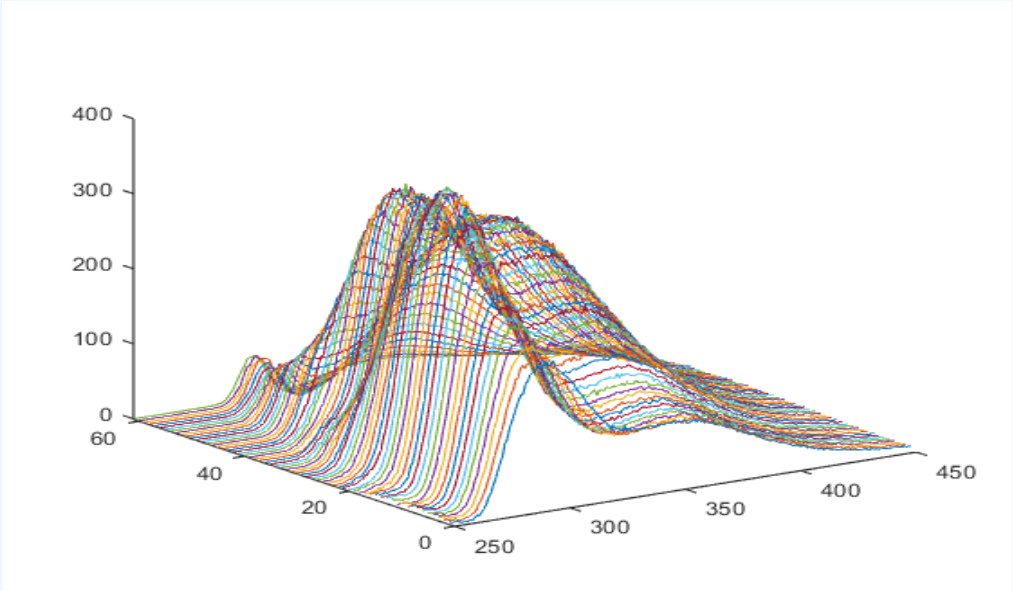
Existence and Algorithms for Best Low-Rank Tensor Approximation

Eric Evert

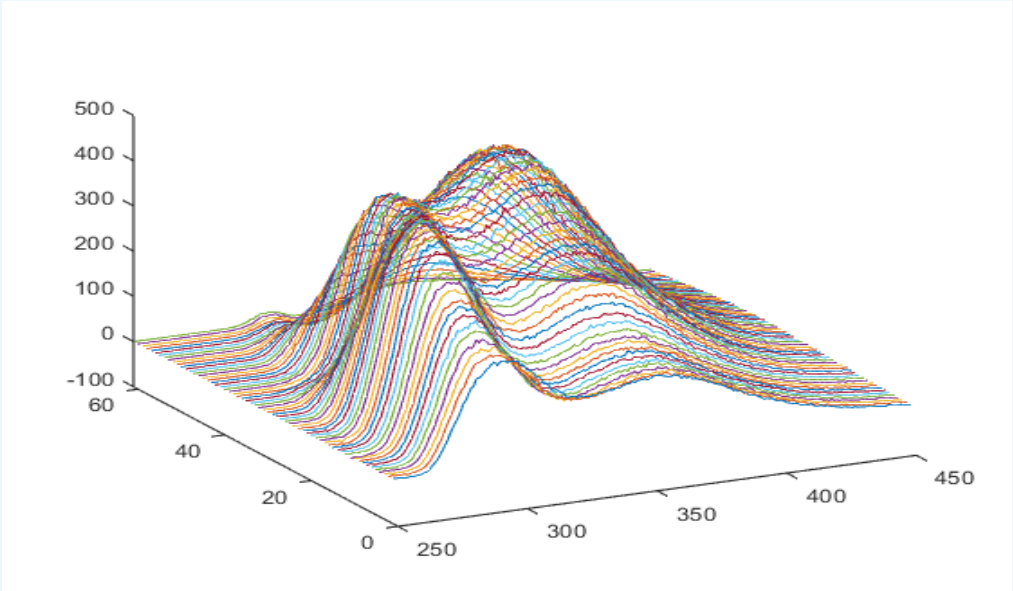
Joint work with Lieven De Lathauwer

UT Austin Data and Algebra, 7 March 2025

Emission Excitation Spectra (due to R. Bro)



Emission Excitation Spectra (due to R. Bro)



Tensor decompositions recover amino acid mixtures

Represent our amino acid data as a multiindexed array $\mathcal{T}$ of size $201 \times 61 \times 2$.

$\mathcal{T}$ approximately has "tensor rank" equal to 3.

