Unmixing of rational functions by tensor computations

Lieven De Lathauwer and Marc Van Barel

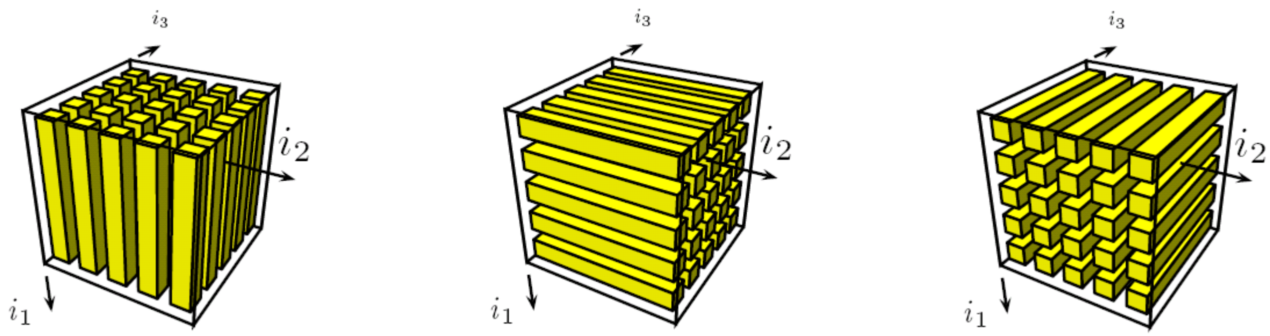
K.U.Leuven
Belgium
Lieven.DeLathauwer@kuleuven-kortrijk.be
Marc.VanBarel@cs.kuleuven.be

Overview

- Preliminaries
- Tensor decompositions
- Factor analysis and signal separation
- Block Term Decompositions and Block Component Analysis

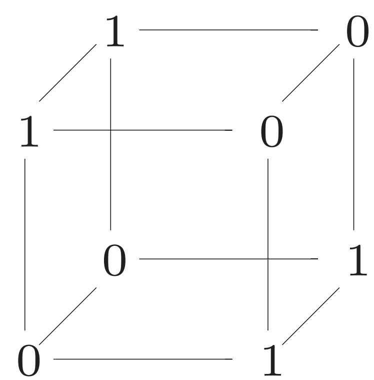
Columns, rows and mode-n vectors

Mode-n vectors of a tensor: generalization of column/row vectors of a matrix



Multilinear rank of a tensor

- The column (row) rank of a matrix $\mathbf{A}$ is equal to the maximal number of columns (rows) of $\mathbf{A}$ that form a linearly independent set
- Mode-n rank of a tensor: dimension of the vector space generated by mode-n vectors
- Mode-n ranks can be mutually different
- Rank-(R_1, R_2, R_3) tensor: $\text{rank}_1(\mathcal{A}) = R_1$, $\text{rank}_2(\mathcal{A}) = R_2$, $\text{rank}_3(\mathcal{A}) = R_3$
- Multilinear rank: (R_1, R_2, R_3)



$R_1 = R_2 = 2 \quad R_3 = 1$

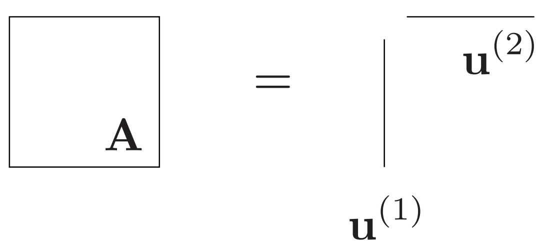
Rank-1 tensor

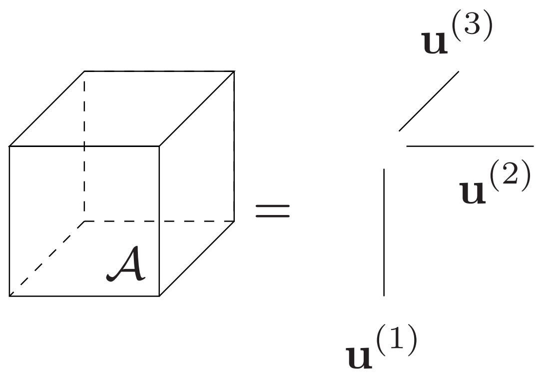
- Rank-1 matrix: outer product of 2 vectors $\mathbf{u}^{(1)}$, $\mathbf{u}^{(2)}$:

$$\begin{aligned} a_{i_1 i_2} &= u^{(1)}_{i_1} u^{(2)}_{i_2} \\ \mathbf{A} &= \mathbf{u}^{(1)} \cdot \mathbf{u}^{(2)^T} \equiv \mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \end{aligned}$$

- Rank-1 tensor: outer product of N vectors $\mathbf{u}^{(1)}$, $\mathbf{u}^{(2)}$, . . . , $\mathbf{u}^{(N)}$:

