Numerical linear algebra and data analysis in large dimensions using tensor format

Ph.D. defence

Centre Inria of the University of Bordeaux, Talence, 09-12-2022

Martina Iannacito

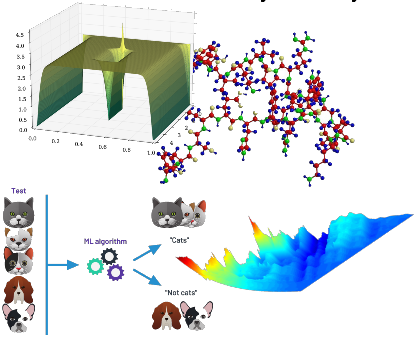
Supervised by O. Coulaud and L. Giraud

Concace - Inria joint team with Airbus Central R&T and Cerfacs

The context

Different disciplines as physics, computer science, chemistry... relay on numerical simulations to study

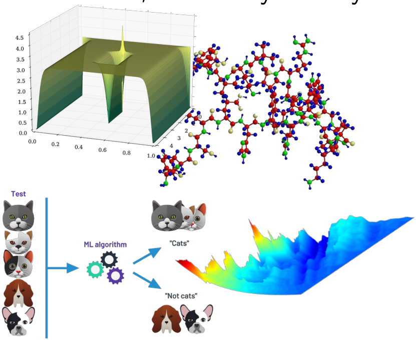
- Stochastic equations
- Uncertainty quantification problems
- Quantum and vibration chemistry
- Optimization
- Machine learning



The context

Different disciplines as physics, computer science, chemistry... relay on numerical simulations to study

- Stochastic equations
- Uncertainty quantification problems
- Quantum and vibration chemistry
- Optimization
- Machine learning

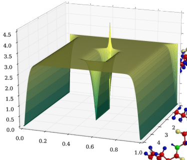


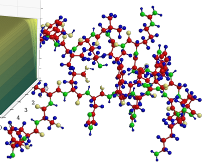
Usually studied with **matrix** linear algebra

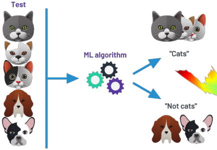
$$\begin{bmatrix} 1 & 8 & 4 \\ 9 & 2 & 2 \\ 7 & 1 & 6 \end{bmatrix}$$

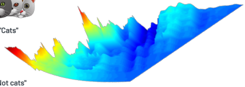
Different disciplines as physics, computer science, chemistry... relay on numerical simulations to study

- Stochastic equations
- Uncertainty quantification problems
- Quantum and vibration chemistry
- Optimization
- Machine learning





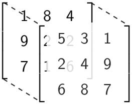




Usually studied with **matrix** linear algebra

$$\begin{bmatrix} 1 & 8 & 4 \\ 9 & 2 & 2 \\ 7 & 1 & 6 \end{bmatrix}$$

since the last century → **tensor**



Inria

- Better representation of structured problems and data

Tensors

+ Better representation of structured problems and data
- Curse of dimensionality

\+ Better representation of structured problems and data

\- Curse of dimensionality

Over the years different approximation techniques were proposed

- Canonical Polyadic
- Tucker
- Hierarchical Tucker
- Tensor-Train

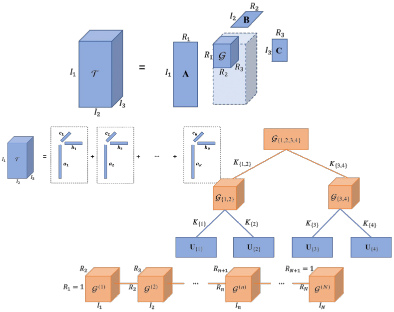


Figure: from [Bi et al. 2022]

+ Better representation of structured problems and data
− Curse of dimensionality

Over the years different approximation techniques were proposed

- Canonical Polyadic
- Tucker
- Hierarchical Tucker
- Tensor-Train

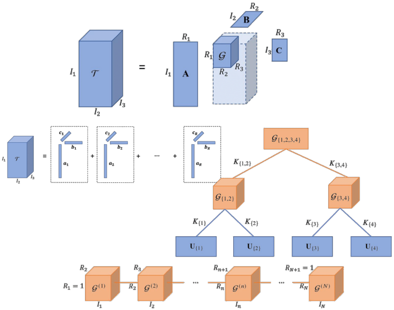


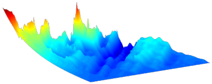
Figure: from [Bi et al. 2022]

These approximation techniques introduce compression errors, so **what are their effects inside classical algorithms?**

What are the effects of tensor representation and compression inside classical algorithms?

What are the effects of tensor representation and compression inside classical algorithms?

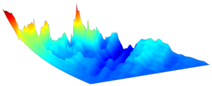
Numerical linear algebra



- How to solve a high dimensional linear system of equations, represented by low rank tensors?
- How to construct an orthogonal basis of a tensor subspace?
- What is the effect of tensor compression on the final solution?

What are the effects of tensor representation and compression inside classical algorithms?

Numerical linear algebra



- How to solve a high dimensional linear system of equations, represented by low rank tensors?
- How to construct an orthogonal basis of a tensor subspace?
- What is the effect of tensor compression on the final solution?

- How to relate and interpret point clouds from tensor data?
- Is it mathematically meaningful to visualize simultaneously more than two point clouds?

Data analysis

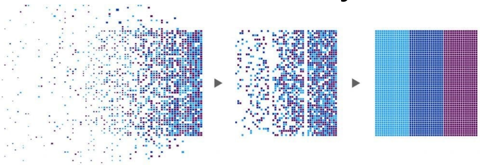


Table of contents

2

Numerical linear algebra

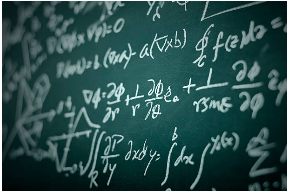
- The variable accuracy approach
- Generalized Minimal RESidual (GMRES)
  - > numerical results in matrix computation
  - > numerical results in Tensor-Train (TT) format
- Orthogonalization kernels
  - > TT-algorithms
  - > numerical loss of orthogonality in TT-format

Mathematical world

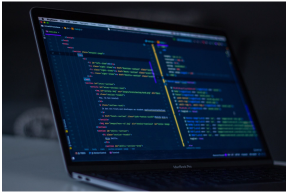
- $\pi = 3.14159265358979323846264 33...$

Computer world

```
>>> π = 3.141592653589793
```



?



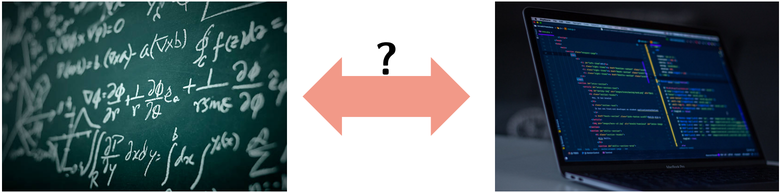
Mathematical world

- $\pi = 3.141592653589793238462643\ldots$
- $x = 0.1$ and $y = 0.2$, then $x + y = 0.3$

Computer world

>>> $\overline{\pi}$ = 3.141592653589793

>>> $\overline{x}$ = 0.1 and $\overline{y}$ = 0.2, then $\overline{x+y}$ = 0.30000000000000004



Representation and computational error

Denoting by u the **unit roundoff** of the working precision

Standard IEEE model [Higham 2002]

$$fl(x) = x(1+\xi) \qquad \text{[storage perturbation]}$$

$$fl(x \operatorname{op} y) = (x \operatorname{op} y)(1+\varepsilon) \qquad \text{[computational perturbation]}$$

with $|\xi| \leq u$, $|\varepsilon| \leq u$ and $\operatorname{op} \in \{+, -, \times, \div\}$.

Inria

Representation and computational error

Denoting by u the **unit roundoff** of the working precision

Standard IEEE model [Higham 2002]

$$fl(x) = x(1+\xi) \qquad \text{[storage perturbation]}$$
$$fl(x \operatorname{op} y) = (x \operatorname{op} y)(1+\varepsilon) \qquad \text{[computational perturbation]}$$

with $|\xi| \leq u$, $|\varepsilon| \leq u$ and $\operatorname{op} \in \{+, -, \times, \div\}$.

Example

Assuming to work in floating point 64, with $u_{64} = 10^{-16}$

- $\overline{\pi} = 3.141592653589793 = \pi(1+\xi)$ with $|\xi| \leq u_{64}$
- $\overline{\texttt{x}} = \texttt{0.1}$ and $\overline{\texttt{y}} = \texttt{0.2}$, then

$$\overline{\texttt{x + y}} = \texttt{0.30000000000000004} = (0.2 + 0.1)(1+\varepsilon)$$

with $|\varepsilon| \leq u_{64}$

The variable accuracy approach

The storage precision and the computational one are decoupled

The variable accuracy approach

The storage precision and the computational one are decoupled

δ-componentwise storage

$$\delta\text{-storage}(x) = x(1+\xi)$$

with $|\xi| \leq \delta$ with δ the storage precision.

Standard IEEE model [Higham 2002]

$$fl(x \operatorname{op} y) = (x \operatorname{op} y)(1+\varepsilon)$$

with $|\varepsilon| \leq u$ with u the computation precision.

The variable accuracy approach

The storage precision and the computational one are decoupled

δ-componentwise storage

$$\delta\text{-storage}(x) = x(1+\xi)$$

with $|\xi| \leq \delta$ with δ the storage precision.

Standard IEEE model [Higham 2002]

$$fl(x \operatorname{op} y) = (x \operatorname{op} y)(1+\varepsilon)$$

with $|\varepsilon| \leq u$ with u the computation precision.