$\mathcal{T}$ decomposes as

$$a_1 \otimes b_1 \otimes \begin{pmatrix} 0.409 \\ 0.312 \end{pmatrix} + a_2 \otimes b_2 \otimes \begin{pmatrix} 0.284 \\ 0.307 \end{pmatrix} + a_3 \otimes b_3 \otimes \begin{pmatrix} 0.409 \\ 0.363 \end{pmatrix}$$

Tensor decompositions recover amino acid mixtures

Represent our amino acid data as a multiindexed array $\mathcal{T}$ of size $201 \times 61 \times 2$.

$\mathcal{T}$ approximately has "tensor rank" equal to 3.

$\mathcal{T}$ decomposes as

$$a_1 \otimes b_1 \otimes \begin{pmatrix} 0.409 \\ 0.312 \end{pmatrix} + a_2 \otimes b_2 \otimes \begin{pmatrix} 0.284 \\ 0.307 \end{pmatrix} + a_3 \otimes b_3 \otimes \begin{pmatrix} 0.409 \\ 0.363 \end{pmatrix}$$

The true concentrations are: Mixture 1: 0.424, 0.293, 0.283
Mixture 2: 0.333, 0.334, 0.333

Multidimensional arrays.

A **tensor** $\mathcal{T}$ is a multiindexed array of size $R \times R \times K$.

$$\mathcal{T} = \quad \in \mathbb{R}^{R \times R \times K}$$

E.g., $\mathcal{T} \in \mathbb{R}^{3\times3\times2}$ defined by $\mathcal{T}(i,j,k) = i + j + k$ is the tensor with **frontal slices**

$$\mathbf{T}_1 := \mathcal{T}(:,:,1) = \begin{pmatrix} 3 & 4 & 5 \\ 4 & 5 & 6 \\ 5 & 6 & 7 \end{pmatrix} \qquad \mathbf{T}_2 := \mathcal{T}(:,:,2) = \begin{pmatrix} is4 & 5 & 6 \\ 5 & 6 & 7 \\ 6 & 7 & 8 \end{pmatrix}$$

The tensor product

Let $\otimes$ denote the **tensor outer product**. That is, for vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^R$ and $\mathbf{c} \in \mathbb{R}^k$ where $\mathbf{a} = (\mathbf{a}(1), \mathbf{a}(2), \ldots, \mathbf{a}(R))$, the tensor

$$\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} \in \mathbb{R}^{R \times R \times K}$$

has i, j, k entry equal to

$$\mathbf{a}(i)\mathbf{b}(j)\mathbf{c}(k)$$

The tensor product

Let $\otimes$ denote the **tensor outer product**. That is, for vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^R$ and $\mathbf{c} \in \mathbb{R}^k$ where $\mathbf{a} = (\mathbf{a}(1), \mathbf{a}(2), \ldots, \mathbf{a}(R))$, the tensor

$$\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} \in \mathbb{R}^{R \times R \times K}$$

has i, j, k entry equal to

$$\mathbf{a}(i)\mathbf{b}(j)\mathbf{c}(k)$$

E.g. the tensor product between vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^R$ is equal to the matrix

$$\mathbf{a} \otimes \mathbf{b} = \mathbf{a}\mathbf{b}^\mathsf{T} \in \mathbb{R}^{R \times R}.$$

which has i, j entry equal to $\mathbf{a}(i)\mathbf{b}(j)$.

A tensor of the form $\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}$ is called a **rank one tensor**.

Decompose tensor into canonical components.

Every tensor can be expressed as a sum of rank one tensors. E.g,

$$\mathcal{T} = \sum \mathcal{T}(i,j,k)\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k.$$

However, this is not a minimal decomposition.

Canonical **P**olyadic **D**ecomp. (CPD) expresses $\mathcal{T}$ as minimal sum of rank 1 terms.

$$\mathcal{T} = \sum_{\ell=1}^{L} \mathbf{a}_\ell \otimes \mathbf{b}_\ell \otimes \mathbf{c}_\ell =$$

If L is as small as possible, then L is called the **rank** of $\mathcal{T}$.

Many differences between matrix and tensor rank

A tensor in $\mathbb{R}^{R \times R \times R}$ is expected to have rank $\approx R^2/3$.

Tensor rank depends on whether decomposition is over reals or complexes. E.g. there exist tensors with complex rank 2 but real rank 3.

For low rank tensors (e.g. tensors with rank $\leq R$) which satisfy light assumptions, CPD is unique.

The set of tensors of rank $\leq L$ is not closed unless $L = 1$ or L is sufficiently large. E.g. there exists a sequence of rank 2 tensors which converges to a rank 3 tensor.

Low rank CPD computation is a big industry...

Uniqueness of low rank canonical polyadic decompositions makes CPD a big tool in applications.

Often tensor $\mathcal{T}$ is some low rank signal tensor. Decomposing this signal with CPD can reveal component information. One example problem is blind source separation.

CPD has applications in machine learning, artificial intelligence, signal processing, data science, chemometrics, biomath, etc.

but there can be some challenges.