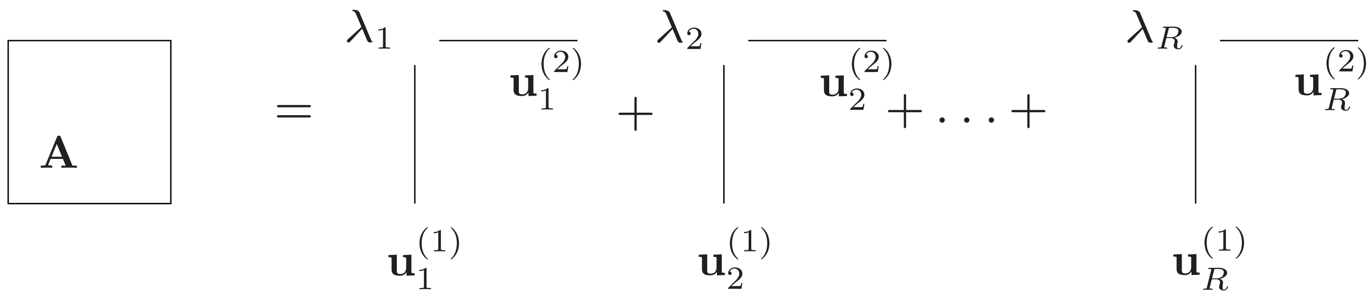
$$\begin{aligned} a_{i_1 i_2 \dots i_N} &= u^{(1)}_{i_1} u^{(2)}_{i_2} \dots u^{(N)}_{i_N} \\ \mathcal{A} &= \mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ \dots \circ \mathbf{u}^{(N)} \end{aligned}$$



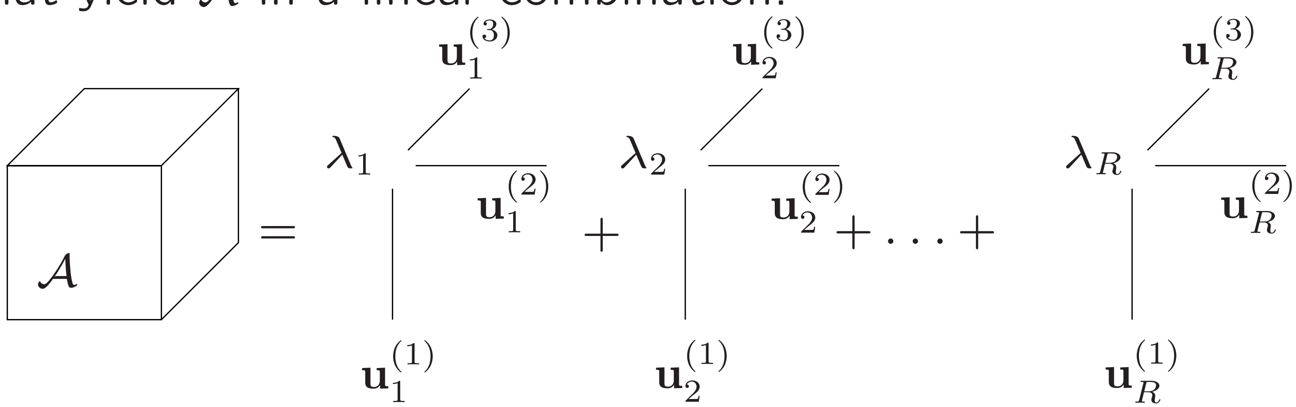


Rank of a tensor

- The rank R of a matrix $\mathbf{A}$ is minimal number of rank-1 matrices that yield $\mathbf{A}$ in a linear combination.



- The rank R of an Nth-order tensor $\mathcal{A}$ is the minimal number of rank-1 tensors that yield $\mathcal{A}$ in a linear combination.



[*Hitchcock, 1927*]

Overview

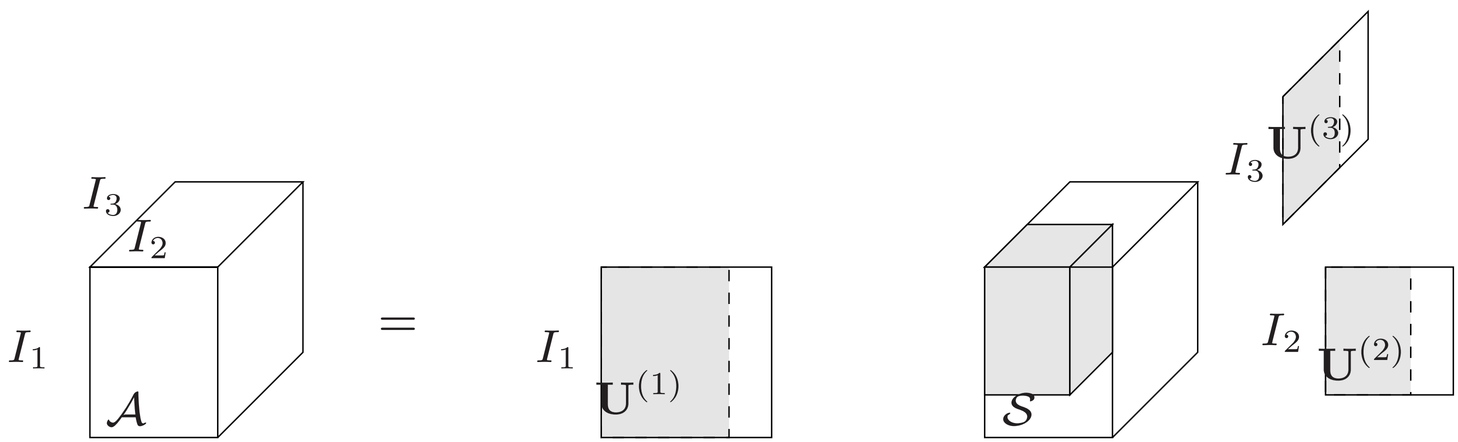
- Preliminaries
- Tensor decompositions:
  - Tucker decomposition / Multilinear SVD
  - Parallel Factor Decomposition
- Factor analysis and signal separation
- Block Term Decompositions and Block Component Analysis