δ-storage

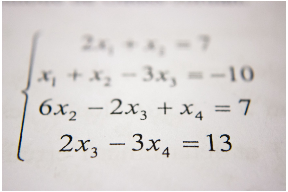
$$fl_\delta(x \operatorname{op} y) = \delta\text{-storage}\big(fl(x \operatorname{op} y)\big) = (x \operatorname{op} y)(1+\varepsilon+\xi)$$

with $|\varepsilon| \leq u$, $|\xi| \leq \delta$ and $\operatorname{op} \in \{+,-,\times,\div\}$.

Iterative solver

- Generalized Minimal RESidual (GMRES)

in matrix and tensor format



Orthogonalization kernels

- Classical and Modified Gram-Schmidt (CGS, MGS)
- Gram approach
- Householder transformation

only in tensor format

Background on GMRES [Saad and Schultz 1986]

To solve $Ax = b$ with initial guess $x_0 = 0$, at the k-th iteration GMRES minimizes the norm of residual

$$||r_k|| = \min_{x \in \mathcal{K}_k(A,b)} ||Ax - b||$$

in the Krylov space $\mathcal{K}_k(A, b) = \mathrm{span}\{b, Ab, \ldots, A^{k-1}b\}$

Background on GMRES [Saad and Schultz 1986]

To solve $Ax = b$ with initial guess $x_0 = 0$, at the k-th iteration GMRES minimizes the norm of residual

$$||r_k|| = \min_{x \in \mathcal{K}_k(A,b)} ||Ax - b||$$

in the Krylov space $\mathcal{K}_k(A, b) = \text{span}\{b, Ab, \dots, A^{k-1}b\}$
Practically, let $V_k = [v_1, \dots, v_k]$ such that

$$\mathcal{K}_k(A, b) = \text{span}\{v_1, \dots, v_k\}$$

thanks to the **Arnoldi relation**

$$AV_k = V_{k+1}\overline{H}_k \qquad \text{with} \qquad V_{k+1}^\top V_{k+1} = \mathbb{I}_{k+1}$$

Background on GMRES [Saad and Schultz 1986]

To solve $Ax = b$ with initial guess $x_0 = 0$, at the k-th iteration GMRES minimizes the norm of residual

$$||r_k|| = \min_{x \in \mathcal{K}_k(A,b)} ||Ax - b||$$

in the Krylov space $\mathcal{K}_k(A, b) = \text{span}\{b, Ab, \dots, A^{k-1}b\}$
Practically, let $V_k = [v_1, \dots, v_k]$ such that

$$\mathcal{K}_k(A, b) = \text{span}\{v_1, \dots, v_k\}$$

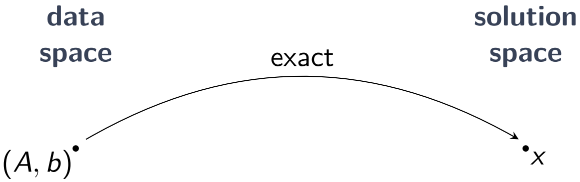
thanks to the **Arnoldi relation**

$$AV_k = V_{k+1}\overline{H}_k \qquad \text{with} \qquad V_{k+1}^\top V_{k+1} = \mathbb{I}_{k+1}$$

Commonly Householder or Modified Gram-Schmidt algorithms are used to construct V_k

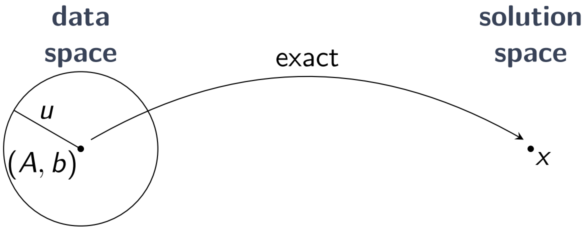
Backward error analysis [Wilkinson 1963]

Given the linear system $Ax = b$ and a working precision u, then



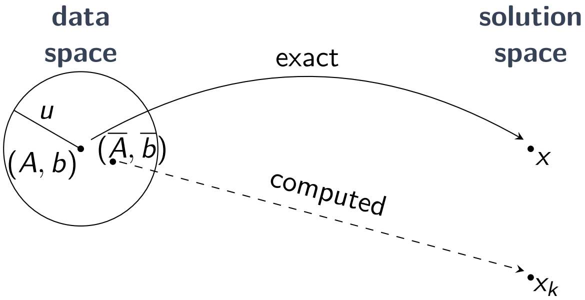
Backward error analysis [Wilkinson 1963]

Given the linear system $Ax = b$ and a working precision u, then



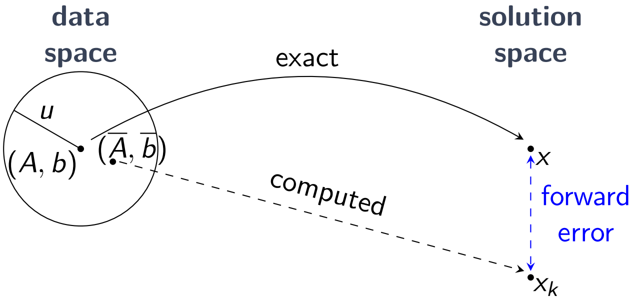
Backward error analysis [Wilkinson 1963]

Given the linear system $Ax = b$ and a working precision u, then



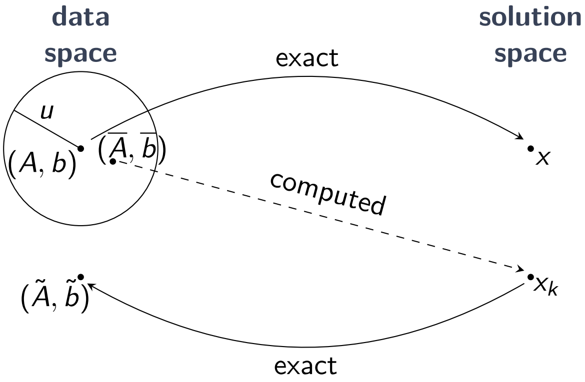
Backward error analysis [Wilkinson 1963]

Given the linear system $Ax = b$ and a working precision u, then

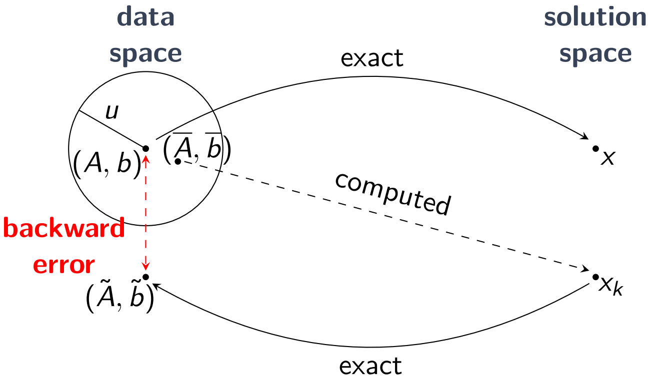


Backward error analysis [Wilkinson 1963]

Given the linear system $Ax = b$ and a working precision u, then

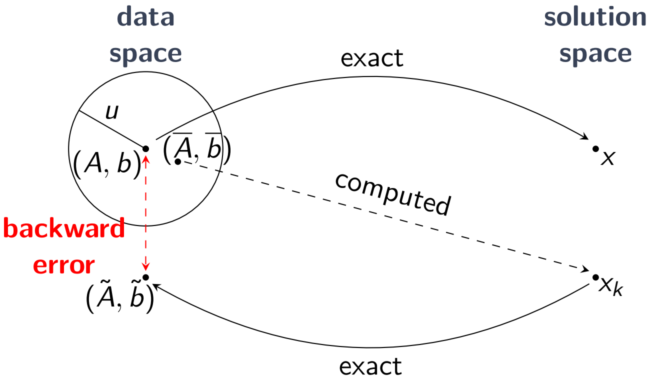


Given the linear system $Ax = b$ and a working precision u, then



Backward error analysis [Wilkinson 1963]

Given the linear system $Ax = b$ and a working precision u, then



If the backward error

$$\eta_{A,b}(x_k) = \frac{\|Ax_k - b\|}{||A||||x_k|| + ||b||} \sim \mathcal{O}(u)$$

than the algorithm is backward stable

GMRES backward stability

Denoting by u the unit roundoff of the working precision for solving the linear system

$$Ax = b$$

then in [Drkosova, Greenbaum, Rozložník, and Strakoš 1995] and in [Paige, Rozložník, and Strakoš 2006], it is proved that

$$\eta_{A,b}(x_k) = \mathcal{O}(u)$$

GMRES backward stability

Denoting by u the unit roundoff of the working precision for solving the linear system

$$Ax = b$$

then in [Drkosova, Greenbaum, Rozložník, and Strakoš 1995] and in [Paige, Rozložník, and Strakoš 2006], it is proved that

$$\eta_{A,b}(x_k) = \mathcal{O}(u)$$

$$\left\|\mathbb{I}_k - V_k^\top V_k\right\| \eta_{A,b}(x_k) = \mathcal{O}(u)$$

GMRES backward stability

Denoting by u the unit roundoff of the working precision for solving the linear system

$$Ax = b$$

then in [Drkosova, Greenbaum, Rozložník, and Strakoš 1995] and in [Paige, Rozložník, and Strakoš 2006], it is proved that

$$\eta_{A,b}(x_k) = \mathcal{O}(u)$$
$$\left\|\mathbb{I}_k - V_k^\top V_k\right\| \eta_{A,b}(x_k) = \mathcal{O}(u)$$

So, the backward error $\eta_{A,b}$ is suggested as **stopping criterion** for GMRES