Key issue: In practice only have access to a measurement $\mathcal{T} + \mathcal{N}$ where $\mathcal{N}$ is noise. However, $\mathcal{T} + \mathcal{N}$ is not low rank.

Must compute a best low rank approximation to $\mathcal{T} + \mathcal{N}$, but a best low rank approximation can fail to exist due to nonclosedness of the set of low rank tensors.

If a best low rank approximation does not exist, then near optimal low rank approximations exhibit undesirable properties.

Even if a best low rank approximation exists, if is NP-hard to compute. A popular strategy is to use optimization initialized by an algebraic approximation.

Example of tensor that does not have a best low rank approximation

Consider the $2 \times 2 \times 2$ tensor $\mathcal{S}$ defined by

$$\mathcal{S}(:,:,1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \mathcal{S}(:,:,2) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$\mathcal{S}$ has rank 3 but is a limit of rank 2 tensors. In particular

$$\mathcal{S} = \lim_{n\to\infty} -n(\mathbf{e}_1)^{\otimes 3} + n\left(\mathbf{e}_1 + \frac{\mathbf{e}_2}{n}\right)^{\otimes 3} = \lim_{n\to\infty} \left(\begin{pmatrix} 0 & 1 \\ 1 & 1/n \end{pmatrix}, \begin{pmatrix} 1 & 1/n \\ 1/n & 1/n^2 \end{pmatrix} \right)$$

This is bad news for interpreting component information, as the two components $-n(\mathbf{e}_1)^{\otimes 3}$ and $n\left(\mathbf{e}_1 + \frac{\mathbf{e}_2}{n}\right)^{\otimes 3}$ each approach having infinite magnitude as n grows.

Tensors without best low rank approximations must exhibit diverging components

Suppose $\mathcal{T} \in \mathbb{R}^{R \times R \times K}$ has rank L but is a limit of rank $\ell < L$ tensors $\mathcal{T}^{(n)}$. Then the $\mathcal{T}^{(n)}$ must have (at least two) rank one terms whose norm goes to infinity.

Suppose toward a contradiction that

$$\mathcal{T}^{(n)} = \sum_{j=1}^{\ell} \mathcal{T}_j^{(n)}$$

where each $\mathcal{T}_j^{(n)}$ is a rank one tensor and where $\sup_n \{\|\mathcal{T}_j^{(n)}\|_F\} < \infty$ for each ℓ.

Tensors without best low rank approximations must exhibit diverging components

Suppose $\mathcal{T} \in \mathbb{R}^{R \times R \times K}$ has rank L but is a limit of rank $\ell < L$ tensors $\mathcal{T}^{(n)}$. Then the $\mathcal{T}^{(n)}$ must have (at least two) rank one terms whose norm goes to infinity.

Suppose toward a contradiction that

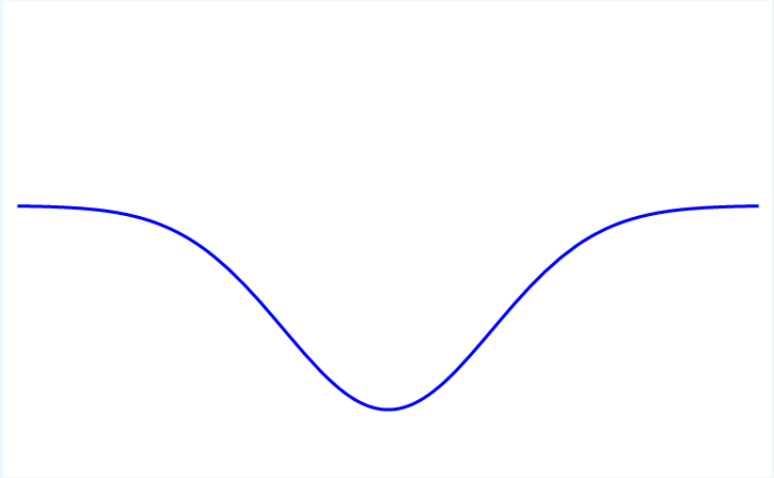
$$\mathcal{T}^{(n)} = \sum_{j=1}^{\ell} \mathcal{T}_j^{(n)}$$

where each $\mathcal{T}_j^{(n)}$ is a rank one tensor and where $\sup_n \{\|\mathcal{T}_j^{(n)}\|_F\} < \infty$ for each ℓ.