Multilinear rank and associated decomposition

Definition:

$$\mathcal{A} = \mathcal{S} \bullet_1 \mathbf{U}^{(1)} \bullet_2 \mathbf{U}^{(2)} \bullet_3 \ldots \bullet_N \mathbf{U}^{(N)}$$

in which $\mathcal{S}$ is all-orthogonal and ordered
$\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \ldots, \mathbf{U}^{(N)}$ are orthogonal



[*Tucker '64*], [*De Lathauwer '00*]

Computation

$$\mathcal{A} = \mathcal{S} \bullet_1 \mathbf{U}^{(1)} \bullet_2 \mathbf{U}^{(2)} \bullet_3 \mathbf{U}^{(3)}$$

- $(I_1 \times I_2 I_3)$ matrix $\mathbf{A}^{(1)}$ in which all the columns are stacked

$$\text{SVD: } \mathbf{A}^{(1)} = \mathbf{U}^{(1)} \cdot \mathbf{\Sigma}^{(1)} \cdot \mathbf{V}^{(1)^T}$$

- $(I_2 \times I_3 I_1)$ matrix $\mathbf{A}^{(2)}$ in which all the row vectors are stacked

$$\text{SVD: } \mathbf{A}^{(2)} = \mathbf{U}^{(2)} \cdot \mathbf{\Sigma}^{(2)} \cdot \mathbf{V}^{(2)^T}$$

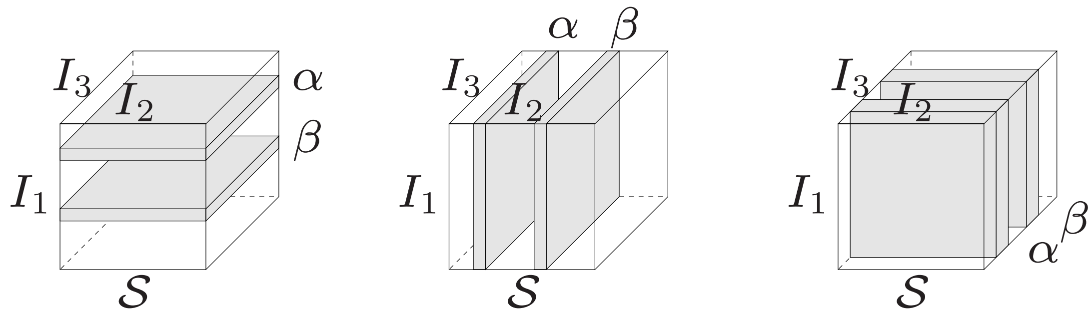
- $(I_3 \times I_1 I_2)$ matrix $\mathbf{A}^{(3)}$ in which all the mode-3 vectors are stacked

$$\text{SVD: } \mathbf{A}^{(3)} = \mathbf{U}^{(3)} \cdot \mathbf{\Sigma}^{(3)} \cdot \mathbf{V}^{(3)^T}$$

- Compute $\mathcal{S}$:

$$\mathcal{S} = \mathcal{A} \bullet_1 \mathbf{U}^{(1)^T} \bullet_2 \mathbf{U}^{(2)^T} \bullet_3 \mathbf{U}^{(3)^T}$$

All-orthogonality:

