(Bio)Tensor Decompositions: Introduction and Overview

Lieven De Lathauwer

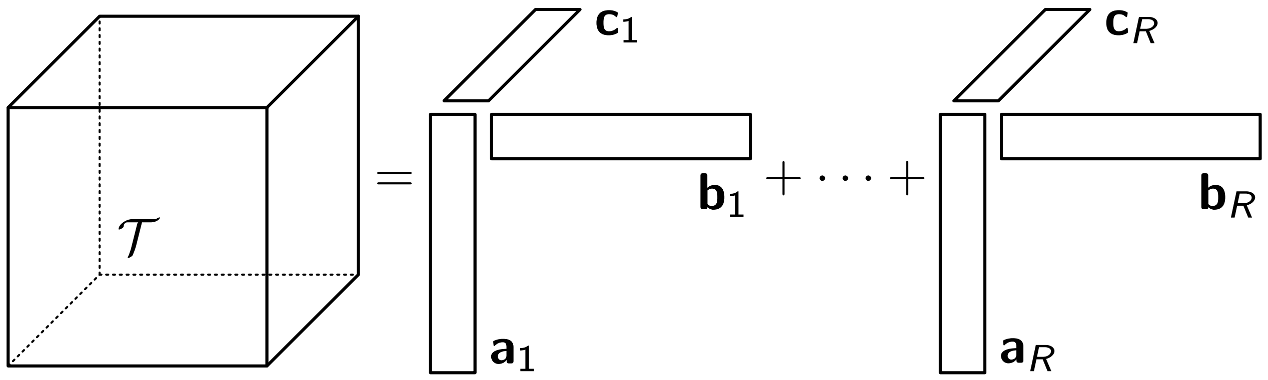
KU Leuven
Belgium

Lieven.DeLathauwer@kuleuven-kulak.be

Canonical Polyadic Decomposition

Rank: minimal number of rank-1 terms [*Hitchcock, 1927*]

Canonical Polyadic Decomposition (CPD): decomposition in minimal number of rank-1 terms [*Harshman '70*], [*Carroll and Chang '70*]

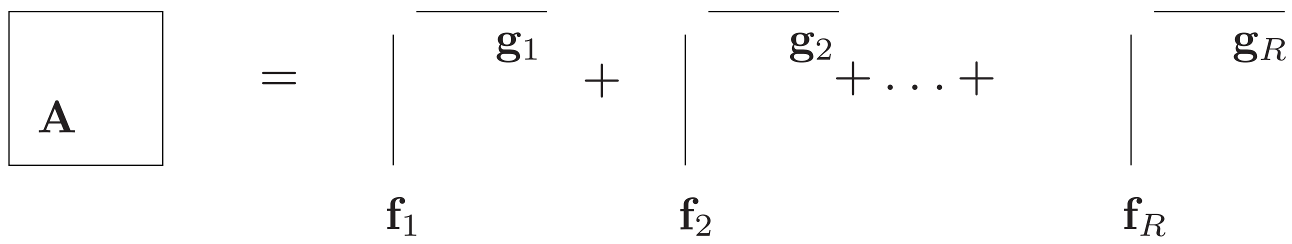


- Unique under mild conditions on number of terms and differences between terms
- Orthogonality (triangularity, . . .) not required (but may be imposed)

Factor Analysis and Blind Source Separation

- Decompose a data matrix in rank-1 terms that can be interpreted
  E.g. statistics, telecommunication, biomedical applications, chemometrics, data analysis, . . .