GMRES backward stability

Denoting by u the unit roundoff of the working precision for solving the linear system

$$Ax = b$$

then in [Drkosova, Greenbaum, Rozložník, and Strakoš 1995] and in [Paige, Rozložník, and Strakoš 2006], it is proved that

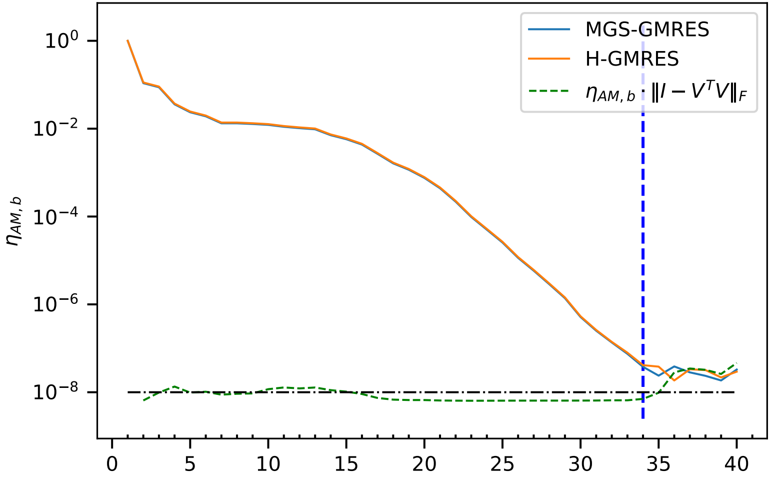
$$\eta_{A,b}(x_k) = \mathcal{O}(u)$$
$$\left\|\mathbb{I}_k - V_k^\top V_k\right\| \eta_{A,b}(x_k) = \mathcal{O}(u)$$

So, the backward error $\eta_{A,b}$ is suggested as **stopping criterion** for GMRES

For MGS-GMRES the backward stability holds if $k \in \mathbb{N}$ is such that

$$\kappa(V_{k+1}) > \frac{4}{3}$$

32bit floating point arithmetic

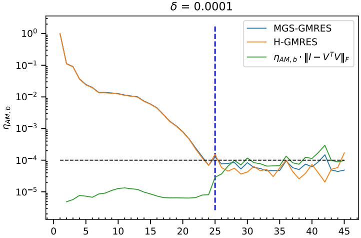


δ-componentwise storage

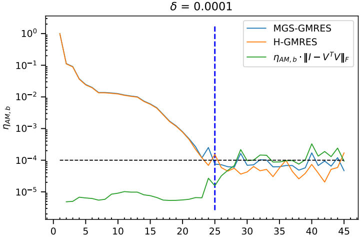
Any vector $z \in \mathbb{R}^n$ is replaced by $\bar{z} \in \mathbb{R}^n$ such that

$$fl\big(z(i)\big) = \bar{z}(i) \qquad \text{such that} \qquad \frac{|z(i) - \bar{z}(i)|}{|z(i)|} \leq \delta.$$

fp32 computation



fp64 computation



$$\delta = 10^{-4}$$

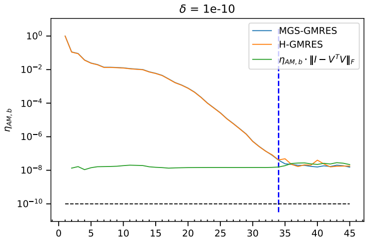
Numerical experiments δ-componentwise storage

δ-componentwise storage

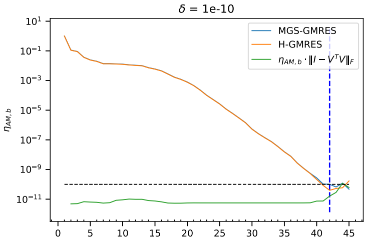
Any vector $z \in \mathbb{R}^n$ is replaced by $\bar{z} \in \mathbb{R}^n$ such that

$$fl\big(z(i)\big) = \bar{z}(i) \quad \text{such that} \quad \frac{|z(i) - \bar{z}(i)|}{|z(i)|} \leq \delta.$$

fp32 computation



fp64 computation



$$\delta = 10^{-10}$$

δ-normwise storage model

Length n vectors are stored in compressed format with

δ-component storage

$$\frac{|z(i) - \bar{z}(i)|}{|z|} \leq \delta.$$

Remark

[Drkosova, Greenbaum, Rozložník, and Strakoš 1995; Paige, Rozložník, and Strakoš 2006]

do

readily apply

δ-normwise storage model

Length n vectors are stored in compressed format with

δ-component storage

$$\frac{|z(i) - \bar{z}(i)|}{|z|} \leq \delta.$$

δ-normwise storage

$$\frac{||z - \bar{z}||}{||z||} \leq \delta.$$

Remark

[Drkosova, Greenbaum, Rozložník, and Strakoš 1995; Paige, Rozložník, and Strakoš 2006]

do

readily apply

Remark

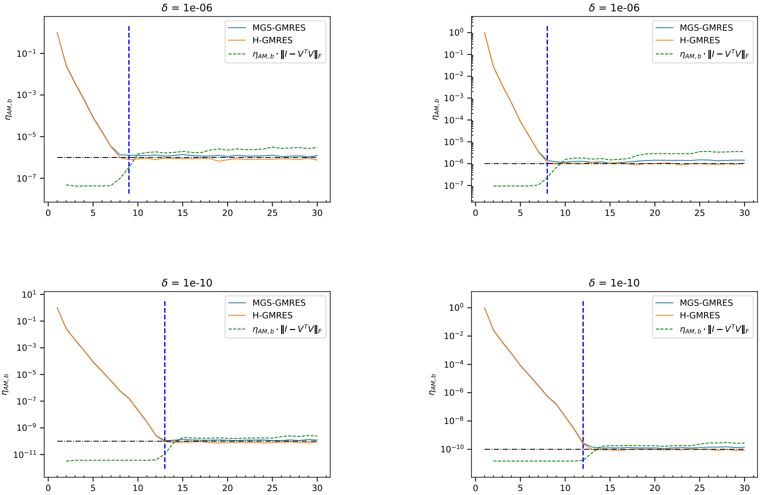
[Drkosova, Greenbaum, Rozložník, and Strakoš 1995; Paige, Rozložník, and Strakoš 2006]

do not

readily apply

Numerical experiment comparison

δ-componentwise δ-normwise



The δ-componentwise and δ-normwise perturbation lead to similar results.

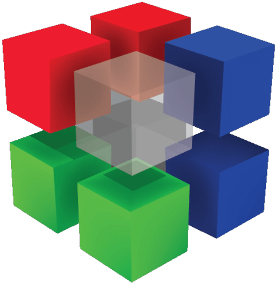
Inria

δ-normwise perturbation: practical cases

Practical cases where the variable accuracy approach with **normwise perturbation** find an application

- Lossy compressor

- Tensor problem



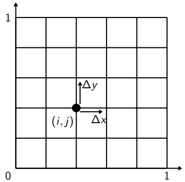
Context

The problem

$$\begin{cases} \mathcal{L}(u) & = f \quad \text{in} \quad \Omega \\ u & = f_0 \quad \text{in} \quad \partial\Omega \end{cases}$$

for $\Omega \subseteq \mathbb{R}^{n_1 \times \cdots \times n_d}$.

discretization

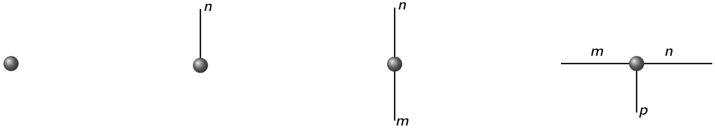


$$\mathbf{Ax} = \mathbf{b}$$

where $\mathbf{A} : \mathbb{R}^{n_1 \times \cdots \times n_d} \to \mathbb{R}^{n_1 \times \cdots \times n_d}$ is a multilinear operator and $\mathbf{b} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ a tensor.

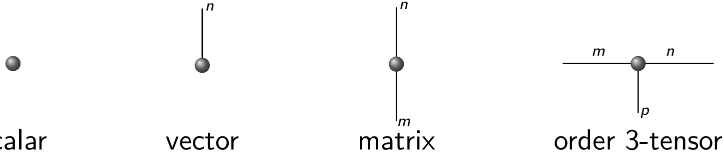
- Curse of dimensionality
- TT-formalism [Oseledets 2011]

Tensor Network formalism

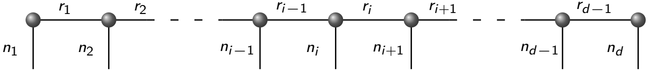


scalar vector matrix order 3-tensor

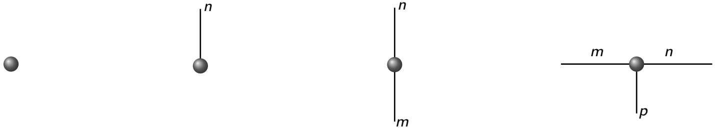
Tensor Network formalism



The TT-vector $\mathbf{x} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ becomes

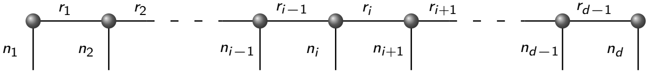


Tensor Network formalism

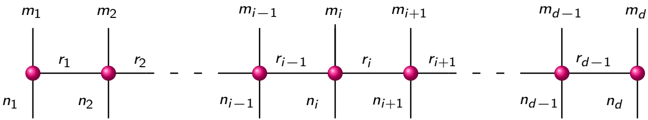


scalar vector matrix order 3-tensor

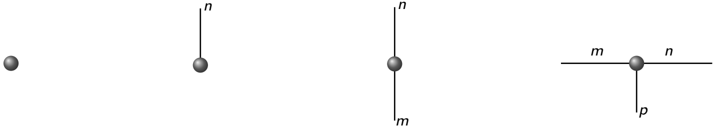
The TT-vector $\mathbf{x} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ becomes