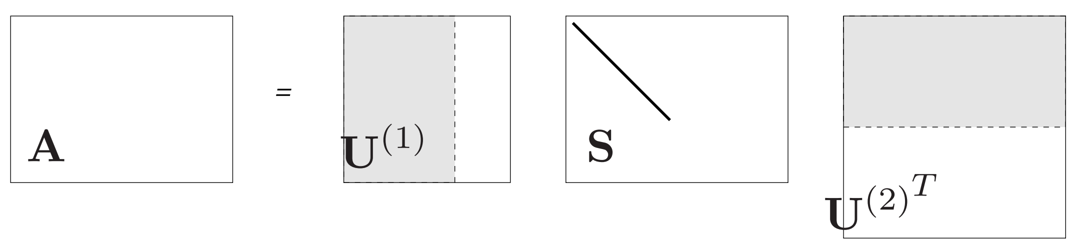


All-orthogonality is a generalization of diagonality

Ordering: slices have decreasing Frobenius norm

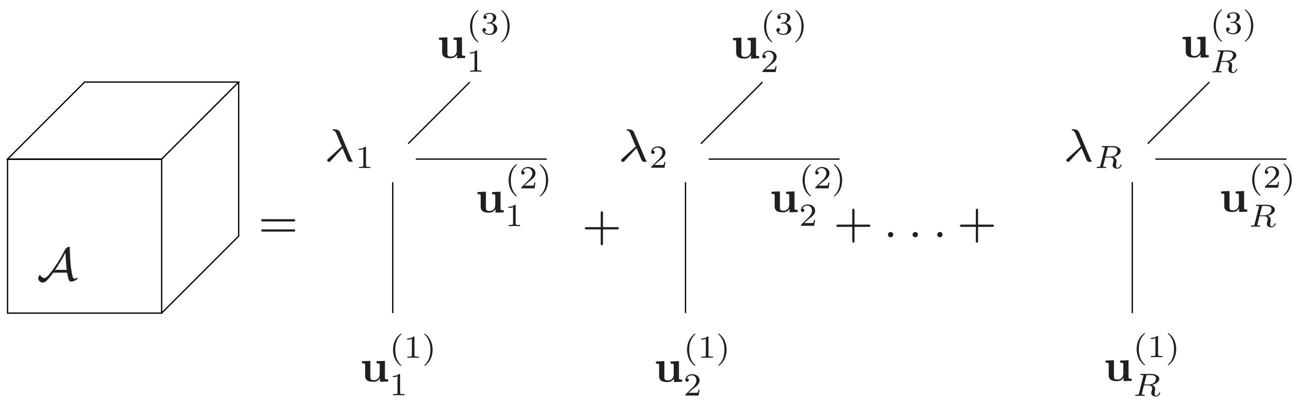
Norms of slices = mode-n singular values

Matrix SVD:



CANDECOMP/PARAFAC

Canonical Decomposition / Parallel Factor Decomposition / Canonical Polyadic Decomposition of a tensor $\mathcal{A}$ is its decomposition in a minimal sum of rank-1 tensors



[*Harshman '70*], [*Carroll and Chang '70*]

Unique under mild conditions

Algorithms

$$f(\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \mathbf{U}^{(3)}) = \|\mathcal{A} - \sum_{r=1}^{R} \mathbf{u}_r^{(1)} \circ \mathbf{u}_r^{(2)} \circ \mathbf{u}_r^{(3)}\|^2$$

- Alternating least squares (ALS) [*Harshman '70*]
- ALS with Exact Line Search [*Rajih et al. '08*], [*Nion and De Lathauwer '08*]
- ALS with regularization [*Navasca et al. '08*]
- general-purpose optimization:
  - Levenberg-Marquardt
  - conjugate gradient [*Paatero '99*], [*Acar et al. '09*]
  - ...
- EVD [*Leurgans et al. '93*], ...
- simultaneous generalized Schur [*De Lathauwer et al. '04*]
- simultaneous matrix diagonalization [*De Lathauwer '06*]
- ...

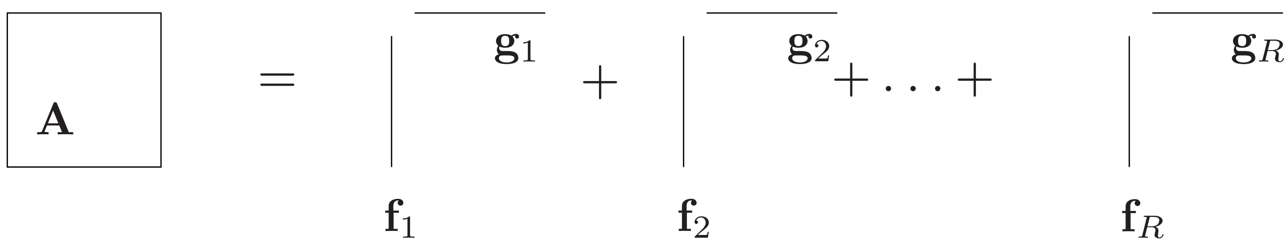
Overview

- Preliminaries
- Tensor decompositions
- Factor analysis and signal separation:
  - Principal Component Analysis
  - Parallel Factor Analysis
- Block Term Decompositions and Block Component Analysis

Factor analysis and signal separation

- Decompose a data matrix in rank-1 terms
  E.g. independent component analysis, telecommunications, biomedical applications, chemometrics, data analysis, . . .

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



- Decomposition in rank-1 terms is not unique

$$\begin{aligned} \mathbf{A} &= (\mathbf{FM}) \cdot (\mathbf{M}^{-1}\mathbf{G}^T) \\ &= \tilde{\mathbf{F}} \cdot \tilde{\mathbf{G}}^T \end{aligned}$$

Principal Component Analysis

Exploitation of prior knowledge

PCA, SVD: uniqueness obtained by adding orthogonality constraints

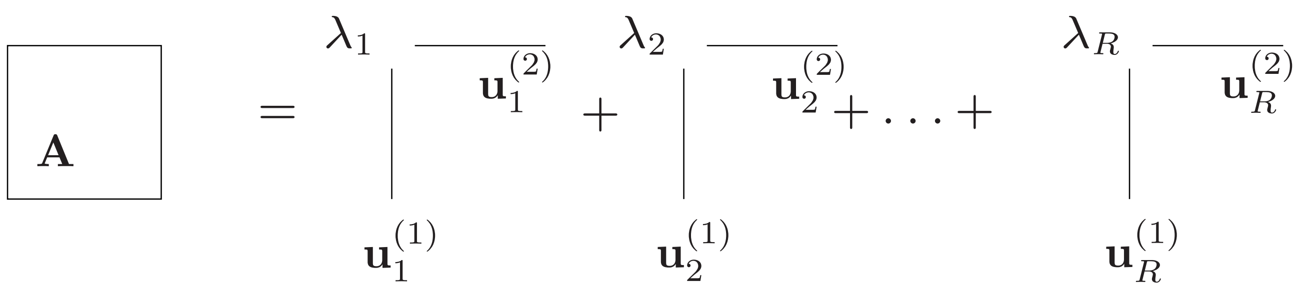
$$\mathbf{A} = \mathbf{U}^{(1)} \cdot \mathbf{\Sigma} \cdot \mathbf{U}^{(2)^T}$$

