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
Mode-$n$ vectors of a tensor: generalization of column/row vectors of a matrix
The column (row) rank of a matrix $A$ is equal to the maximal number of columns (rows) of $A$ 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**: $\text{rank}_1(A) = R_1$, $\text{rank}_2(A) = R_2$, $\text{rank}_3(A) = R_3$.

- **Multilinear rank**: $(R_1, R_2, R_3)$

```
R_1 = R_2 = 2   R_3 = 1
```
**Rank-1 tensor**

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

  $$a_{i_1 i_2} = u^{(1)}_{i_1} u^{(2)}_{i_2}$$

  $A = u^{(1)} \cdot u^{(2)^T} \equiv u^{(1)} \circ u^{(2)}$

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

  $$a_{i_1 i_2 \ldots i_N} = u^{(1)}_{i_1} u^{(2)}_{i_2} \ldots u^{(N)}_{i_N}$$

  $A = u^{(1)} \circ u^{(2)} \circ \ldots \circ u^{(N)}$

![Diagram of rank-1 matrix and tensor]
Rank of a tensor

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

$$A = \lambda_1 \begin{bmatrix} u_1^{(2)} \\ u_1^{(1)} \end{bmatrix} + \lambda_2 \begin{bmatrix} u_2^{(2)} \\ u_2^{(1)} \end{bmatrix} + \ldots + \lambda_R \begin{bmatrix} u_R^{(2)} \\ u_R^{(1)} \end{bmatrix}$$

- The rank $R$ of an $N$th-order tensor $A$ is the minimal number of rank-1 tensors that yield $A$ in a linear combination.

$$A = \lambda_1 \begin{bmatrix} u_1^{(3)} \\ u_1^{(2)} \\ u_1^{(1)} \end{bmatrix} + \lambda_2 \begin{bmatrix} u_2^{(3)} \\ u_2^{(2)} \\ u_2^{(1)} \end{bmatrix} + \ldots + \lambda_R \begin{bmatrix} u_R^{(3)} \\ u_R^{(2)} \\ u_R^{(1)} \end{bmatrix}$$

[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:**

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

in which \( S \) is all-orthogonal and ordered

\( U^{(1)}, U^{(2)}, \ldots, U^{(N)} \) are orthogonal

[\textit{Tucker '64}, \textit{De Lathauwer '00}]
Computation

\[ \mathbf{A} = S \cdot_1 \mathbf{U}^{(1)} \cdot_2 \mathbf{U}^{(2)} \cdot_3 \mathbf{U}^{(3)} \]

- \((I_1 \times I_2 I_3)\) 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 \(\mathbf{S}\):
  \[ \mathbf{S} = \mathbf{A} \cdot_1 \mathbf{U}^{(1)^T} \cdot_2 \mathbf{U}^{(2)^T} \cdot_3 \mathbf{U}^{(3)^T} \]
All-orthogonality:

All-orthogonality is a generalization of diagonality

Ordering: slices have decreasing Frobenius norm

Norms of slices = mode-\(n\) singular values

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

$$\mathcal{A} = \lambda_1 u_1^{(1)} + \lambda_2 u_2^{(1)} + \ldots + \lambda_R u_R^{(1)}$$

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

Unique under mild conditions
Algorithms

\[ f(U^{(1)}, U^{(2)}, U^{(3)}) = \| A - \sum_{r=1}^{R} u_r^{(1)} \circ u_r^{(2)} \circ u_r^{(3)} \|^2 \]

- Alternating least squares (ALS) \cite{Harshman:1970}
- ALS with Exact Line Search \cite{Rajih:2008, Nion:2008}
- ALS with regularization \cite{Navasca:2008}
- general-purpose optimization:
  - Levenberg-Marquardt
  - conjugate gradient \cite{Paatero:1999, Acar:2009}
  - ...
- EVD \cite{Leurgans:1993}, ...
- simultaneous generalized Schur \cite{DeLathauwer:2004}
- simultaneous matrix diagonalization \cite{DeLathauwer:2006}
- ...

