Existence of best low rank tensor approximations

Eric Evert

Joint work with Lieven De Lathauwer

2 July 2021

KU LEUVEN kulak

fwo

EOS
THE EXCELLENCE
OF SCIENCE

fnrs
LA LIBERTÉ DE CHERCHER

Decompose signal into canonical components.

A **tensor** $\mathcal{T}$ is a multiindexed array.

$\mathcal{T} =$  $\in \mathbb{R}^{N_1 \times N_2 \times N_3}$

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

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

R is the **rank** of $\mathcal{T}$.

Notation: $\mathbf{a}_r, \mathbf{b}_r, \mathbf{c}_r$ are vectors of length N_1, N_2, N_3, respectively.

In practice measured signal tensors are corrupted by noise. Must compute a low rank approximation

A common problem in applications: Given $\mathcal{M} = \mathcal{T} + \mathcal{N}$ of size $I_1 \times I_2 \times I_3$ where $\mathcal{T}$ is a rank R signal tensor, compute a best rank (less than or equal to) R approximation of $\mathcal{M}$.

Intuition: Computing rank R approximation allows (approximate) recovery of $\mathcal{T}$ from $\mathcal{M}$.

Low rank tensor approximation is ill-posed: A best rank $\leq R$ approximation may not exist.

This leads to considering the set of border rank $\leq R$ tensors. This is the closure of the set of rank R tensors.

What is border rank and why?

Why care:

1. Optimization always has solution over border rank R tensors.
2. Decomposition is uninterpretable if solution has rank $>$ border rank.

Tensor phenomena: Limit of rank 2 tensors could have rank 3.

E.g.

$$\lim_{n\to\infty} n(\mathbf{e}_1)^{\otimes 3} - n\left(\mathbf{e}_1 + \frac{\mathbf{e}_2}{n}\right)^{\otimes 3}$$

In general: $\mathcal{T}$ has border rank R means:

1. $\mathcal{T}$ is a limit of rank R tensors.
2. $\mathcal{T}$ is not a limit of tensors having rank $<$ R.

Commonly proposed "solutions" if best approximation has rank $>$ border rank have issues.

Common suggestions are:

1. Take close to optimal approximation: Suffers from diverging components
2. Increase rank: Can lose uniqueness of decomposition.
3. Impose constraints: Solution will always be on boundary of constraint.

Recap

Tensor approximation always has best approximation over border rank R solutions

Rank $>$ border rank leads to big problems for interpretation

1. Diverging components
2. Loss of uniqueness
3. Artificially chosen solution.

Need good, meaningful solutions!
Need guarantee solution has rank $=$ border rank.

Start by understanding $R \times R \times 2$ examples. A good case: Distinct generalized eigenvalues

Write $\mathbf{T}_1 = \mathcal{T}(:,:,1)$ and $\mathbf{T}_2 = \mathcal{T}(:,:,2)$.

Let $\mathcal{T}$ be defined by

$$\mathbf{T}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{T}_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{has} \qquad \mathbf{T}_2^{-1}\mathbf{T}_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$\mathbf{T}_2^{-1}\mathbf{T}_1$ eigenvalues 1 and -1 and eigenvectors $\mathbf{e}_1$ and $\mathbf{e}_2$.

The tensor has real rank 2.

A bad case: Repeated generalized eigenvalues

Let $\mathcal{T}$ be defined by

$$\mathbf{T}_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{T}_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{has} \qquad \mathbf{T}_2^{-1}\mathbf{T}_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$\mathbf{T}_2^{-1}\mathbf{T}_1$ eigenvalue 0 (with algebraic multiplicity 2) and eigenvector $\mathbf{e}_1$.

The tensor has rank > 2 and border rank 2.

This tensor is the limit of the first example I showed.

Story for $R \times R \times 2$ tensors is completely told by generalized eigenvalues.