The TT-matrix $\mathbf{A} \in \mathbb{R}^{(n_1 \times m_1) \cdots \times (n_d \times m_d)}$ becomes

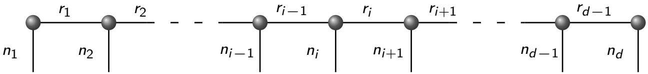


Tensor Network formalism

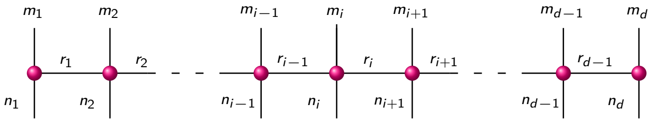


scalar | vector | matrix | order 3-tensor

The TT-vector $\mathbf{x} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ becomes

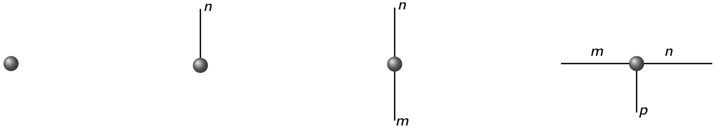


The TT-matrix $\mathbf{A} \in \mathbb{R}^{(n_1 \times m_1) \cdots \times (n_d \times m_d)}$ becomes



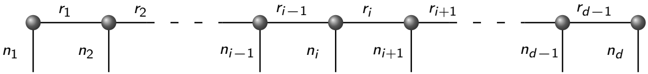
Summation and contraction increase the TT-rank r_i values

Tensor Network formalism

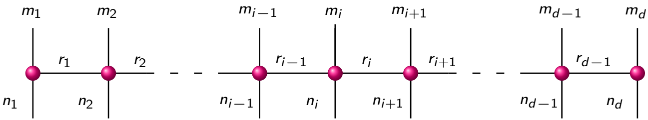


scalar vector matrix order 3-tensor

The TT-vector $\mathbf{x} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ becomes



The TT-matrix $\mathbf{A} \in \mathbb{R}^{(n_1 \times m_1) \cdots \times (n_d \times m_d)}$ becomes



Summation and contraction increase the TT-rank r_i values

TT-rounding [Oseledets 2011]

Robustness of TT-GMRES

TT-rounding at precision δ

For every TT-vector $\mathbf{z} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ we compress it getting $\overline{\mathbf{z}} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ such that

$$\frac{\|\mathbf{z} - \overline{\mathbf{z}}\|}{\|\mathbf{z}\|} \leq \delta.$$

Robustness of TT-GMRES

TT-rounding at precision δ

For every TT-vector $\mathbf{z} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ we compress it getting $\bar{\mathbf{z}} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ such that

$$\frac{\|\mathbf{z} - \bar{\mathbf{z}}\|}{\|\mathbf{z}\|} \leq \delta.$$

Convection-Diffusion problem

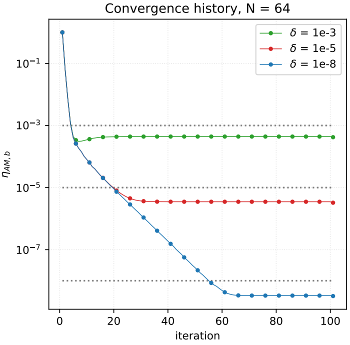
$$\begin{cases} -\Delta \mathbf{u} + \mathbf{v} \cdot \nabla \mathbf{u} & = 0 \\ \mathbf{u}_{\{y=1\}} & = 1 \end{cases} \quad \text{in} \quad \Omega = [-1, 1]^3$$

Robustness of TT-GMRES

TT-rounding at precision δ

For every TT-vector $\mathbf{z} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ we compress it getting $\overline{\mathbf{z}} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ such that

$$\frac{\|\mathbf{z} - \overline{\mathbf{z}}\|}{\|\mathbf{z}\|} \leq \delta.$$



Convection-Diffusion problem

$$\begin{cases} -\Delta \mathbf{u} + \mathbf{v} \cdot \nabla \mathbf{u} & = 0 \\ \mathbf{u}_{\{y=1\}} & = 1 \end{cases} \quad \text{in} \quad \Omega = [-1, 1]^3$$

Parameter dependent problem

Let the parametric convection-diffusion problem be

$$\begin{cases} -\alpha \Delta \mathbf{u} + \mathbf{v} \cdot \nabla \mathbf{u} & = 0 \\ \mathbf{u}_{\{y=1\}} & = 1 \end{cases} \quad \text{in} \quad \Omega = [-1, 1]^d \text{ and } \alpha \in [1, 10]$$

Solve independently p tensor linear systems of order d

$$\mathbf{A}_{\alpha_i} \mathbf{x}_{\alpha_i} = \mathbf{b}_{\alpha_i}$$

$\forall \alpha_i \in \{\alpha_1, \ldots, \alpha_p\}$.

Parameter dependent problem

Let the parametric convection-diffusion problem be

$$\begin{cases} -\alpha\Delta\mathbf{u} + \mathbf{v}\cdot\nabla\mathbf{u} & = 0 \\ \mathbf{u}_{\{y=1\}} & = 1 \end{cases} \quad \text{in} \quad \Omega = [-1,1]^d \text{ and } \alpha \in [1,10]$$

Solve independently p tensor linear systems of order d

$$\mathbf{A}_{\alpha_i}\mathbf{x}_{\alpha_i} = \mathbf{b}_{\alpha_i}$$

$\forall \alpha_i \in \{\alpha_1, \dots, \alpha_p\}$.

Solve once the tensor linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ of order $(d+1)$

$$\begin{bmatrix} \mathbf{A}_{\alpha_1} & & \\ & \ddots & \\ & & \mathbf{A}_{\alpha_p} \end{bmatrix} \begin{bmatrix} \mathbf{x}^{[\alpha_1]} \\ \vdots \\ \mathbf{x}^{[\alpha_p]} \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{\alpha_i} \\ \vdots \\ \mathbf{b}_{\alpha_i} \end{bmatrix}$$

Parameter dependent problem

Let the parametric convection-diffusion problem be

$$\begin{cases} -\alpha\Delta\mathbf{u} + \mathbf{v}\cdot\nabla\mathbf{u} & = 0 \\ \mathbf{u}_{\{y=1\}} & = 1 \end{cases} \quad \text{in} \quad \Omega = [-1,1]^d \text{ and } \alpha \in [1,10]$$

Solve independently p tensor linear systems of order d

$$\mathbf{A}_{\alpha_i}\mathbf{x}_{\alpha_i} = \mathbf{b}_{\alpha_i}$$

$\forall\alpha_i \in \{\alpha_1, \ldots, \alpha_p\}$.

Solve once the tensor linear system $\mathbf{Ax} = \mathbf{b}$ of order $(d+1)$

$$\begin{bmatrix} \mathbf{A}_{\alpha_1} & & \\ & \ddots & \\ & & \mathbf{A}_{\alpha_p} \end{bmatrix} \begin{bmatrix} \mathbf{x}^{[\alpha_1]} \\ \vdots \\ \mathbf{x}^{[\alpha_p]} \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{\alpha_i} \\ \vdots \\ \mathbf{b}_{\alpha_i} \end{bmatrix}$$

What is the numerical quality of $\mathbf{x}^{[\alpha_i]}$ slice of $\mathbf{x}$ solution of the $(d+1)$ tensor linear system and by construction the solution of the problem $\mathbf{A}_{\alpha_i}\mathbf{x}_{\alpha_i} = \mathbf{b}_{\alpha_i}$?

Order d vs $(d+1)$ solution bounds

Let the backward error of the $(d+1)$ system $\mathbf{Ax} = \mathbf{b}$ be

$$\eta_{\mathbf{A},\mathbf{b}}(\mathbf{x}) = \frac{||\mathbf{b} - \mathbf{Ax}||}{||\mathbf{A}||||\mathbf{x}|| + ||\mathbf{b}||}.$$

Let $\mathbf{A}_i$, $\mathbf{b}_i$ be the multilinear operator and the right-hand side of the parametric problem for the parameter value α_i, then define

$$\eta_{\mathbf{A}_i,\mathbf{b}_i}(\mathbf{x}^{[i]}) = \frac{||\mathbf{b}_i - \mathbf{A}_i\mathbf{x}^{[i]}||}{||\mathbf{A}_i||||\mathbf{x}^{[i]}|| + ||\mathbf{b}_i||}$$

with $\mathbf{x}^{[i]}$ denotes the i-th slice of $\mathbf{x}$ on the parameter mode for $i \in \{1, \dots p\}$. Then we prove that

$$\eta_{\mathbf{A},\mathbf{b}}(\mathbf{x})\, \rho_i(\mathbf{x}) \geq \eta_{\mathbf{A}_i,\mathbf{b}_i}(\mathbf{x}^{[i]}) \qquad \text{where} \qquad \rho_i(\mathbf{x}) = \frac{\|\mathbf{A}\|\|\mathbf{x}\| + \sqrt{p}}{\|\mathbf{A}_i\mathbf{x}^{[i]}\| + 1}.$$