$\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$ orthogonal, $\mathbf{\Sigma}$ diagonal

Example: excitation-emission fluorescence in chemometrics

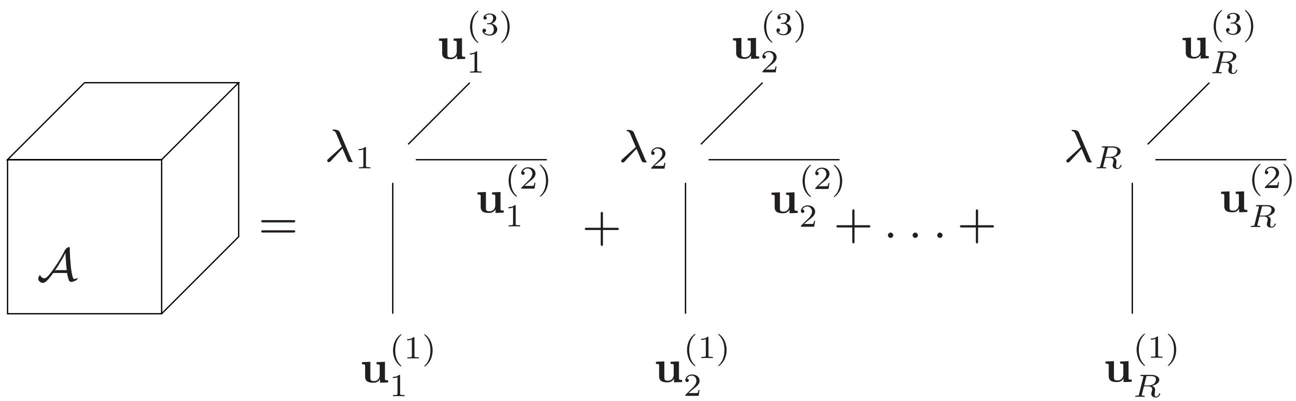
Matrix approach

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



Tensor solution: Parallel Factor Analysis

row vector ∼ emission spectrum
column vector ∼ excitation spectrum
coefficients ∼ concentrations



Applications

- Speech and audio
- Image processing

  feature extraction, image reconstruction, video
- Telecommunications

  OFDM, CDMA, . . .
- Biomedical applications

  functional Magnetic Resonance Imaging, electromyogram, electro-encephalogram, (fetal) electrocardiogram, mammography, pulse oximetry, (fetal) magnetocardiogram, . . .
- Other applications

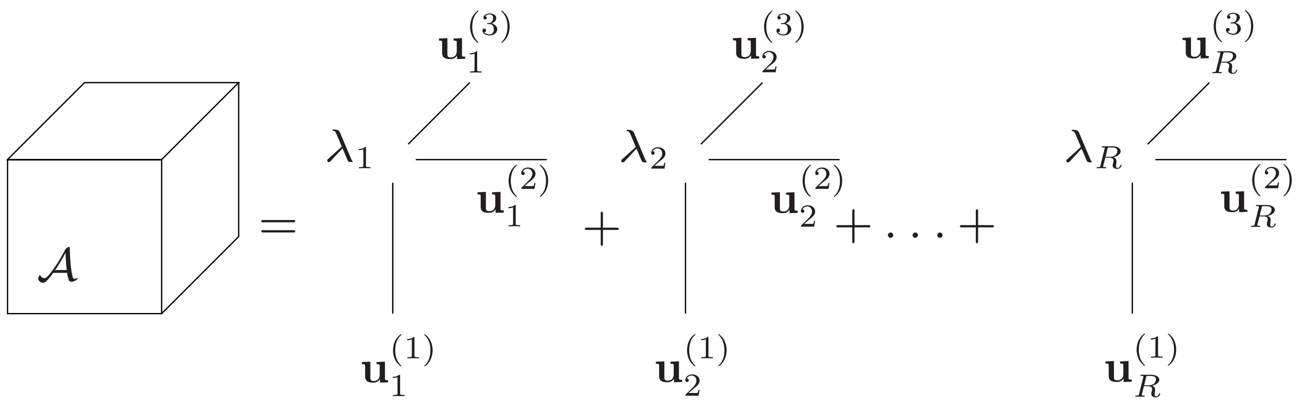
  text classification, vibratory signals generated by termites (!), electron energy loss spectra, astrophysics, . . .