The generalized eigenvalues and generalized eigenvectors of $\mathcal{T}$ are (essentially) equal to the classical eigenvalues and eigenvectors of the matrix

$$\mathbf{T}_2^{-1}\mathbf{T}_1$$

Theorem: $\mathcal{T} \in \mathbb{R}^{R\times R\times 2}$ has rank R IFF $\mathcal{T}$ has a basis of generalized eigenvectors.

Idea: To guarantee rank $=$ border rank, use perturbation theory for generalized eigenvalues to guarantee perturbation has distinct generalized eigenvalues.

An existence bound for $R \times R \times 2$ tensors

Theorem

Let $\mathcal{T}$ and $\hat{\mathcal{T}}$ be tensors of size $R \times R \times 2$. Assume that $\mathcal{T}$ has $\mathbb{R}$-rank R with CPD $[\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$. If

$$\|\mathcal{T} - \hat{\mathcal{T}}\|_{sp} < \frac{\sigma_{\min}(\mathbf{A})\sigma_{\min}(\mathbf{B}) \min_{i \neq j} \chi(\mathbf{C}_i, \mathbf{C}_j)}{2},$$

then $\hat{\mathcal{T}}$ has $\mathbb{R}$-rank R and

$$md[\mathcal{T}, \hat{\mathcal{T}}] < \frac{\|\mathcal{T} - \hat{\mathcal{T}}\|_{sp}}{\sigma_{\min}(\mathbf{A})\sigma_{\min}(\mathbf{B})}. \tag{1}$$

A perturbation bound for $R \times R \times K$ tensors. The tensor Bauer-Fike theorem.

Theorem

Let $\mathcal{T}, \hat{\mathcal{T}} \in \mathbb{R}^{R \times R \times K}$ be rank R tensors and let $\mathcal{T} = [\![\mathbf{A}, \mathbf{B}, \mathbf{C}]\!]$. Then

$$sv[\mathcal{T}, \hat{\mathcal{T}}] \le \sqrt{R} \|(\mathcal{T} - \tilde{\mathcal{T}}) \cdot_1 \mathbf{A}^{-1} \cdot_2 \mathbf{B}^{-1}\|_{sp} \le \frac{\sqrt{R}\|\mathcal{T} - \hat{\mathcal{T}}\|_{sp}}{\sigma_{\min}(\mathbf{A})\sigma_{\min}(\mathbf{B})}.$$

Getting an existence bound in the $R \times R \times K$ setting

Key ideas:

1. Show best border rank R approximation $\hat{\mathcal{T}}$ has rank R.

2. If $\hat{\mathcal{T}}$ has border rank $<$ rank, then every subpencil of $\hat{\mathcal{T}}$ must have a repeated generalized eigenvalue.

3. If $\hat{\mathcal{T}}$ has border rank $<$ rank, then so does $\hat{\mathcal{T}} \cdot_3 \mathbf{U}$ for any invertible matrix $\mathbf{U}$.

A multiple pencil based bound for existence

Theorem

Let $\mathcal{M} \in \mathbb{R}^{R \times R \times K}$ be any tensor. For each $i = 1, \ldots, \lfloor K/2 \rfloor$, let $\epsilon_i \geq 0$ the bound computed using the $K = 2$ theorem for the pencil

$$(\mathbf{M}_{2i-1}, \mathbf{M}_{2i})$$

and set $\epsilon = ||(\epsilon_1, \ldots, \epsilon_{\lfloor K/2 \rfloor})||_2$. If there exists some $\mathbb{R}$-rank R tensor $\mathcal{T}'$ such that

$$\|\mathcal{M} - \mathcal{T}'\|_F < \epsilon,$$

then $\mathcal{M}$ has a best $\mathbb{R}$-rank R approximation and any best $\mathbb{R}$-rank R approximation of $\mathcal{M}$ has a unique CPD.