$$\mathbf{A} = \mathbf{F} \cdot \mathbf{G}^T$$



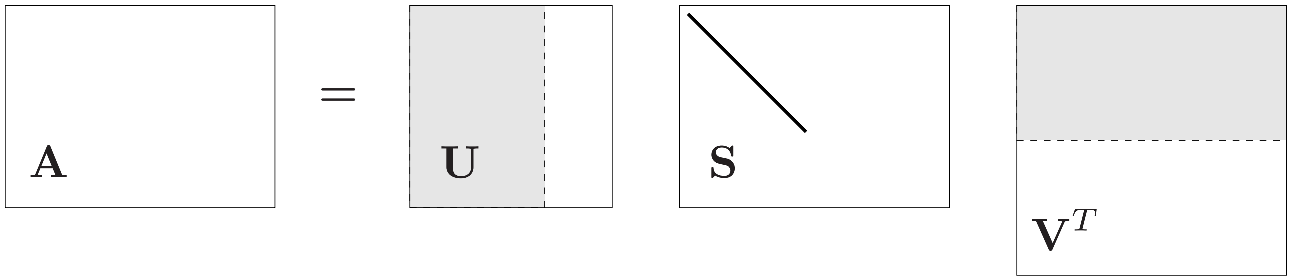
- $\mathbf{F}$: mixing matrix
  $\mathbf{G}$: source signals
  Decompose a data matrix in rank-1 terms that can be interpreted

What about SVD?

- SVD is unique
- . . . thanks to orthogonality constraints

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V}^T = \sum_{r=1}^{R} s_{rr} \mathbf{u}_r \mathbf{v}_r^T$$

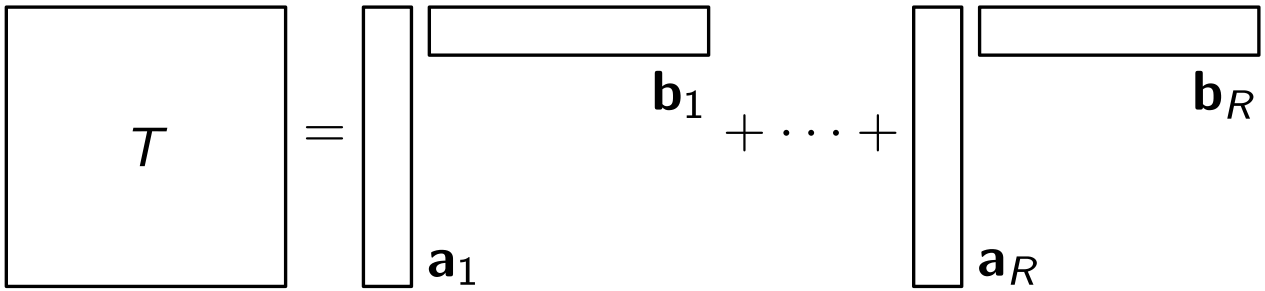
  $\mathbf{U}$, $\mathbf{V}$ orthogonal, $\mathbf{S}$ diagonal
- Whether these constraints make sense, depends on the application
- SVD is great for dimensionality reduction
  best rank-R approximation $\leftarrow$ truncated SVD