Overview

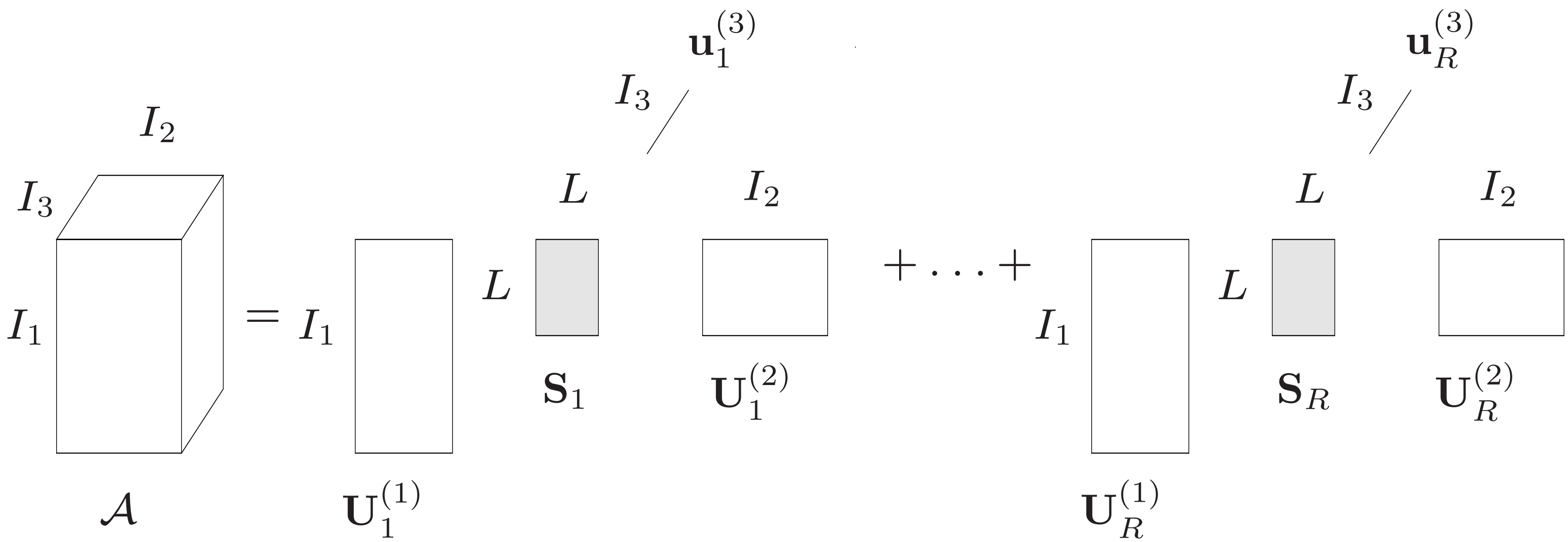
- Preliminaries
- Tensor decompositions
- Factor analysis and signal separation
- Block Term Decompositions and Block Component Analysis

CANDECOMP/PARAFAC

Canonical Decomposition / Parallel Factor Decomposition of a tensor $\mathcal{A}$ is its decomposition in a minimal sum of rank-1 tensors



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



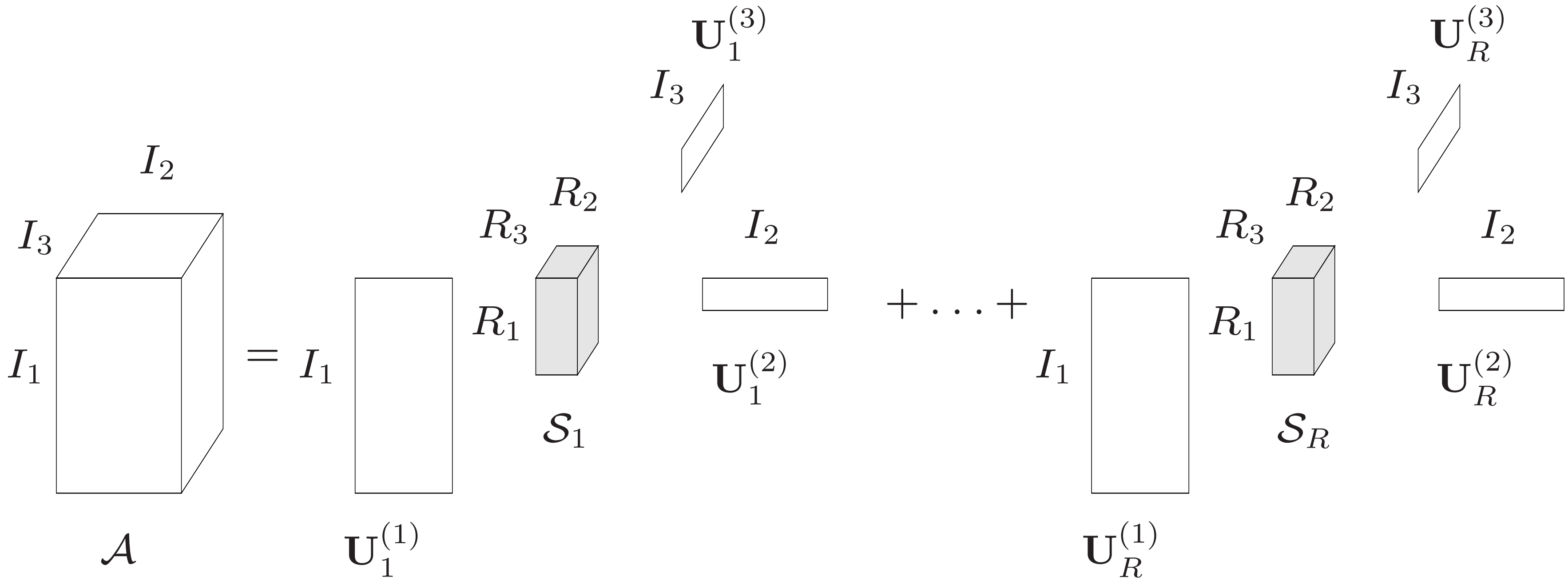
Uniqueness

$$\min\left(\left\lfloor \frac{I_1}{L} \right\rfloor, R\right) + \min\left(\left\lfloor \frac{I_2}{L} \right\rfloor, R\right) + \min(I_3, R) \geqslant 2R + 2$$

cf. $\min(I_1, R) + \min(I_2, R) + \min(I_3, R) \geqslant 2R + 2$ (PARAFAC)

