Underdetermined MIMO Channel Estimation Using Nonnegative Tensor Factorization Algorithms

Carlos Estêvão Rolim Fernandes

Wireless Research Telecommunications Group – GTEL
Federal University of Ceará / Center of Technology
http://www.gtel.ufc.br

June 5th, 2012
Presentation Outline

1. Multilinear Data Analysis
2. Nonnegative Tensor Factorization
3. Current Research
4. Simulation results
5. Future Work
Multilinear Data Analysis
WHY TENSORS?

1. Tensor factorizations are unique under mild conditions
2. Tensor rank is not bounded by its dimensions (more inputs than outputs)
3. Lack of orthogonality constraints (no prewhitening needed)
4. Fully exploit the multidimensional nature of data
Multilinear Data Analysis

Tensor Decompositions in Data Analysis

WHY TENSORS?

1. Tensor factorizations are unique under mild conditions
2. Tensor rank is not bounded by its dimensions (more inputs than outputs)
3. Lack of orthogonality constraints (no prewhitening needed)
4. Fully exploit the multidimensional nature of data

Tensor Decomposition Models

CP Model

\[
\mathbf{Y} = \sum_{i=1}^{p} \mathbf{a}_i \mathbf{b}_i \mathbf{c}_i + \mathbf{E}
\]

\((I \times T \times Q)\)
Tensor Decompositions in Data Analysis

Solving the 3rd-order CP Model

Scalar form: \( t_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} \)

Tensor Rank is \( F \)

\[
\begin{bmatrix}
T \\
I \\
J \\
K
\end{bmatrix} = \begin{bmatrix}
c_1 \\
b_1 \\
a_1
\end{bmatrix} + \ldots + \begin{bmatrix}
c_F \\
b_j \\
a_F
\end{bmatrix}
\]
Tensor Decompositions in Data Analysis

Solving the 3rd-order CP Model

Scalar form: \[ t_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} \]

Tensor Rank is \( F \)

Cost Functions:
\[
J(A, B, C) = \| T[1] - (B \odot C) A^T \|_2^2 = \| T[2] - (C \odot A) B^T \|_2^2 = \| T[3] - (A \odot B) C^T \|_2^2
\]

Factors \( A, B, \) and \( C \) are unique under Kruskal conditions:
\[ k_A + k_B + k_C \geq 2F + 2 \]

Unfolded representations:
\[
\begin{align*}
T[1] &= B D_i(A) C^T \\
T[2] &= C D_j(B) A^T \\
T[3] &= A D_k(C) B^T
\end{align*}
\]
Tensor Decompositions in Data Analysis

Solving the 3rd-order CP Model

Scalar form: \[ t_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} \]

Tensor Rank is \( F \)

Unfolded representations

Cost Functions:
\[
J(A, B, C) = \| T^{[1]} - (B \mathbin{\hat{\otimes}} C) A^T \|_F
\]
\[
= \| T^{[2]} - (C \mathbin{\hat{\otimes}} A) B^T \|_F
\]
\[
= \| T^{[3]} - (A \mathbin{\hat{\otimes}} B) C^T \|_F
\]

Factors \( A \), \( B \) and \( C \) are unique under Kruskal conditions:
\[
k_A + k_B + k_C \geq 2F + 2
\]
Tensor Decompositions in Data Analysis

Solving the 3rd-order CP Model

Scalar form: \( t_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} \)

Tensor Rank is \( F \)

\[ T = \begin{bmatrix} c_1 \\ b_1 \\ a_1 \\ \vdots \\ c_F \\ b_F \\ a_F \end{bmatrix} \]

Unfolded representations

Cost Functions:

\[ J(A, B, C) = \| T_{[1]} - (B \odot C)A^T \|_F^2 = \| T_{[2]} - (C \odot A)B^T \|_F^2 = \| T_{[3]} - (A \odot B)C^T \|_F^2 \]
Tensor Decompositions in Data Analysis

Solving the 3rd-order CP Model

Scalar form: \[ t_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} \]

Tensor Rank is \( F \)

Cost Functions:
\[ J(A, B, C) = \| T_{[1]} - (B \odot C)A^T \|_F^2 = \| T_{[2]} - (C \odot A)B^T \|_F^2 = \| T_{[3]} - (A \odot B)C^T \|_F^2 \]

Factors \( A, B \) and \( C \) are unique under Kruskal conditions: \( k_A + k_B + k_C \geq 2F + 2 \)
Nonnegative Tensor Factorization
Nonnegativity and Tensor Decompositions

WHY NOT TENSORS?

1. **Complexity**: Numerically difficult to solve
2. **Degeneracy**: convergence to global minimum is not guaranteed
Nonnegativity and Tensor Decompositions

WHY NOT TENSORS?