...
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, ...

\[
A = F \cdot G^T
\]

\[
\begin{bmatrix}
A \\
\end{bmatrix} = \begin{bmatrix}
g_1 \\
g_2 \\
\vdots \\
g_R \\
\end{bmatrix} + \begin{bmatrix}
f_1 \\
f_2 \\
\vdots \\
f_R \\
\end{bmatrix} + \ldots + \begin{bmatrix}
\end{bmatrix}
\]

- Decomposition in rank-1 terms is not unique

\[
A = (FM) \cdot (M^{-1}G^T) \\
= \tilde{F} \cdot \tilde{G}^T
\]
Exploitation of prior knowledge

PCA, SVD: uniqueness obtained by adding orthogonality constraints

\[ A = U^{(1)} \cdot \Sigma \cdot U^{(2)T} \]

\( U^{(1)} \), \( U^{(2)} \) orthogonal, \( \Sigma \) diagonal
Example: excitation-emission fluorescence in chemometrics

Matrix approach

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

$$A = \lambda_1 \begin{pmatrix} u_1^{(1)} \\ u_1^{(2)} \end{pmatrix} + \lambda_2 \begin{pmatrix} u_2^{(1)} \\ u_2^{(2)} \end{pmatrix} + \ldots + \lambda_R \begin{pmatrix} u_R^{(1)} \\ u_R^{(2)} \end{pmatrix}$$
Tensor solution: Parallel Factor Analysis

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

\[
\mathcal{A} = \lambda_1 \frac{\mathbf{u}_1^{(3)}}{\mathbf{u}_1^{(2)}} + \lambda_2 \frac{\mathbf{u}_2^{(3)}}{\mathbf{u}_2^{(2)}} + \ldots + \lambda_R \frac{\mathbf{u}_R^{(3)}}{\mathbf{u}_R^{(2)}}
\]
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
Canonical Decomposition / Parallel Factor Decomposition of a tensor $\mathcal{A}$ is its decomposition in a minimal sum of rank-1 tensors:

$$
\mathcal{A} = \lambda_1 u_1^{(1)} u_1^{(2)} + \lambda_2 u_2^{(1)} u_2^{(2)} + \ldots + \lambda_R u_R^{(1)} u_R^{(2)}
$$
Decomposition in rank-\((L, L, 1)\) terms

\[
A = I_1 \begin{bmatrix} L & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \end{bmatrix} \begin{bmatrix} S_1 & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \end{bmatrix} \begin{bmatrix} U_1^{(1)} & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \end{bmatrix} + \ldots + I_1 \begin{bmatrix} L & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \end{bmatrix} \begin{bmatrix} S_R & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \end{bmatrix} \begin{bmatrix} U_R^{(1)} & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \end{bmatrix}
\]

Uniqueness

\[
\min\left(\begin{bmatrix} I_1 \\ L \end{bmatrix}, R\right) + \min\left(\begin{bmatrix} I_2 \\ L \end{bmatrix}, R\right) + \min(I_3, R) \geq 2R + 2
\]

\[
\text{cf. } \min(I_1, R) + \min(I_2, R) + \min(I_3, R) \geq 2R + 2 \quad \text{(PARAFAC)}
\]

[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 $\delta_i$, $i = 1, \ldots, q$.

\[(\text{deg}(r) = \max\{\text{deg(numerator)}, \text{deg(denominator)}\})\]

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 $W$

\[m(z_i) = W r(z_i), \quad i = 1, \ldots, N\]

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 \( \delta \).

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

\[
L(r) = \begin{bmatrix}
    r(x_i) - r(y_j) \\
    x_i - y_j
\end{bmatrix}_{i=1,\ldots,\alpha; j=1,\ldots,N-\alpha}
\]

has rank \( \delta \).

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-$(\delta, \delta, 1)$ 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