[*De Lathauwer '06*]

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



Unifies: "mode-n rank" and "rank"
Tucker and PARAFAC

[*De Lathauwer '06*]

Block Component Analysis, a new concept for signal separation

- rank-1 structure is very restrictive
- a decomposition in terms of low multilinear rank is under certain conditions essentially unique
- analysis of tensor data by means of BTD: Block Component Analysis
- low multilinear rank is often a very natural structure
- low multilinear rank means that the signal has a relatively small “intrinsic dimension”

Example: unmixing of rational functions (1)

Let the rational functions $r_i(z)$ having degree δ_i, $i = 1, \ldots, q$.
$(deg(r) = \max\{\deg(numerator), \deg(denominator)\})$
Let r be the column vector of these rational functions $r = [r_1, r_2, \ldots, r_q]$.
Let W be a $p \times q$ mixture matrix with $p \geq q$.
Suppose we have only access through these rational functions r via the mixture matrix W

$$m(z_i) = W r(z_i), \qquad i = 1, \ldots, N$$

for N different points z_i in the complex plane.
Problem: recover the rational functions r_i

Example: unmixing of rational functions (2)

Connection between the degree of a rational function and the rank of the corresponding Loewner matrix

Let $r(z)$ be a rational function of degree δ.

Take N points z_i in the complex plane and split this set into two subsets: $X = \{x_i = z_i, i = 1, \ldots, \alpha\}$ and $Y = \{y_{i-\alpha} = z_i, i = \alpha + 1, \ldots, N\}$ with $\alpha, N - \alpha \geq \delta$.

Corresponding Loewner matrix

$$L(r) = \left[\frac{r(x_i) - r(y_j)}{x_i - y_j}\right]_{i=1,\ldots,\alpha;\ j=1,\ldots,N-\alpha}$$

has rank δ.

Moreover, if $N - \alpha = \delta + 1$ and $Lc = 0$, the denominator polynomial $v(z)$ can be written as

$$v(z) = \sum_{j=1,\ldots,\delta+1} c_j \prod_{i \neq j} (z - y_i).$$

[*Antoulas and Anderson '86*]

Example: unmixing of rational functions (3)

Solution of the unmixing problem: use block term decomposition of a corresponding "Loewner tensor"
Loewner tensor $\mathcal{L}$ is defined as

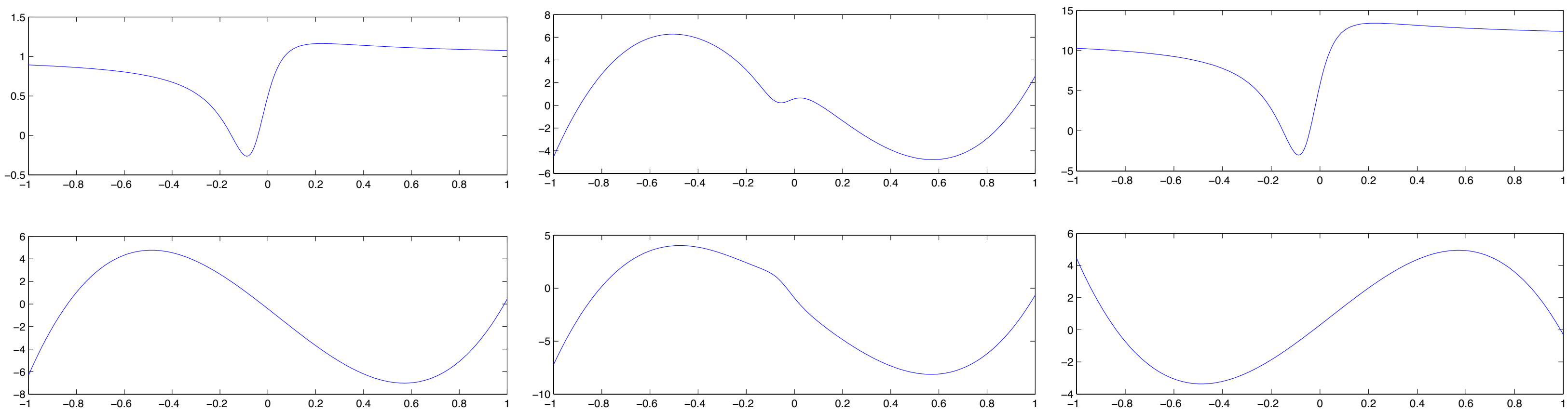
$$\mathcal{L}_{i,j,k} = \left[\frac{m_k(x_i) - m_k(y_j)}{x_i - y_j}\right]$$

Decomposition of $\mathcal{L}$ in rank-(δ,δ,1) terms gives the Loewner matrices of the different rational functions r_i.