Passing to a subsequence, each $\mathcal{T}_j^{(n)}$ converges to some $\mathcal{T}_j$, which also has rank one, hence

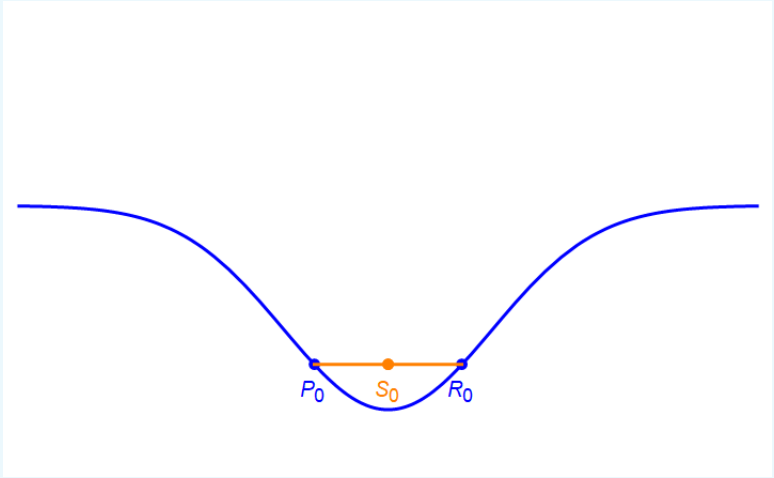
$$\mathcal{T} = \sum_{j=1}^{\ell} \mathcal{T}_j.$$

Nonclosedness of tensor of rank $\leq R$ is due to geometry of rank one tensors



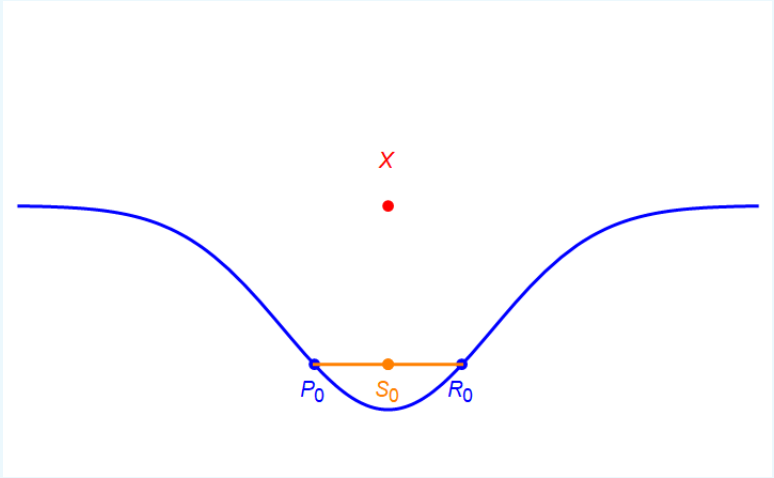
Nonclosedness of tensor of rank $\leq R$ is due to geometry of rank one tensors

A tensor has rank 2 means it is a linear combination of rank 1 tensors. I.e. it is on a line between two rank one tensors.



Nonclosedness of tensor of rank $\leq R$ is due to geometry of rank one tensors

The tensor X below is not rank 2 due to the horizontal asymptote for the set of rank 1 tensors. However, X is a limit of rank 2 tensors.



Changing the point of view

Bad instances for CPD have lead to the mathematical perspective that low rank CPD approximation is a challenging, ill-posed problem. In practice CPD is often very successful. Can we bridge the gap in perspective?

Changing the point of view

Bad instances for CPD have lead to the mathematical perspective that low rank CPD approximation is a challenging, ill-posed problem. In practice CPD is often very successful. Can we bridge the gap in perspective?

"Theorem" (E-De Lathauwer) For many tensors occurring in applications, best low-rank tensor approximation is well-posed in a mathematically quantifiable neighborhood around the tensor.

Symmetric slices and the spectral norm

Say a tensor $\mathcal{T} \in \mathbb{R}^{R \times R \times K}$ has symmetric frontal slices if the frontal slice $\mathbf{T}_r$ is symmetric for each $r = 1, \ldots, K$.

The spectral norm $\|\mathcal{T}\|_{sp}$ of $\mathcal{T}$ is the Frobenius norm of a best rank one approximation to $\mathcal{T}$.

Spectral norm bound guaranteeing existence of best low rank approximation

Theorem [E-De Lathauwer]

Let $\mathcal{T}, \mathcal{N} \in \mathbb{R}^{R \times R \times K}$ and assume $\mathcal{T}$ has rank R and has SFS. If

$$\|\mathcal{N}\|_{sp} < \max_{\|\mathbf{w}\|=1} \min_{\|\mathbf{v}\|=1} \mathbf{v}^T \left(\sum_{r=1}^{K} \mathbf{w}(r)(\mathbf{T}_r + \mathbf{N}_r) \right) \mathbf{v}$$

then $\mathcal{T} + \mathcal{N}$ has a best rank R approximation among SFS tensors.

Here $\mathbf{T}_r + \mathbf{N}_r$ denotes the rth frontal slice of $\mathcal{T} + \mathcal{N}$.