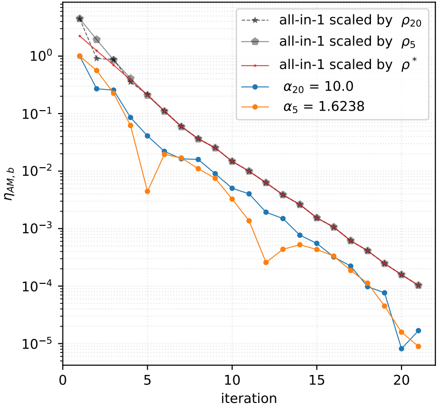


Figure: 4-d Parametric convection diffusion $\eta_{\mathbf{AM,b}}$ for and $n = 255$, bound using $\delta = \varepsilon = 10^{-5}$ and $p = 20$ uniformly logarithmically distributed parameter values $\alpha_i \in [1, 10]$

Numerical linear algebra summary I

- Backward stability of GMRES holds numerically when data are stored with
  - componentwise perturbations
  - normwise perturbations

  both have practical implementations
- GMRES in TT-format with TT-rounding seems to be backward stable
- Parameter dependent TT-problems of order d can be solved at once through an order $(d+1)$ problem, guaranteeing backward stable bounds linking the $(d+1)$ and d solutions
- Loss of orthogonality bounds for six widely used orthogonalization kernels appears to hold true when generalized to TT-format with normwise perturbation

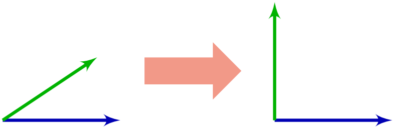
Inria

Application of the variable accuracy approach II

Iterative solver

- Generalized Minimal RESidual (GMRES)

in matrix and tensor format



Orthogonalization kernels

- Classical and Modified Gram-Schmidt (CGS, MGS)
- Gram approach
- Householder transformation

only in tensor format

Classical and Modified Gram-Schmidt

$Q, R = \mathtt{CGS}(A, \delta)$

Input: $A = [a_1, \ldots, a_m]$, $\delta \in \mathbb{R}_+$

for $i = 1, \ldots, m$ **do**
  $p = a_i$
  for $j = 1, \ldots, i-1$ **do**
    $R(i,j) = \langle a_i, q_j \rangle$
    $p = p - R(i,j) q_j$
  end
  $p = \delta\text{-}\mathtt{storage}(p)$
  $R(i,i) = ||p||$
  $q_i = 1/R(i,i)\, p$
end

Output: $Q = [q_1, \ldots, q_m]$, R

$Q, R = \mathtt{MGS}(A, \delta)$

Input: $A = [a_1, \ldots, a_m]$, $\delta \in \mathbb{R}_+$

for $i = 1, \ldots, m$ **do**
  $p = a_i$
  for $j = 1, \ldots, i-1$ **do**
    $R(i,j) = \langle p, q_j \rangle$
    $p = p - R(i,j) q_j$
  end
  $p = \delta\text{-}\mathtt{storage}(p)$
  $R(i,i) = ||p||$
  $q_i = 1/R(i,i)\, p$
end

Output: $Q = [q_1, \ldots, q_m]$, R

Classical and Modified Gram-Schmidt

$\mathcal{Q}, R = \texttt{CGS}(\mathcal{A}, \delta)$

Input: $\mathcal{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$, $\delta \in \mathbb{R}_+$

for $i = 1, \ldots, m$ **do**
  $\mathbf{p} = \mathbf{a}_i$
  for $j = 1, \ldots, i-1$ **do**
    $R(i,j) = \langle \mathbf{a}_i, \mathbf{q}_j \rangle$
    $\mathbf{p} = \mathbf{p} - R(i,j)\mathbf{q}_j$
  end
  $\mathbf{p} = \texttt{TT-rounding}(\mathbf{p}, \delta)$
  $R(i,i) = ||\mathbf{p}||$
  $\mathbf{q}_i = 1/R(i,i)\,\mathbf{p}$
end

Output: $\mathcal{Q} = \{\mathbf{q}_1, \ldots, \mathbf{q}_m\}$, R

$\mathcal{Q}, R = \texttt{MGS}(\mathcal{A}, \delta)$

Input: $\mathcal{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$, $\delta \in \mathbb{R}_+$

for $i = 1, \ldots, m$ **do**
  $\mathbf{p} = \mathbf{a}_i$
  for $j = 1, \ldots, i-1$ **do**
    $R(i,j) = \langle \mathbf{p}, \mathbf{q}_j \rangle$
    $\mathbf{p} = \mathbf{p} - R(i,j)\mathbf{q}_j$
  end
  $\mathbf{p} = \texttt{TT-rounding}(\mathbf{p}, \delta)$
  $R(i,i) = ||\mathbf{p}||$
  $\mathbf{q}_i = 1/R(i,i)\,\mathbf{p}$
end

Output: $\mathcal{Q} = \{\mathbf{q}_1, \ldots, \mathbf{q}_m\}$, R

They readily write in TT-format.

Householder transformation

Given a vector $x \in \mathbb{R}^n$ and a direction $y \in \mathbb{R}^n$, the Householder reflector H reflects x along y, i.e.,

$$Hx = ||x||y \qquad \text{with} \qquad ||y|| = 1.$$

Thanks to its properties, H writes as

$$H = \mathbb{I}_n - \frac{2}{||u||^2} u \otimes u \qquad \text{with} \qquad u = (x - ||x||y).$$

The practical implementation of the Householder transformation kernel uses the components of the input vectors.

Remark

The Householder algorithm does **not** readily apply to tensor in TT-formats, because of the compressed nature of this format.

Loss Of Orthogonality

Let $Q_k = [q_1, \dots, q_k]$ be the orthogonal basis produced by an orthogonalization kernel, then the **Loss Of Orthogonality** is

$$||I_k - Q_k^\top Q_k||.$$

It measures the quality in terms of orthogonality of the computed basis. It is linked with the linearly dependency of the input vectors $A_k = [a_1, \dots, a_k]$, estimated through $\kappa(A_k)$.

Matrix		
Source	*Algorithm*	$\left\|\mathbb{I}_k - Q_k^\top Q_k\right\|$
[Stathopoulos and Wu 2002]	**Gram**	$\mathcal{O}(u\kappa^2(A_k))$
[Giraud, Langou, and Rozložník 2005]	**CGS**	$\mathcal{O}(u\kappa^2(A_k))$
[Björck 1967]	**MGS**	$\mathcal{O}(u\kappa(A_k))$
[Giraud, Langou, and Rozložník 2005]	**CGS2**	$\mathcal{O}(u)$
[Giraud, Langou, and Rozložník 2005]	**MGS2**	$\mathcal{O}(u)$
[Wilkinson 1965]	**Householder**	$\mathcal{O}(u)$

Numerical experiments in TT-format

$$\mathbf{x}_{k+1} = \texttt{TT-rounding}(\boldsymbol{\Delta}_d \mathbf{a}_k, \texttt{max\_rank} = 1) \quad \text{with} \quad \mathbf{a}_{k+1} = \frac{1}{\|\mathbf{x}_{k+1}\|}\mathbf{x}_{k+1}$$

Numerical experiments in TT-format

$$\mathbf{x}_{k+1} = \texttt{TT-rounding}(\boldsymbol{\Delta}_d \mathbf{a}_k, \texttt{max\_rank} = 1) \quad \text{with} \quad \mathbf{a}_{k+1} = \frac{1}{\|\mathbf{x}_{k+1}\|}\mathbf{x}_{k+1}$$

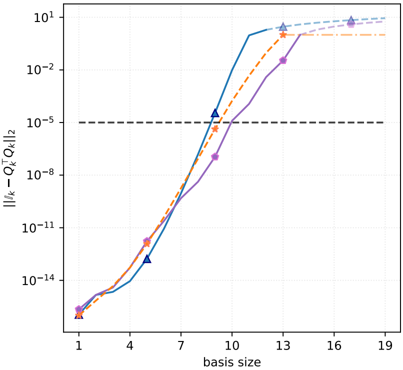


Figure: Loss of orthogonality for $m = 20$ TT-vectors of order $d = 3$ and mode size $n = 15$, rounding precision $\delta = 10^{-5}$

- **Gram approach**
- **CGS**
- $\kappa^2(A_k)$

$$\mathcal{O}(u\kappa^2(A_k))$$

Numerical experiments in TT-format

$$\mathbf{x}_{k+1} = \texttt{TT-rounding}(\boldsymbol{\Delta}_d \mathbf{a}_k, \texttt{max\_rank} = 1) \quad \text{with} \quad \mathbf{a}_{k+1} = \frac{1}{\|\mathbf{x}_{k+1}\|}\mathbf{x}_{k+1}$$

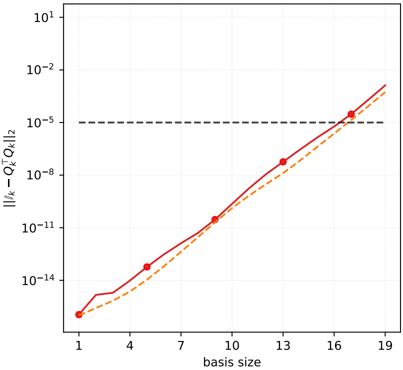


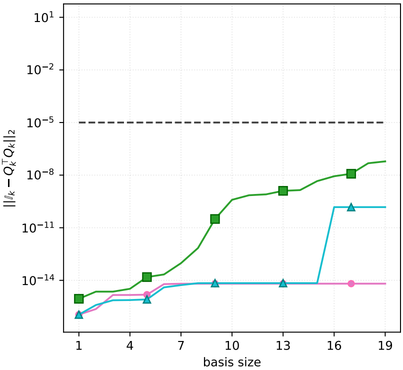
Figure: Loss of orthogonality for $m = 20$ TT-vectors of order $d = 3$ and mode size $n = 15$, rounding precision $\delta = 10^{-5}$