Example: unmixing of rational functions (4)

Numerical experiment

Mixing of 2 rational functions of degree 2 and 3 respectively



Conclusion

Tensor decompositions:

- Tucker decomposition / multilinear Singular Value Decomposition
- Parallel factor decomposition
- Block term decompositions

Factor analysis and signal separation:

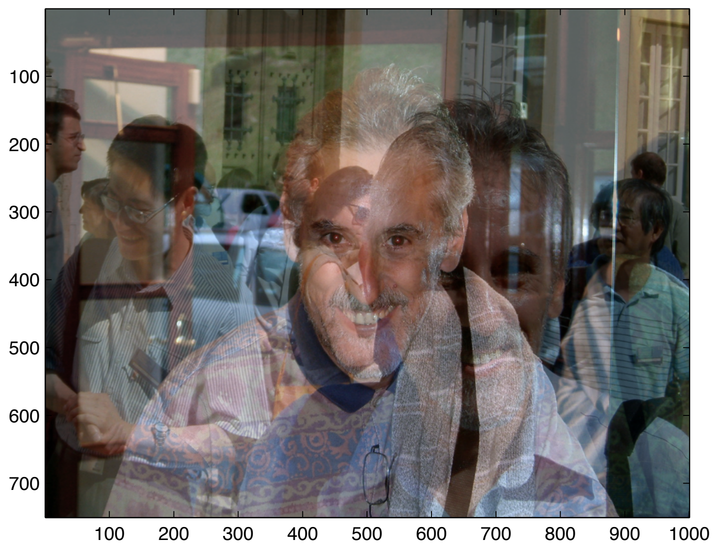
- Principal Component Analysis
- Parallel Factor Analysis
- Block Component Analysis

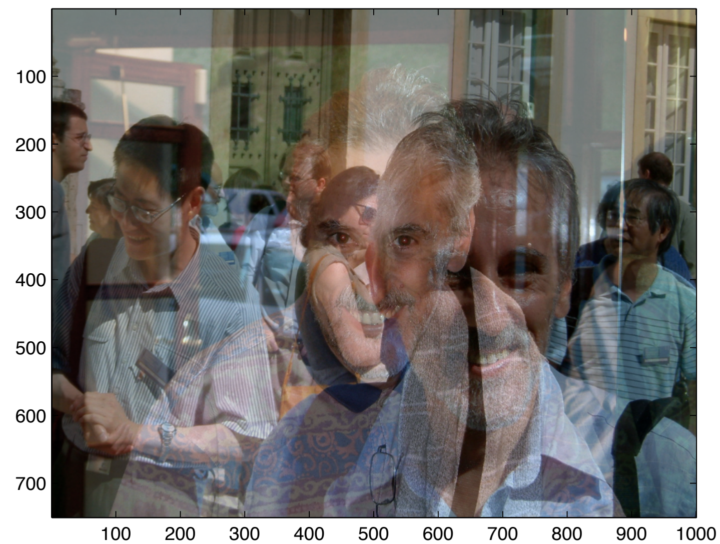
Signal separation on the basis of low intrinsic dimensionality

Example: unmixing of images (1)

Numerical experiment

Mixed images

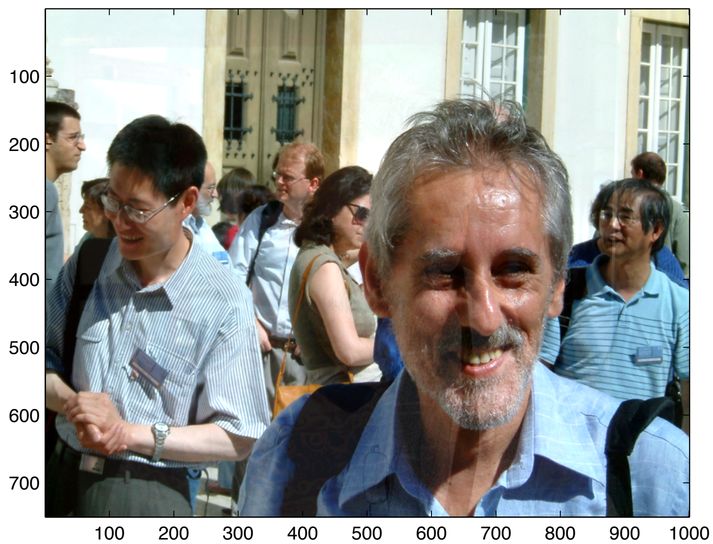


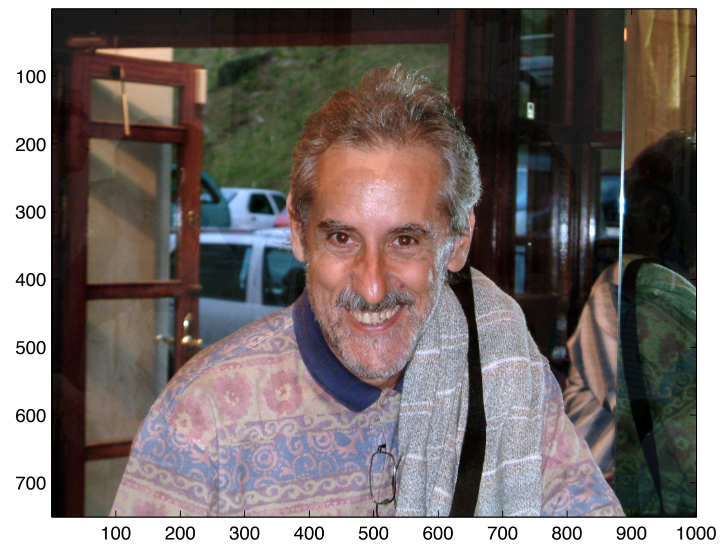


Example: unmixing of images (2)

Numerical experiment

Unmixed images





Example: unmixing of images (3)

Numerical experiment

Original images