Explanation of the Theorem

1. A best border rank R approximation $\hat{\mathcal{T}}$ of $\mathcal{M}$ is guaranteed to exist.

2. Have $\|\hat{\mathcal{T}} - \mathcal{M}\|_F \leq \|\mathcal{T}' - \mathcal{M}\|_F < \epsilon$

3. This implies there is an index i so that

$$\|(\hat{\mathbf{T}}_{2i-1}, \hat{\mathbf{T}}_{2i}) - (\mathbf{M}_{2i-1}, \mathbf{M}_{2i})\| < \epsilon_i$$

4. The subpencil $(\hat{\mathbf{T}}_{2i-1}, \hat{\mathbf{T}}_{2i})$ cannot have a repeated generalized eigenvalue.

5. The tensor $\hat{\mathcal{T}}$ must have rank $=$ border rank $= R$.

A multiple pencil based bound for existence (improved version)

Theorem

Let $\mathcal{M} \in \mathbb{R}^{R\times R\times K}$ be any tensor. Let $\mathbf{U} \in \mathbb{K}^{K\times K}$ be a unitary matrix and set $\mathcal{S} = \mathcal{M} \cdot_3 \mathbf{U}$. For each $i = 1, \ldots, \lfloor K/2 \rfloor$, let $\epsilon_i \geq 0$ the bound computed using the $K = 2$ theorem for the pencil

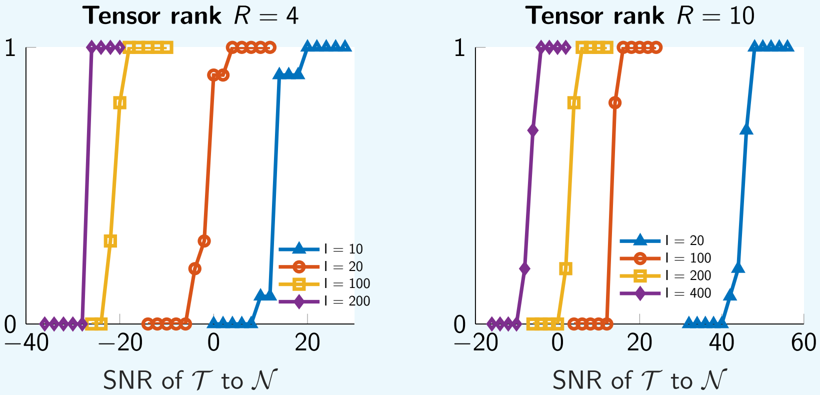
$$(\mathbf{S}_{2i-1}, \mathbf{S}_{2i})$$

and set $\epsilon = ||(\epsilon_1, \ldots, \epsilon_{\lfloor K/2 \rfloor})||_2$. If there exists some $\mathbb{R}$-rank R tensor $\mathcal{T}'$ such that

$$\|\mathcal{M} - \mathcal{T}'\|_F < \epsilon,$$

then $\mathcal{M}$ has a best $\mathbb{R}$-rank R approximation and any best $\mathbb{R}$-rank R approximation of $\mathcal{M}$ has a unique CPD.

SNR at which tensors of various sizes are guaranteed
(in experiments) to have a best rank R approximation



Proportion of $I \times I \times I$ tensors $\mathcal{T} + \mathcal{N}$ with truncated MLSVD guaranteed to have a best rank R approximation

Bonus: The tensor Procrustes problem

This talk: $R \times R \times K$ tensors. In general $I_1 \times I_2 \times I_3$ is okay.

Key ingredient: The tensor Procrustes problem.

Theorem

Let $\mathcal{T}, \hat{\mathcal{T}} \in \mathbb{R}^{I_1 \times \cdots \times I_\ell}$ be tensors having the same multilinear rank. Then there exist orthogonal compressions $\mathcal{T}^c, \hat{\mathcal{T}}^c$ of $\mathcal{T}$ and $\hat{\mathcal{T}}$, respectively, such that

$$\|\mathcal{T}^c - \hat{\mathcal{T}}^c\|_F \leq \|\mathcal{T} - \hat{\mathcal{T}}\|_F.$$

Contributions

Deterministic method for showing that a tensor has a best rank R approximation.

Tensor decomposition is stable and well-posed in a computable neighborhood around a given rank R tensor.

Guarantees that a decomposition will not suffer from the dreaded diverging components or loss of uniqueness.