Spectral norm bound guaranteeing existence of best low rank approximation

Theorem [E-De Lathauwer]

Let $\mathcal{T}, \mathcal{N} \in \mathbb{R}^{R \times R \times K}$ and assume $\mathcal{T}$ has rank R and has SFS. If

$$\|\mathcal{N}\|_{sp} < \max_{\|\mathbf{w}\|=1} \min_{\|\mathbf{v}\|=1} \mathbf{v}^T \left(\sum_{r=1}^{K} \mathbf{w}(r)(\mathbf{T}_r + \mathbf{N}_r) \right) \mathbf{v}$$

then $\mathcal{T} + \mathcal{N}$ has a best rank R approximation among SFS tensors.

Here $\mathbf{T}_r + \mathbf{N}_r$ denotes the rth frontal slice of $\mathcal{T} + \mathcal{N}$.

Intuitively: If $\mathcal{T}$ has a positive definite slice mix, and the noise is small enough that it cannot destroy the positivity, then a best low rank approximation exists.

Spectral norm bound guaranteeing existence of best low rank approximation

Theorem [E-De Lathauwer]

Let $\mathcal{T}, \mathcal{N} \in \mathbb{R}^{R\times R\times K}$ and assume $\mathcal{T}$ has rank R and has SFS. If

$$\|\mathcal{N}\|_{sp} < \max_{\|\mathbf{w}\|=1} \min_{\|\mathbf{v}\|=1} \mathbf{v}^T \left(\sum_{r=1}^{K} \mathbf{w}(r)(\mathbf{T}_r + \mathbf{N}_r) \right) \mathbf{v}$$

then $\mathcal{T} + \mathcal{N}$ has a best rank R approximation among SFS tensors.

Consequence: Suppose you have some noisy rank R tenor $\mathcal{T} + \mathcal{N} \in \mathbb{R}^{R\times R\times K}$, and let $\hat{\mathcal{T}}$ be any rank R approximation to $\mathcal{T} + \mathcal{N}$. If

$$\|\mathcal{T} + \mathcal{N} - \hat{\mathcal{T}}\|_{sp} < \max_{\|\mathbf{w}\|=1} \min_{\|\mathbf{v}\|=1} \mathbf{v}^T \left(\sum_{r=1}^{K} \mathbf{w}(r)(\mathbf{T}_r + \mathbf{N}_r) \right) \mathbf{v}$$

then $\mathcal{T} + \mathcal{N}$ has a best rank R approximation.

Computing our bound

Theorem [E-De Lathauwer]

Let $\mathcal{T} \in \mathbb{R}^{R \times R \times K}$ and assume $\mathcal{T}$ has SFS. The quantity

$$\max_{\|\mathbf{w}\|=1} \min_{\|\mathbf{v}\|=1} \mathbf{v}^T \left(\sum_{r=1}^{K} \mathbf{w}(r)(\mathbf{T}_r + \mathbf{N}_r) \right) \mathbf{v}$$

is computable via semidefinite programming

Sharpness of the bound

Theorem [E-De Lathauwer]

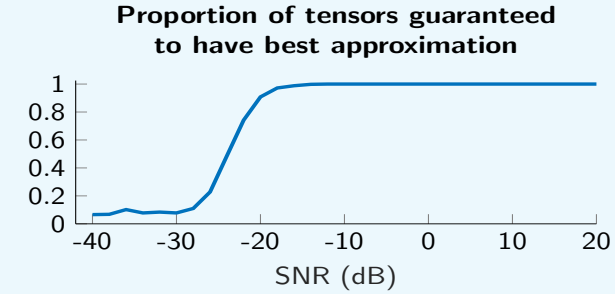
Let $\mathcal{T} \in \mathbb{R}^{R\times R\times K}$ and assume $\mathcal{T}$ has SFS rank R. Set

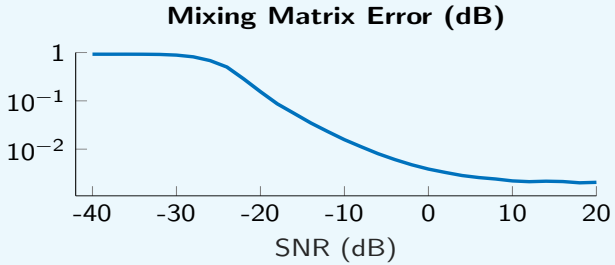
$$\lambda_* = \max_{\|\mathbf{w}\|=1} \min_{\|\mathbf{v}\|=1} \mathbf{v}^T \left(\sum_{r=1}^{K} \mathbf{w}(r)(\mathbf{T}_r + \mathbf{N}_r) \right) \mathbf{v}$$

and assume $\lambda_* \geq 0$. Then there exists a tensor $\mathcal{N}_* \in \mathbb{R}^{R\times R\times K}$ with $\|\mathcal{N}_*\|_{sp} = \lambda_*$ such that no linear combination of frontal slices of $\mathcal{T} + \mathcal{N}_*$ is positive definite.

Furthermore, if $K = 2$, then any open set containing $\mathcal{T} + \mathcal{N}_*$ contains a tensor which does not have a best rank R approximation.

