In: pp. 1–18. arXiv: 2106.13034. URL: http://arxiv.org/abs/2106.13034.
- Eckart, C. & G. Young (Sept. 1936). "The approximation of one matrix by another of lower rank". en. In: *Psychometrika* 1.3, pp. 211–218. ISSN: 0033-3123, 1860-0980. DOI: 10.1007/BF02288367. URL: http://link.springer.com/10.1007/BF02288367 (visited on 10/19/2022).
- Nocedal, J. & S. J. Wright (2006). *Numerical Optimization*. New York: Springer New York. ISBN: 978-0-387-30303-1. DOI: 10.1007/978-0-387-40065-5. URL: http://link.springer.com/10.1007/978-0-387-40065-5.
- Rice, J. R. (June 1966). "A Theory of Condition". In: *SIAM Journal on Numerical Analysis* 3.2, pp. 287–310. ISSN: 0036-1429. DOI: 10.1137/0703023. URL: http://epubs.siam.org/doi/10.1137/0703023.

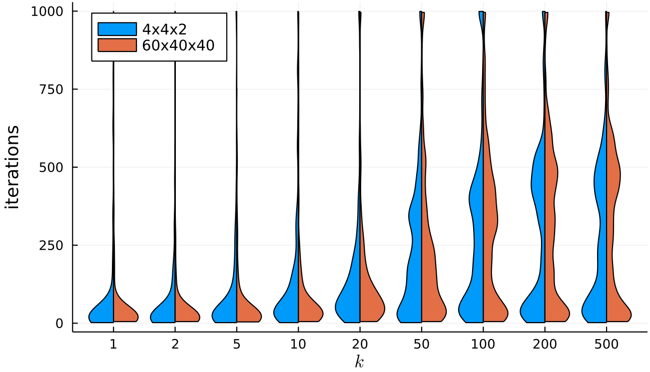
5 References III

Schmidt, E. (Dec. 1907). "Zur Theorie der linearen und nichtlinearen Integralgleichungen: I. Teil: Entwicklung willkuerlicher Funktionen nach Systemen vorgeschriebener". de. In: *Mathematische Annalen* 63.4, pp. 433–476. ISSN: 0025-5831, 1432-1807. DOI: 10.1007/BF01449770. URL: http://link.springer.com/10.1007/BF01449770 (visited on 10/19/2022).

Vervliet, N., O. Debals & L. De Lathauwer (2017). "Tensorlab 3.0 - Numerical optimization strategies for large-scale constrained and coupled matrix/tensor factorization". In: *Conference Record - Asilomar Conference on Signals, Systems and Computers* 32, pp. 1733–1738. ISSN: 10586393. DOI: 10.1109/ACSSC.2016.7869679.

5 Iterative methods

The invariance of the condition number indicates that the convergence rate of iterative algorithms is unaffected by compression, even though the search space is reduced.



Performance of Tensorlab's `ll1_nls` algorithm [Vervliet, Debals & De Lathauwer 2017] to compute the two-term $(2, 2, 1)$ block term decomposition.
As k increases from 1 to 500, the condition number increases from ~ 10 to $\sim 10^8$ on average

KU LEUVEN