- **MGS**
- $\kappa(A_k)$

$\mathcal{O}(u\kappa(A_k))$

Numerical experiments in TT-format

$$\mathbf{x}_{k+1} = \texttt{TT-rounding}(\boldsymbol{\Delta}_d \mathbf{a}_k, \texttt{max\_rank} = 1) \quad \text{with} \quad \mathbf{a}_{k+1} = \frac{1}{\|\mathbf{x}_{k+1}\|}\mathbf{x}_{k+1}$$



- **CGS2**
- **MGS2**
- **Householder transformation**

$\mathcal{O}(u)$

Figure: Loss of orthogonality for $m = 20$ TT-vectors of order $d = 3$ and mode size $n = 15$, rounding precision $\delta = 10^{-5}$

Numerical experiments in TT-format

$$\mathbf{x}_{k+1} = \texttt{TT-rounding}(\boldsymbol{\Delta}_d \mathbf{a}_k, \texttt{max\_rank} = 1) \quad \text{with} \quad \mathbf{a}_{k+1} = \frac{1}{\|\mathbf{x}_{k+1}\|}\mathbf{x}_{k+1}$$

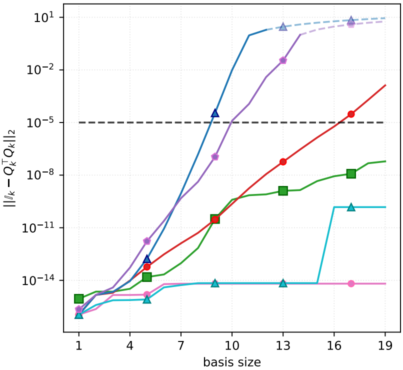


Figure: Loss of orthogonality for $m = 20$ TT-vectors of order $d = 3$ and mode size $n = 15$, rounding precision $\delta = 10^{-5}$

- **Gram approach**
- **CGS**
- **MGS**
- **CGS2**
- **MGS2**
- **Householder transformation**

Loss of orthogonality: matrix vs tensor

	Matrix, theoretical	TT-format, conjecture
Algorithm	$\left\|\mathbb{I}_k - Q_k^\top Q_k\right\|$	$\left\|\mathbb{I}_k - \mathcal{Q}_k^\top \mathcal{Q}_k\right\|$
Gram	$\mathcal{O}(u\kappa^2(A_k))$	$\mathcal{O}(\delta\kappa^2(\mathcal{A}_k))$
CGS	$\mathcal{O}(u\kappa^2(A_k))$	$\mathcal{O}(\delta\kappa^2(\mathcal{A}_k))$
MGS	$\mathcal{O}(u\kappa(A_k))$	$\mathcal{O}(\delta\kappa(\mathcal{A}_k))$
CGS2	$\mathcal{O}(u)$	$\mathcal{O}(\delta)$
MGS2	$\mathcal{O}(u)$	$\mathcal{O}(\delta)$
Householder	$\mathcal{O}(u)$	$\mathcal{O}(\delta)$

Numerical linear algebra summary II

- Backward stability of GMRES holds numerically when data are stored with
  - componentwise perturbations
  - normwise perturbations

  both have practical implementations

- GMRES in TT-format with TT-rounding seems to be backward stable

- Parameter dependent TT-problems of order d can be solved at once through an order $(d+1)$ problem, guaranteeing backward stable bounds linking the $(d+1)$ and d solutions

- Loss of orthogonality bounds for six widely used orthogonalization kernels appears to hold true when generalized to TT-format with normwise perturbation

3

Data analysis

- Background on Correspondence Analysis
- MultiWay Correspondence Analysis
- A study case: the Malabar dataset

Joint work with A. Franc

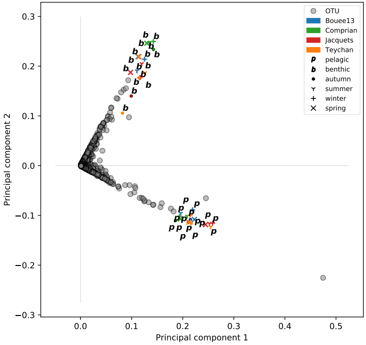
Malabar dataset [Auby et al. 2022]

$d = 4$ mode dataset

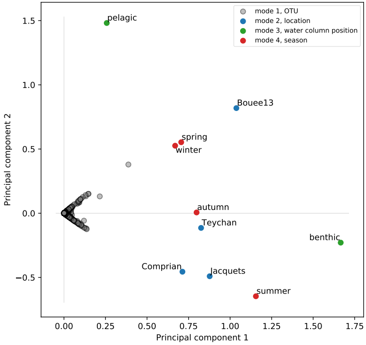
- 1st mode of Operational Taxonomic Units (OTUs) with size $n_1 = 3539$
- 2nd mode of locations with size $n_2 = 4$, namely Bouee13, Comprian, Jacquets, Teychan
- 3rd mode of water column position with size $n_3 = 2$, that are pelagic and benthic
- 4th mode of seasons with size $n_4 = 4$, that are spring, summer, autumn and winter

Figure: Aerial tour of the Arcachon basin, France

The addressed question



Correspondence Analysis on mode 1



MultiWay Correspondence Analysis

Is it mathematically meaningful to display point clouds together?

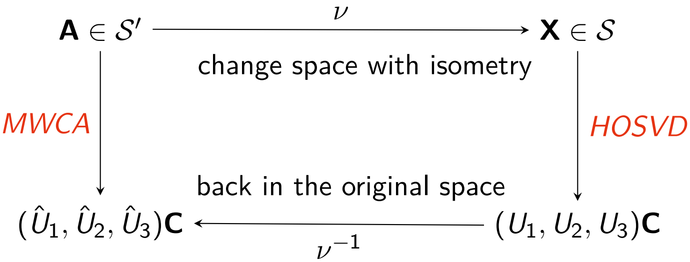
Correspondence Analysis (CA) is a Principal Component Analysis (PCA) meant to investigate contingency tables through a **specific norm**

$$\begin{array}{ccc} A \in \mathcal{S}' & \xrightarrow{\quad \nu \quad} & X \in \mathcal{S} \\ & \text{change space with isometry} & \\ \Big\downarrow \textit{CA} & & \Big\downarrow \textit{SVD} \\ & \text{back in the original space} & \\ \hat{U}\Sigma\hat{V}^\top & \xleftarrow[\quad \nu^{-1} \quad]{} & U\Sigma V^\top \end{array}$$

Geometrical meaning

The h-th Principal Coordinate of the i-th category of the row variable is the barycentre of the h-th Principal Coordinate of all the column variable categories

> MWCA is a Principal Components Analysis of a multiway table with a **specific norm**



Tucker model

Let $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ be an order 3 tensor, then

Tucker format [Tucker 1963]

$$\mathbf{A} = \sum_{\alpha,\beta,\gamma=1}^{r_1,\ldots,r_d} \mathbf{C}(\alpha,\beta,\gamma) u_\alpha^{(1)} \otimes u_\beta^{(2)} \otimes u_\gamma^{(3)}$$

where $\mathbf{C}$ is the **core tensor**, $U_k = [u_1^{(k)}, \ldots, u_{r_k}^{(k)}]$ is an orthogonal $(n_k \times r_k)$ matrix with $r_k = \text{rank}(A^{(k)})$. More compactly

$$\mathbf{A} = (U_1, U_2, U_3)\mathbf{C}$$

Tucker model

Let $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ be an order 3 tensor, then

Tucker format [Tucker 1963]

$$\mathbf{A} = \sum_{\alpha,\beta,\gamma=1}^{r_1,\dots,r_d} \mathbf{C}(\alpha,\beta,\gamma) u_\alpha^{(1)} \otimes u_\beta^{(2)} \otimes u_\gamma^{(3)}$$

where $\mathbf{C}$ is the **core tensor**, $U_k = [u_1^{(k)}, \dots, u_{r_k}^{(k)}]$ is an orthogonal $(n_k \times r_k)$ matrix with $r_k = \mathrm{rank}(A^{(k)})$. More compactly

$$\mathbf{A} = (U_1, U_2, U_3)\mathbf{C}$$

Tucker decomposition can be computed through the HOSVD [De Lathauwer, De Moor, and Vandewalle 2000], that performs a SVD for each matricization

$$\mathbf{A}^{(k)} = U_k \Sigma_k V_k^\top$$

and the core tensor is obtained projecting the tensor $\mathbf{A}$ in the new basis, i.e., $\mathbf{C} = (U_1^\top, U_2^\top, U_3^\top)\mathbf{A}$

MWCA inner product

Let $\mathbf{A}$ be a relative frequencies multiway contingency table, the 1st mode marginal is $a_1 \in \mathbb{R}^{n_1}$ such that

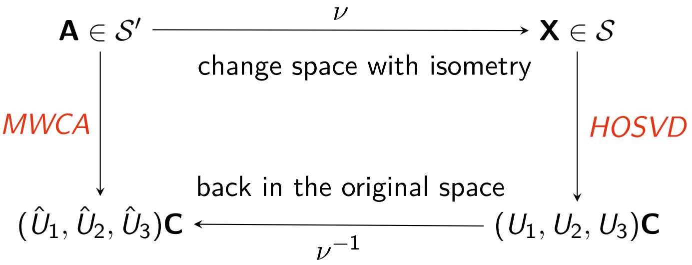
$$a_1(i) = \sum_{j,k=1}^{n_2,n_3} \mathbf{A}(i,j,k)$$

- $\mathcal{S}'$ denotes the Euclidean space $\mathbb{R}^{n_1 \times n_2 \times n_3}$ with the inner product induced by $(D_1^{-1}, D_2^{-1}, D_3^{-1})$, with $D_k = \mathrm{diag}(\sqrt{a_k})$
- $\mathcal{S}$ denotes the Euclidean space $\mathbb{R}^{n_1 \times n_2 \times n_3}$ with the standard inner product.