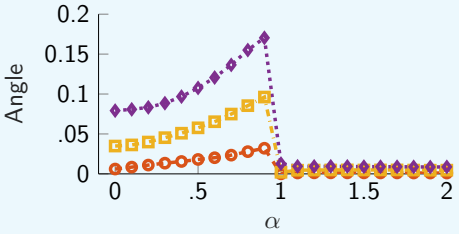
Numerical experiments: Second order blind identification



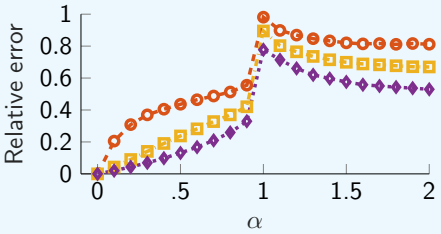


Bound sharpness vs. $4 \times 4 \times 2$ tensors. Approximations of $\mathcal{T} + \alpha \mathcal{N}_*$

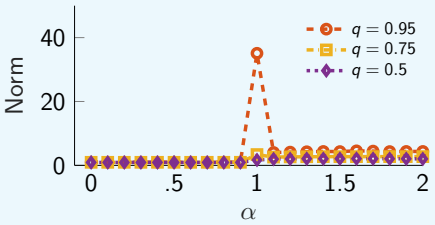
Min column angle quantiles



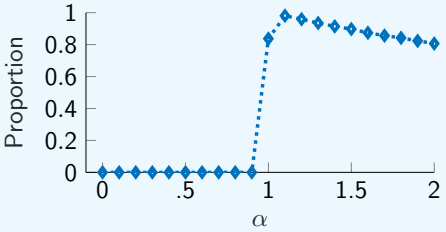
Factor matrix error quantiles



Max column norm quantiles

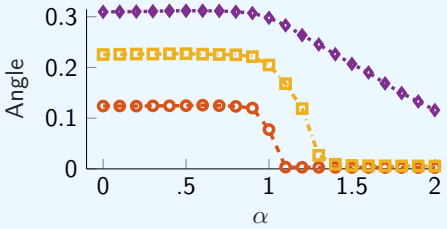


Proportion of MLSVD cores with repeated or complex generalized eigenvalues

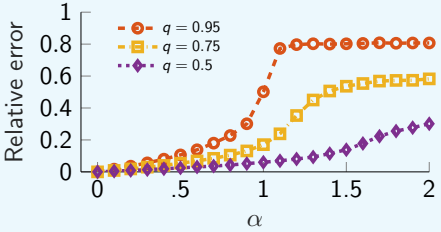


Bound sharpness vs. $4 \times 4 \times 4$ tensors. Approximations of $\mathcal{T} + \alpha \mathcal{N}_*$

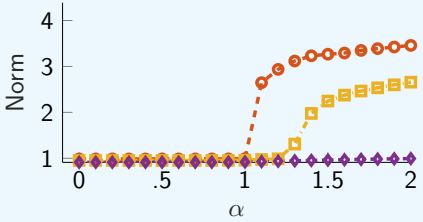
Min column angle quantiles



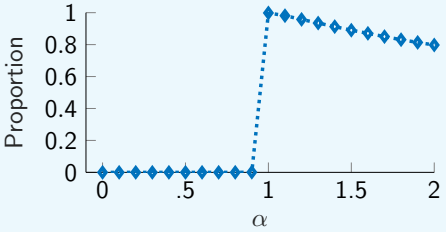
Factor matrix error quantiles



Max column norm quantiles



Proportion of MLSVD cores with repeated or complex generalized eigenvalues



Existence guarantees for unconstrained tensor decompositions

If $\mathcal{T} \in \mathbb{R}^{R \times R \times K}$ has rank R but does not have symmetric frontal slices, the the CPD of $\mathcal{T}$ can be seen as a joint generalized eigenvalue decomposition of $(\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_K)$

We show that if $\mathcal{T}$ has rank R but is a limit of tensors of rank $r < R$, then $\mathcal{T}$ is defective in the sense of this joint generalized eigenvalue problem.

In this case, every subpencil $(\mathcal{T}_i, \mathcal{T}_j)$ is defective in the sense of the generalized eigenvalue problem (i.e, has eigenvalues with algebraic multiplicity greater than geometric multiplicity).

Perturbation theoretic bounds for the generalized eigenvalue problem therefore lead to existence guarantees for the best low rank approximations of $\mathcal{T} + \mathcal{N}$ where $\mathcal{T}$ has rank R.

Generalized eigenvector decomposition (GEVD) gives CPD of a rank R tensor.