Motivating example: excitation-emission fluorescence in chemometrics

Matrix approach

row vector ∼ emission spectrum
column vector ∼ excitation spectrum

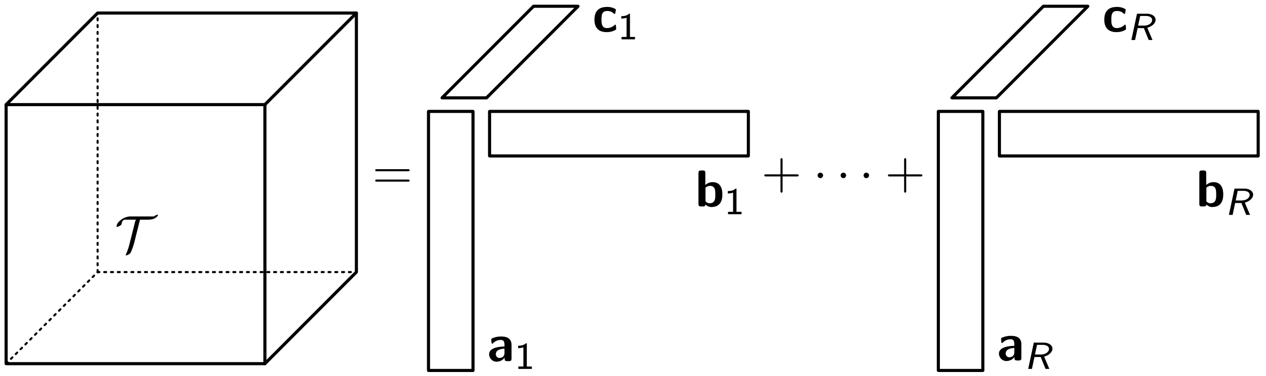


NMF not unique in general

Tensor solution: CP Analysis

Tensorization: one matrix $\rightarrow$ several matrices, stacked in tensor

row vector $\sim$ emission spectrum
column vector $\sim$ excitation spectrum
coefficients $\sim$ concentrations

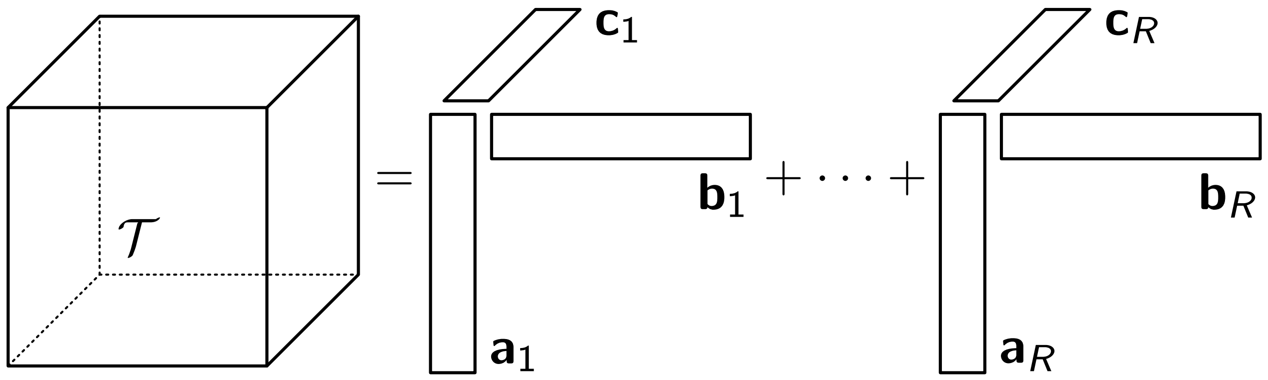


[*Smilde, Bro, Geladi '04*]

Canonical Polyadic Decomposition

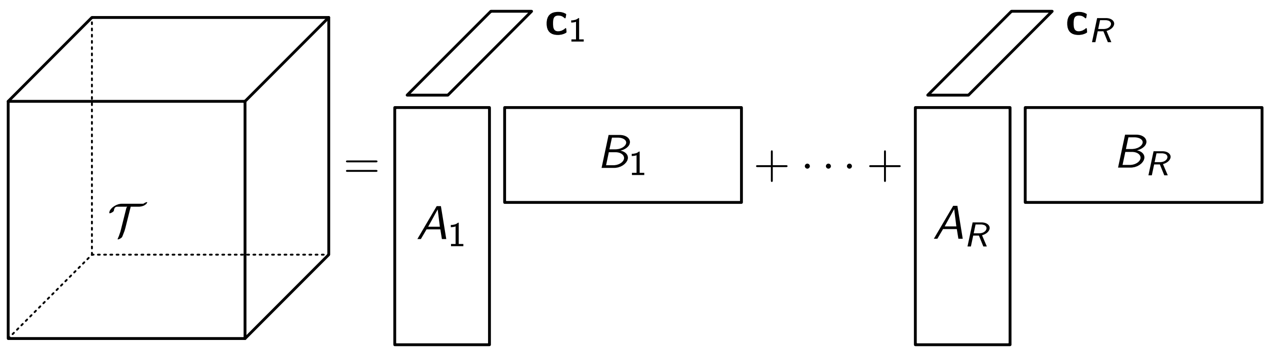
Rank: minimal number of rank-1 terms [*Hitchcock, 1927*]