1. Complexity: Numerically difficult to solve
2. Degeneracy: convergence to global minimum is not guaranteed

Nonnegativity in Signal Processing

- Nonnegative quantities in nature and engineering applications
- Part-based representation of data that is compositional in nature
- Nonnegativity may be the key for information extraction
Nonnegativity and Tensor Decompositions

WHY NOT TENSORS?

1. Complexity: Numerically difficult to solve
2. Degeneracy: convergence to global minimum is not guaranteed

Nonnegativity in Signal Processing

- Nonnegative quantities in nature and engineering applications
- Part-based representation of data that is compositional in nature
- Nonnegativity may be the key for information extraction
- Consequences of exploiting nonnegativity:
  
  The low rank approximation degeneracy of the CP model will not happen!
WHY NOT TENSORS?

1. Complexity: Numerically difficult to solve
2. Degeneracy: convergence to global minimum is not guaranteed

Nonnegativity in Signal Processing

- Nonnegative quantities in nature and engineering applications
- Part-based representation of data that is compositional in nature
- Nonnegativity may be the key for information extraction
- Consequences of exploiting nonnegativity:
  The low rank approximation degeneracy of the CP model will not happen!

What else to expect?

Faster and simpler algorithms for physical implementation in complex and possibly distributed systems
Nonnegative Tensor Factorization

Algorithms for Nonnegative Component Estimation

Efficient Algorithms to Compute the Nonnegative CP Model Approximation Factors

1. Nonnegative Least Squares (NNLS) methods [Paatero:1997]:
   - Active Set strategy
   - Weighted Cost Function with Penalty terms

2. Loading matrices parametrization methods [Lim.Comon:2009]: Nonnegativity is imposed to loading matrices through parametrization. Cost function is unchanged:

   \[ J(\tilde{A}, \tilde{B}, \tilde{C}) = \| T[1] - [(\tilde{B} \odot \tilde{B}) \odot (\tilde{C} \odot \tilde{C})](\tilde{A} \odot \tilde{A})^T \|^2_F \]

   where \( A = \tilde{A} \odot \tilde{A}, B = \tilde{B} \odot \tilde{B}, C = \tilde{C} \odot \tilde{C} \) and \( \odot \) denotes the Hadamard product

   - Gradient and Quasi-Gradient
   - Gauss-Newton
   - Levenberg-Marquadt
Underdetermined MIMO Channel Estimation
NMF-based algorithms for channel estimation

**Underdetermined MIMO Scenario**

- User power loading concept
- $M \times Q$ MU-SIMO channel model
- Quasi-static propagation (flat fading within a data block)
- Data blocks of $N$ symbols split in $P$ substreams

**Proposed Approach:** Tensor of Covariance Matrices
Current Research

NTF-based algorithms for channel estimation

Underdetermined MIMO Scenario

- User power loading concept
- \( M \times Q \) MU-SIMO channel model
- Quasi-static propagation (flat fading within a data block)
- Data blocks of \( N \) symbols split in \( P \) substreams

Proposed Approach: Tensor of Covariance Matrices
NTF-based algorithms for channel estimation

Underdetermined MIMO Scenario

- User power loading concept
- $M \times Q$ MU-SIMO channel model
- Quasi-static propagation (flat fading within a data block)
- Data blocks of $N$ symbols split in $P$ substreams

Proposed Approach: Tensor of Covariance Matrices
NTF-based algorithms for channel estimation

Underdetermined MIMO Scenario

- User power loading concept
- $M \times Q$ MU-SIMO channel model
- Quasi-static propagation (flat fading within a data block)
- Data blocks of $N$ symbols split in $P$ substreams

Proposed Approach: Tensor of Covariance Matrices

$Y^{(1)} \in \mathbb{C}^{M \times \frac{N}{P}}$

$R^{(p)} = E\{Y^{(p)}Y^{(p)\dagger}\} = H D_p(\tilde{\Lambda}) H^H$
NTF-based algorithms for channel estimation

Underdetermined MIMO Scenario

- User power loading concept
- $M \times Q$ MU-SIMO channel model
- Quasi-static propagation (flat fading within a data block)
- Data blocks of $N$ symbols split in $P$ substreams

Proposed Approach: Tensor of Covariance Matrices

$$\mathbf{Y}^{(1)} \in \mathbb{C}^{M \times \frac{N}{P}}$$

$$\mathbf{Y}^{(p)} \in \mathbb{C}^{M \times \frac{N}{P}}$$
NTF-based algorithms for channel estimation