Recall the CPD of of rank R tensor $\mathcal{T} \in \mathbb{R}^{R \times R \times K}$ is

$$\sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r.$$

Key idea: Columns of

$$\mathbf{B}^{-\mathsf{T}} := \begin{pmatrix} \uparrow & & \uparrow \\ \mathbf{b}_1 & \cdots & \mathbf{b}_R \\ \downarrow & & \downarrow \end{pmatrix}^{-\mathsf{T}} \in \mathbb{R}^{R \times R}$$

are equal to eigenvectors of $\mathbf{T}_k^{-1}\mathbf{T}_\ell$ which in turn are equal to generalized eigenvectors of the **matrix pencil** $(\mathbf{T}_k, \mathbf{T}_\ell)$, i.e. vectors $\mathbf{x}$ such that

$$\mathbf{T}_k\mathbf{x} = \lambda_k\mathbf{y} \qquad \text{and} \qquad \mathbf{T}_\ell\mathbf{x} = \lambda_\ell\mathbf{y}$$

$\implies$ Generalized eigenvector decomp. of $(\mathbf{T}_k, \mathbf{T}_\ell)$ leads to CPD of $\mathcal{T}$.

Small eigenvalue gaps lead to inaccuracy.

Gen. eigenvalues of $(\mathbf{T}_k, \mathbf{T}_\ell)$ are interpreted as points on the unit circle. The pencil $(\mathbf{T}_k, \mathbf{T}_\ell)$ has R generalized eigenvalues.

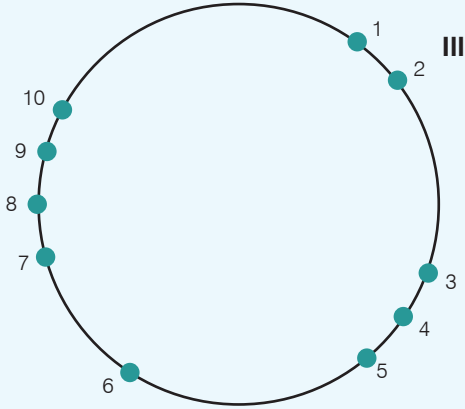


Illustration of generalized eigenvalues of $(\mathbf{T}_k, \mathbf{T}_\ell)$

● = generalized eigenvalue of $(\mathbf{T}_k, \mathbf{T}_\ell)$.

The small gap between generalized eigenvalues 1 and 2 leads to instability in computing the generalized eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$.

Similar issues occur in the other clusters of generalized eigenvalues.

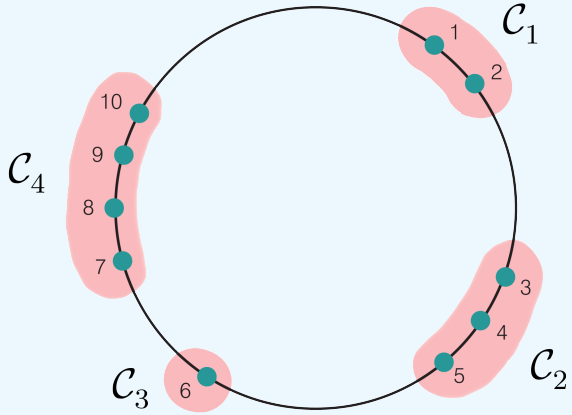
Using a single pencil causes instability

Projection of $\mathcal{T} \cong (\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_K)$ to $(\mathbf{T}_k, \mathbf{T}_\ell)$ is fundamental source of instability in GEVD (quantified in work of Beltrán, Breiding, Vannieuwenhoven).

Projection $\mathcal{T}$ to a pencil is equivalent to a projection of vectors $\mathbf{c}_r \in \mathbb{R}^K$ to $\mathbb{R}^2$
$\implies$ Information is lost and distance between vectors decreases under projection.

We combat both sources of inaccuracy by using many pencils for CPD computation.

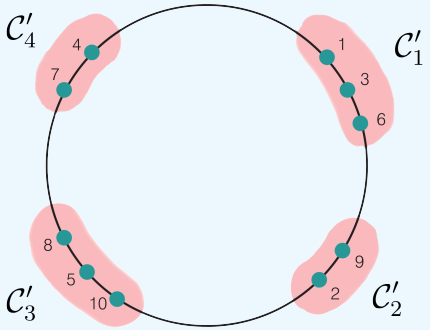
Generalized EigenSpace Decomp: Improve accuracy by computing eigenspaces corresponding to well separated eigenvalue clusters.



Clusters $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4$ are well separated so can improve accuracy by only computing the corresponding eigenspaces $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4$.

Use a new pencil to split eigenspaces!

Consider a new subpencil $(\mathbf{T}_m, \mathbf{T}_n)$. The eigenvectors of this pencil are the same as those of $(\mathbf{T}_k, \mathbf{T}_\ell)$, but the corresponding eigenvalues will lie in new positions on the unit circle.



The clusters $\mathcal{C}'_1, \mathcal{C}'_2, \mathcal{C}'_3, \mathcal{C}'_4$ are well separated, so can compute the eigenspaces $\mathcal{E}'_1, \mathcal{E}'_2, \mathcal{E}'_3, \mathcal{E}'_4$.