Canonical Polyadic Decomposition (CPD): decomposition in minimal number of rank-1 terms [*Harshman '70*], [*Carroll and Chang '70*]



- Unique under mild conditions on number of terms and differences between terms
- Orthogonality (triangularity, . . .) not required (but may be imposed)

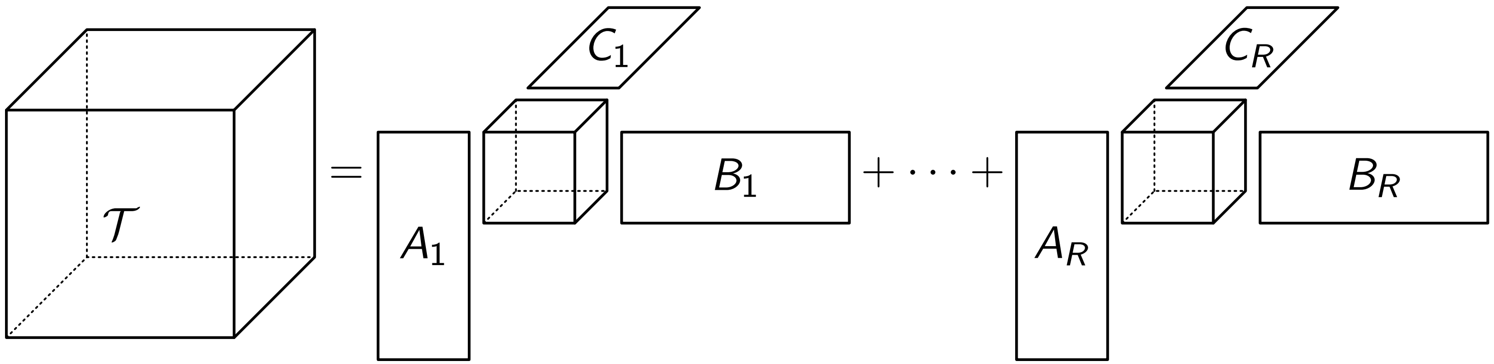
Decomposition in rank-$(L, L, 1)$ terms



Unique under mild conditions

[*DL '08*]

Decomposition in rank-(R_1, R_2, R_3) terms



Unique under mild conditions

Rank-1 term $\sim$ data atom
Block term $\sim$ data molecule

[*DL '08*]

Constraints

Examples: orthogonality [*Sørensen and DL '12*]
nonnegativity [*Cichocki et al. '09*]
Vandermonde [*Sørensen and DL '12*]
independence [*De Vos et al. '12*]
...

Not needed for uniqueness in tensor case

Pro: relaxed uniqueness conditions
easier interpretation
no degeneracy (NN, orthogonality)
higher accuracy

Depending on type of constraints, lower or higher computational cost

Coupled matrix/tensor decompositions

One or more matrices

One or more tensors

Symmetric and nonsymmetric

One or more factors shared (or parts of factors, or generators)

Constraints (orthogonal, nonnegative, exponential, constant modulus, polynomial, rational, Toeplitz, Hankel, . . .)

Data fusion

Numerical optimization

- Most popular: ALS
- Quasi-Newton / NLS
- Exact line / plane search
- Constraints
- Structured data fusion
- Complex optimization

Subproject tensor decompositions

- WP 1: Computing (constrained) tensor decompositions
- WP 2: Updating tensor decompositions
- WP 3: Coupling tensor decompositions
- WP 4: Software platform for tensor-based biomedical blind source separation

Computation: further steps

- (In)equality constraints
- Probabilistic factorizations:
  - Euclidean $\rightarrow$ beta-divergence
  - maximum likelihood $\rightarrow$ maximum a posteriori
- Model selection
  - number of terms
  - block dimensions

Big data variants

- Decompositions of incomplete tensors
- Tools from scientific computing
- Stochastic NLS
- Strategies for (compressive) sampling