The two Euclidean spaces are isometric through ν defined as

$$\nu : \mathcal{S}' \to \mathcal{S} \qquad \text{such that} \qquad \nu(\mathbf{A}) = (D_1^{-1}, D_2^{-1}, D_3^{-1})\mathbf{A}$$

MWCA scheme



Point cloud coordinates in $\mathcal{S}'$

$$W_k = \hat{U}_k \Sigma_k = D_k U_k \Sigma_k$$

Point cloud coordinates in $\mathcal{S}$

$$Y_k = U_k \Sigma_k$$

where Σ_k are the singular values of $\mathbf{X}^{(k)}$

MWCA barycentric relation

If $Z_1 = D_1^{-2} W_1$, then

Barycentric relation

$$Z_1 = D_1^{-2} \mathbf{A}^{(1)} (Z_3 \otimes_{\mathrm{K}} Z_2)(\Sigma_1^{-1} \mathbf{B}_1^{(1)})^\top$$

where

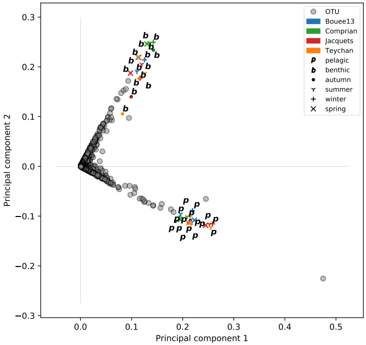
$$\mathbf{B}_1 = (\mathbb{I}_k, \Sigma_2^{-1}, \Sigma_3^{-1}) \mathbf{C}$$

with $\mathbf{C}$ the core tensor of $\mathbf{X} = \nu(\mathbf{A})$

Geometrical meaning

The h_i-th Principal Coordinate of the i-th category of the 1st variable is the barycentre of a linear combination of the h-th Principal Coordinate of all the other two variable categories and coefficients expressed by $\mathbf{B}_1$ entries

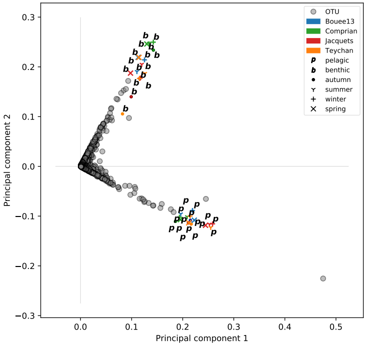
MWCA on Malabar dataset



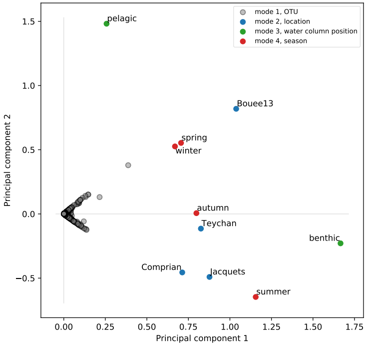
CA on mode 1

- Orthogonality among water column positions
- Distribution of OTU along their directions

MWCA on Malabar dataset



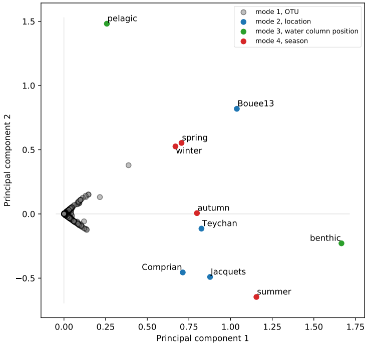
CA on mode 1



MWCA

MWCA on Malabar dataset

- Water column positions still orthogonal, less determinant
- OTU distribution affected by seasons and locations too
- Locations and seasons are clustered
- Correlation between seasons and locations



MWCA

Data analysis summary

- MultiWay Correspondence Analysis has a geometric nature as Correspondence Analysis
- The barycentric relation characterizing CA holds also for MWCA
- MWCA highlights inter-variable interactions
- CA appears more suitable for studying combinations of categories from different variables

4 Conclusion & perspectives

Main contribution

- Variable accuracy approach for GMRES with componentwise and normwise δ-`storage` perturbations

- GMRES in TT-format and parameter dependent bounds

- Numerical study of the loss of orthogonality for six widely used orthogonalization kernels in TT-format

- MultiWay Correspondence Analysis with the Malabar dataset study case

Main contribution

- Variable accuracy approach for GMRES with componentwise and normwise δ-`storage` perturbations
  [Inria RR-9483], joint work with *E. Agullo, O. Coulaud, L. Giraud, G. Marait and N. Schenkels*
- GMRES in TT-format and parameter dependent bounds
  [Inria RR-9484], joint work with *O. Coulaud and L. Giraud*
- Numerical study of the loss of orthogonality for six widely used orthogonalization kernels in TT-format
  [Inria RR-9491], joint work with *O. Coulaud and L. Giraud*
- MultiWay Correspondence Analysis with the Malabar dataset study case
  [Inria RR-9429], joint work with *O. Coulaud and A. Franc*

Perspectives

- Numerical linear algebra
  - Theoretical proof for the normwise δ-`storage` (for future work)
  - Implementation of TT-GMRES with the Householder kernel
  - Study of preconditioner for the tensor case
- Data analysis
  - Statistical interpretation of MultiWay Correspondence Analysis results
  - Combination of isometry and Canonical Polyadic decomposition
  - New visualization techniques

Thanks for the attention.

Questions?

Alternative backward stable stopping criterion

Definition of normwise backward error

$$\eta_{A,b}(x_k) = \min_{\Delta A, \Delta b} \{\tau > 0 : ||\Delta A|| \leq \tau ||A||,\ ||\Delta b|| \leq \tau ||b|| \text{ and } (A + \Delta A)x_k = b + \Delta\}$$
$$= \frac{||Ax_k - b||}{||A||||x_k|| + ||b||}$$

If M is a preconditioner, the previous stopping criterion gets

$$\eta_{AM,b}(t_k) = \frac{||AMt_k - b||}{||AM||||t_k|| + ||b||} \quad \text{and} \quad x_k = Mt_k.$$

Another possible one based just on the right-hand side is

$$\eta_b(x_k) = \min_{\Delta A, \Delta b} \{\tau > 0 : ||\Delta b|| \leq \tau ||b|| \text{ and } Ax_k = b + \Delta b\}$$
$$= \frac{||Ax_k - b||}{||b||}.$$

TT-rounding [Oseledets 2011]

$\underline{\hat{\mathbf{a}}}_1, \ldots, \underline{\hat{\mathbf{a}}}_d = \text{TT-rounding}\big((\underline{\mathbf{a}}_1, \ldots, \underline{\mathbf{a}}_d),\, \delta\big)$

Input: $\underline{\mathbf{a}}_i \in \mathbb{R}^{r_{i-1} \times n_i \times r_i}$ for every $i \in \{1, \ldots, d\}$, $\delta \in \mathbb{R}_+$

$\delta' = \frac{\delta}{\sqrt{d-1}} \|\mathbf{a}\|$

for $i = d, \ldots, 2$ **do**

  $A_i = \texttt{matricize}(\underline{\mathbf{a}}_i, \texttt{mode} = 1)$

  $Q_i, R = \texttt{QR}(A_i^\top)$

  $\underline{\mathbf{a}}_i = \texttt{reshape}(Q_i^\top, [r_{i-1},\, n_i,\, r_i])$

  $\underline{\mathbf{a}}_{i-1} = \underline{\mathbf{a}}_{i-1} \times_3 R^\top$

end

$s_0 = 1$

for $i = 1, \ldots, d-1$ **do**

  $A_i = \texttt{matricize}(\underline{\mathbf{a}}_i, \texttt{mode} = 3)$

  $\hat{U}_i, \hat{\Sigma}_i, \hat{V}_i = \texttt{T-SVD}(A_i^\top, \delta')$

  $\underline{\hat{\mathbf{a}}}_i = \texttt{reshape}(\hat{U}_i, [s_{i-1},\, n_i,\, s_i])$ with s_i equal to the number of columns of $\hat{U}_i$

  $\underline{\mathbf{a}}_{i+1} = \mathbf{a}_{i+1} \times_1 (\hat{\Sigma}_i \hat{V}_i^\top)$

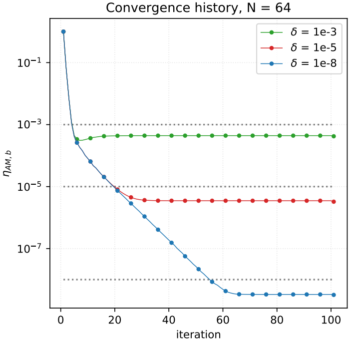
end

$\underline{\hat{\mathbf{a}}}_d = \mathbf{a}_d$

Output: $\underline{\hat{\mathbf{a}}}_1, \ldots, \underline{\hat{\mathbf{a}}}_d$

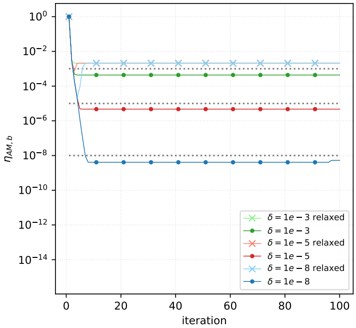
TT-GMRES vs relaxed TT-GMRES from [Dolgov 2013]