Observe $\mathcal{E}_1 = \text{span}\{\mathbf{v}_1, \mathbf{v}_2\}$ and $\mathcal{E}'_1 = \text{span}\{\mathbf{v}_1, \mathbf{v}_3, \mathbf{v}_6\}$. Thus $\mathbf{v}_1 = \mathcal{E}_1 \cap \mathcal{E}'_1$.

GESD recursively deflates tensor rank.

In our implementation, GESD recursively writes $\mathcal{T}$ as a sum of tensors of reduced rank.

In the example, GESD would use $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4$ to write the rank 10 tensor $\mathcal{T}$ as

$$\mathcal{T} = \mathcal{T}^1 + \mathcal{T}^2 + \mathcal{T}^3 + \mathcal{T}^4$$

where $\mathcal{T}^1, \mathcal{T}^2, \mathcal{T}^3$ and $\mathcal{T}^4$ have ranks $2, 3, 1$ and 4, respectively. $\mathcal{T}^1$ can then be decomposed into a sum of rank 1 tensors using the pencil $(\mathcal{T}^1_m, \mathcal{T}^1_n)$, etc.

Variations in GESD are possible. E.g. one could compute intersections of eigenspaces as described above rather than working recursively.

GESD vs direction of arrival retrieval

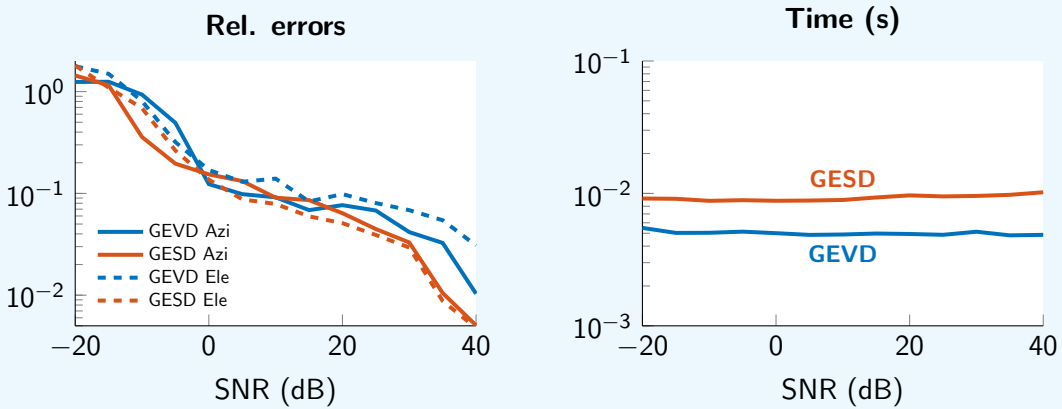
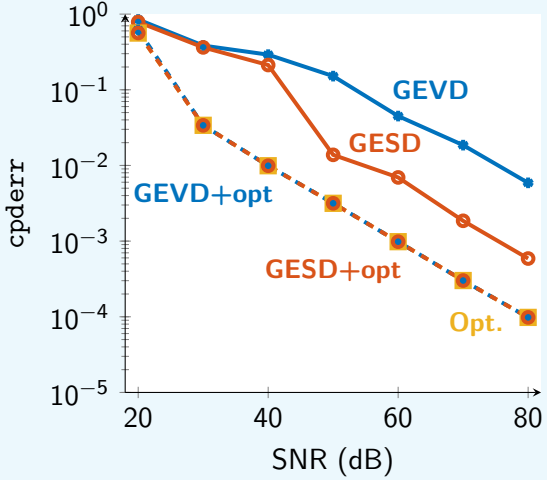
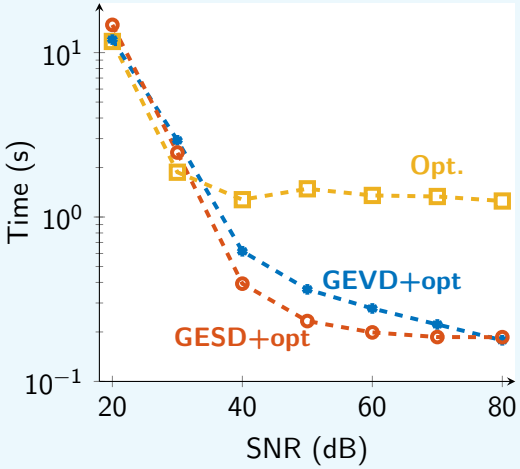


Figure: GESD estimates the azimuths and elevations of the sources more accurately than GEVD, and is only slightly slower. Left: mean relative errors over the sources of the estimated azimuths (——) and elevations (- - -) for GEVD and the estimated azimuths (——) and elevations (- - -) for GESD. Right: computation time for both methods.

GESD vs synthetic data





Accuracy and speed against Rank 10 tensors of size $100 \times 100 \times 100$ with highly correlated factor matrix columns.