Underdetermined MIMO Scenario

- User power loading concept
- $M \times Q$ MU-SIMO channel model
- Quasi-static propagation (flat fading within a data block)
- Data blocks of $N$ symbols split in $P$ substreams

Proposed Approach: Tensor of Covariance Matrices

$$R^{(1)} = E \{ Y^{(1)} Y^{(1)^*} \}$$

$$R^{(p)} = E \{ Y^{(p)} Y^{(p)^*} \} = H D_p (\Lambda) H^H$$
Simulation results

Ongoing Investigation and Results

A few tested scenarios satisfying identifiability conditions

- **K**: Length of symbol burst (3000)
- **P**: Number of substreams (4, 8)
- **L**: Number of Mote Carlo simulations (300)
- **Q**: Number of users (transmitters)
- **M**: Number of receive antennas

<table>
<thead>
<tr>
<th>Q</th>
<th>P</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>M ≥ 3</td>
<td>M ≥ 3</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>M ≥ 4</td>
<td>M ≥ 4</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>M ≥ 5</td>
<td>M ≥ 4</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>M ≥ 7</td>
<td>M ≥ 5</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>M ≥ 9</td>
<td>M ≥ 7</td>
<td></td>
</tr>
</tbody>
</table>

- Identification performance criterion: Mean Squared Error (MSE)

\[
\epsilon = \frac{1}{L} \sum_{\ell=1}^{L} \| \hat{H}^{(\ell)} - H \|^2_F.
\]
Simulation Results

**MSE Performance:** Scenario with $M=3$ receive antennas, $P=8$ substreams

$q = 4$ users

$q = 5$ users
**Simulation Results**

**MSE Performance:** Scenario with $M = 3$ receive antennas, $P = 8$ substreams

- $Q = 4$ users
- $Q = 5$ users

**Highlight observations**

- Faster convergence and lower computational burden
- Improved channel estimation under different environments
- Robustness to challenging scenarios (more users than receive antennas)
Future Work

Perspectives

In the pipeline...

- Loading matrices parametrization approach considering channel structure
- Algorithm efficiency and user synchronization issues: Further computer simulations
- Evaluate impact of initialization on the proposed techniques
- Estimate the number of users (tensor rank)
- Use Fourth- (or higher) order statistics (and higher-order tensors)
- Compute optimal user transmit powers maximizing a given metric (e.g. SINR)
- Use the LTE 3GPP 4G channel model, and perform power allocation in the time-frequency grid.
Perspectives

In the pipeline...

- **Loading matrices parametrization approach** considering channel structure
- Algorithm efficiency and user synchronization issues: Further computer simulations
- Evaluate impact of initialization on the proposed techniques
- Estimate the number of users (tensor rank)
- Use Fourth- (or higher) order statistics (and higher-order tensors)
- **Compute optimal user transmit powers** maximizing a given metric (e.g. SINR)
- Use the LTE 3GPP 4G channel model, and perform power allocation in the time-frequency grid.
Future Work

**Other Perspectives**

A deterministic approach?

![Diagram](image)

- Receive antenna (m)
- Substream (p)
- Discrete time (n)
- \( Y^{(1)} \)
- \( Y^{(2)} \)
- \( \vdots \)
- \( Y^{(p)} \)

**Further Applications and Challenges**

- Blind source separation (biomedical, seismic, ...)
- Radio resource allocation (power allocation and control, link scheduling)
- Wireless sensor networks
- Localization problems
- Blind channel estimation and information retrieval

C. Estêvão R. Fernandes (GTEL-UFC)
Other Perspectives

A deterministic approach?

Future Work
A deterministic approach?

Receive antenna \((m)\)

\[
\begin{align*}
Y^{(1)} & \\
Y^{(2)} & \\
\vdots & \\
Y^{(P)} & \\
\end{align*}
\]

Discrete time \((n)\)

Substream \((p)\)

\[
\begin{align*}
h_1 & \\
s_1 & \\
\vdots & \\
h_Q & \\
\end{align*}
\]

\[
\begin{align*}
\lambda_1 & \\
\lambda_Q & \\
\end{align*} + \cdots + \Psi \quad \text{(noise)}
\]

Further Applications and Challenges

- Blind source separation (biomedical, seismic, ...)
- Radio resource allocation (power allocation and control, link scheduling)
- Wireless sensor networks
- Localization problems
- Blind channel estimation and information retrieval

C. Estêvão R. Fernandes (GTEL-UFC)
Other Perspectives

A deterministic approach?

Further Applications and Challenges

- Blind source separation (biomedical, seismic, ...)
- Radio resource allocation (power allocation and control, link scheduling)
- Wireless sensor networks
- Localization problems
- Blind channel estimation and information retrieval
Thank You