TT-GMRES



- Rounding precision δ constant
- Backward stable stopping criterion

relaxed TT-GMRES

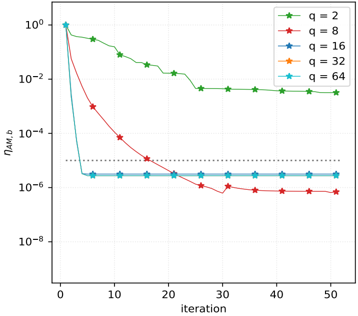


- Rounding precision δ increases with the iterations
- Stopping criterion based on the least square residual

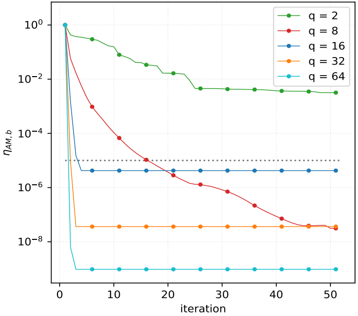
Preconditioner in TT-format

$$\mathbf{M} = \sum_{k=-M}^{M} c_k \exp(-t_k \Delta_1) \otimes \cdots \otimes \exp(-t_k \Delta_1)$$

from the approximation of inverse of $\boldsymbol{\Delta}_d$ [Hackbusch and Khoromskij 2006a,b]



Convergence history for $\tau = 10^{-2}$ and rounding $\delta = 10^{-5}$

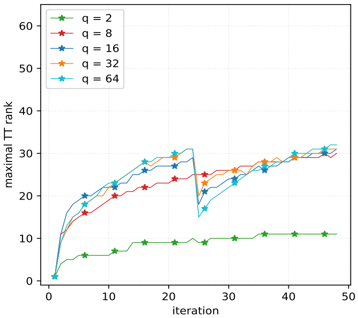


Convergence history for $\tau = 10^{-8}$ and rounding $\delta = 10^{-5}$

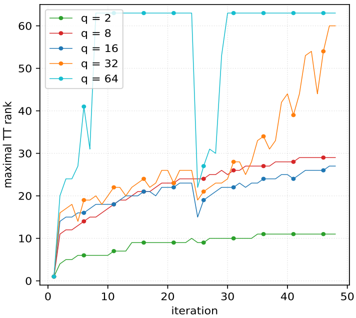
Preconditioner in TT-format

$$\mathbf{M} = \sum_{k=-M}^{M} c_k \exp(-t_k \Delta_1) \otimes \cdots \otimes \exp(-t_k \Delta_1)$$

from the approximation of inverse of $\boldsymbol{\Delta}_d$ [Hackbusch and Khoromskij 2006a,b]



Maximal TT-rank of the last Krylov vector for $\tau = 10^{-2}$



Maximal TT-rank of the last Krylov vector for $\tau = 10^{-8}$

Gram approach

Let $A = [a_1, \ldots, a_m]$, then we look for $A = QR$ with $Q^\top Q = \mathbb{I}_m$ compute the Gram matrix

$$A^\top A = (R^\top Q^\top) QR = R^\top R$$

this is (almost) the **Cholesky** factorization of $A^\top A$ that can be written as

$$A^\top A = R^\top R = LL^\top$$

with the Cholesky factor $L = R^\top$ and then Q gets

$$Q = AR^{-1} = A(L^\top)^{-1}$$

Gram approach

$\mathcal{Q}, R = \texttt{Gram}(\mathcal{A}, \delta)$

Input: $\mathcal{A} = \{a_1, \dots, a_m\}$, $\delta \in \mathbb{R}_+$

```
A = [a_1, ..., a_d]
G = A^⊤ A
L = cholesky(G)
Q = A(L^⊤)^{-1} = [q_1, ..., q_d]
for i = 1, ..., m do
    q_i = δ-storage(q_i)
end
```

Output: $\mathcal{Q} = \{q_1, \dots, q_m\}$, R

In TT-format the following modifications occur

- $G(i,j)$ is the scalar product of $\mathbf{a}_i$ and $\mathbf{a}_j$
- The inverse of $L^\top$ is explicitly computed
- $\mathbf{q}_i$ is constructed as a linear combination of $\mathcal{A}$ elements
- `TT-rounding` is used to compress at precision δ

Computational costs: matrix vs tensor

	cost in fp operations	cost in TT-rounding
Gram	$\mathcal{O}(2nm^2)$	m
CGS	$\mathcal{O}(2nm^2)$	m
MGS	$\mathcal{O}(2nm^2)$	m
CGS2	$\mathcal{O}(4nm^2)$	$2m$
MGS2	$\mathcal{O}(4nm^2)$	$2m$
Householder	$\mathcal{O}(2nm^2 - 2m^3/3)$	$4m$

TT-Subspace iteration

$\#$ orthogonalization as x-axis $\#$ TT-rounding as x-axis

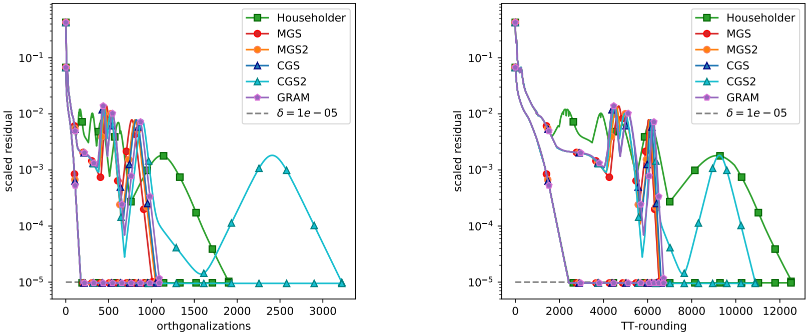
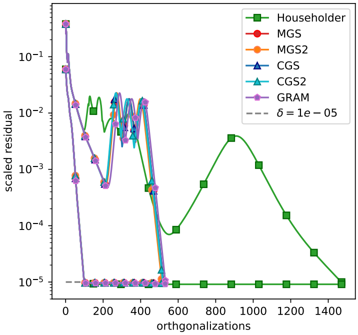


Figure: Residual envelope for $m = 7$ TT-eigenpairs of order $d = 3$ with mode size $\mathbf{n} = [\mathbf{19}, \mathbf{24}, \mathbf{31}]$ and rounding precision $\delta = 10^{-5}$.

TT-Subspace iteration

orthogonalization as x-axis



TT-rounding as x-axis

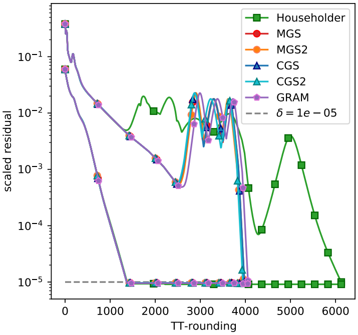


Figure: Residual envelope for $m = 7$ TT-eigenpairs of order $d = 3$ with mode size $\mathbf{n} = [\mathbf{24}, \mathbf{24}, \mathbf{24}]$ and rounding precision $\delta = 10^{-5}$.

CA inner product

Let A be a contingency table with relative frequencies, the row and a_R and column marginals a_C are

$$a_R(i) = \sum_{j=1}^{n} A(i,j) \qquad \text{and} \qquad a_C(j) = \sum_{i=1}^{m} A(i,j).$$

$\mathcal{S}'$ denotes the Euclidean space $\mathbb{R}^{m \times n}$ with the inner product induced by D_R^{-1} and D_C^{-1} such that

$$D_R = \text{diag}(\sqrt{a_R}) \qquad \text{and} \qquad D_C = \text{diag}(\sqrt{a_C}),$$

$\mathcal{S}$ denotes the Euclidean space $\mathbb{R}^{m \times n}$ with the standard inner product

The two Euclidean spaces are isometric through ν defined as

$$\nu : \mathcal{S}' \to \mathcal{S} \qquad \text{such that} \qquad \nu(A) = D_R^{-1} A D_C^{-1}$$

CA construction

$$\begin{array}{ccc} A \in \mathcal{S}' & \xrightarrow{\quad\nu\quad} & X \in \mathcal{S} \\ \color{red}{CA}\downarrow & \text{change space with isometry} & \downarrow\color{red}{SVD} \\ & \text{back in the original space} & \\ \hat{U}\Sigma\hat{V}^\top & \xleftarrow[\quad\nu^{-1}\quad]{} & U\Sigma V^\top \end{array}$$

Point cloud coordinates in $\mathcal{S}'$

$$W_R = \hat{U}\Sigma = D_R U\Sigma$$
$$W_C = \hat{V}\Sigma = D_C V\Sigma$$

Point cloud coordinates in $\mathcal{S}$

$$Y_R = U\Sigma$$
$$Y_C = V\Sigma$$

Barycentric relation [Lebart, 1982]

$$Z_R = D_R^{-2} A Z_C \Sigma^{-1} \quad \text{and} \quad Z_C = D_C^{-2} A^\top Z_R \Sigma^{-1}$$

with $Z_i = D_i^{-2} W_i$

Barycentric relation in CA

$$Z_R(i,h) = \left(D_R^{-2} A Z_C \Sigma^{-1}\right)(i,h) = \frac{1}{\sigma_h} \sum_{j=1}^{n} \frac{A(i,j)}{a_R(i)} Z_C(j,h)$$

$$= \frac{1}{\sigma_h} \sum_{j=1}^{n} \rho_i(j) Z_C(j,h)$$

with σ_h the h-th singular value and $\rho_i \in \mathbb{R}^m$ such that

$$\sum_{j=1}^{n} \rho_i(j) = \sum_{j=1}^{n} \frac{A(i,j)}{a_R(i)} = \frac{a_R(i)}{a_R(i)} = 1$$

Geometrical meaning

The h-th Principal Coordinate of the i-th category of the row variable is the barycentre of the h-th Principal Coordinate of all the column variable categories