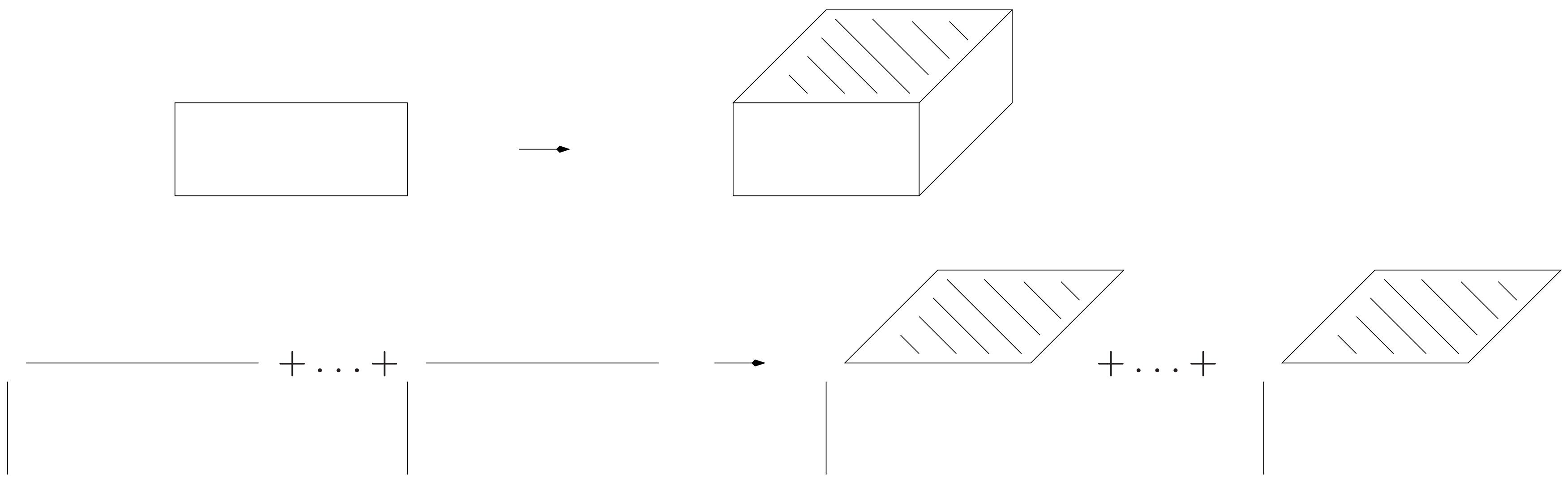
Tensorization

- Hankel: exponential polynomials
- Loewner: rational functions
- Nonnegativity constraints: unique NMF
- Convolutive variants

Deterministic BSS $\leftrightarrow$ ICA

Exponentials, sinusoids, polynomials, exponential polynomials

Principle: Map every row of $\mathbf{T} = \mathbf{A} \cdot \mathbf{B}^T$ to Hankel matrix
Hankel matrices are often very ill-conditioned
Hankel matrices generated by exponential polynomials are exactly low-rank



Tensorlab vX.0

- Core routines in Fortran

Tensorlab — a MATLAB toolbox for tensor decompositions
esat.kuleuven.be/sista/tensorlab

- Elementary operations on tensors
  Multicore-aware and profiler tuned
- Tensor decompositions with structure and/or symmetry
  CPD, LMLRA, MLSVD, block term decompositions
- Global minimization of bivariate polynomials
  Exact line and plane search for tensor optimization
- Cumulants, tensor visualization, estimating a tensor's rank or multilinear rank, . . .

Tensorlab v2.0

Major upgrade which brings:

- Full support for sparse and incomplete tensors
- Major improvements in computational and memory efficiency
- Structured data fusion

  Structured: choose from a large library of constraints to impose on factors (nonnegative, orthogonal, Toeplitz, . . .)

  Data fusion: jointly factorize multiple data sets