Identifiability of tensor rank decompositions

Nick Vannieuwenhoven

KU Leuven / FWO

June 16, 2015

In collaboration with:

- L. Chiantini (Università di Siena)
- G. Ottaviani (Università di Firenze)
Overview

1. An application
2. Identifiability
3. Computations
4. References
Overview

1. An application
2. Identifiability
3. Computations
4. References

Identifiability of tensor rank decompositions

Nick Vannieuwenhoven
Blind source separation is symmetric tensor decomposition

Given some measurements \( Y \in \mathbb{R}^{p \times N} \) in

\[
Y = MX,
\]

find \( M \in \mathbb{R}^{p \times p} \) and \( X \in \mathbb{R}^{p \times N} \).

You cannot solve this uniquely: there is a manifold of solutions!
Blind source separation is symmetric tensor decomposition

Given some measurements $Y \in \mathbb{R}^{p \times N}$ in

$$Y = MX,$$

find $M \in \mathbb{R}^{p \times p}$ and $X \in \mathbb{R}^{p \times N}$.

You cannot solve this uniquely: there is a manifold of solutions!

Now assume the following:

- The rows of $X$ correspond with a random vector $\mathbf{x}$.
- The rows of $Y$ correspond with a random vector $\mathbf{y} = M\mathbf{x}$.
- Every column of $X$ represents a sample of $\mathbf{x}$.
- Every column of $Y$ represents the corresponding sample of $\mathbf{y}$.
- The random variates in $\mathbf{x}$ are (nearly) statistically independent.
The random variates in $\mathbf{x}$ are statistically independent.

This is equivalent with a diagonal $n$th order statistical moment tensor:

$$C^n_{\mathbf{x}} = \sum_{i=1}^{p} c_i \mathbf{e}_i \otimes \cdots \otimes \mathbf{e}_i$$

for every $n$. (Think about the correlation matrix.)
The random variates in $\mathbf{x}$ are statistically independent.

This is equivalent with a diagonal $n$th order statistical moment tensor:

$$C^n_x = \sum_{i=1}^{p} c_i \, \mathbf{e}_i \otimes \cdots \otimes \mathbf{e}_i$$

for every $n$. (Think about the correlation matrix.)

If $\mathbf{y}$ depends linearly on $\mathbf{x}$ through the mixing matrix $\mathbf{M}$, then:

$$C^4_y = \sum_{i=1}^{p} c_i(M\mathbf{e}_i) \otimes (M\mathbf{e}_i) \otimes (M\mathbf{e}_i) \otimes (M\mathbf{e}_i)$$

This is a tensor rank decomposition, which is generally unique.
The cyclically (and seriously) broken camera problem
The cyclically (and seriously) broken camera problem

Recovered

Original

Nick Vannieuwenhoven
Identifiability of tensor rank decompositions
The cyclically (and seriously) broken camera problem

Recovered

Original

Nick Vannieuwenhoven
Identifiability of tensor rank decompositions
Overview

1. An application
2. Identifiability
3. Computations
4. References
Tensor rank decomposition

Every element of $\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d} \cong \mathbb{C}^{n_1 \cdots n_d}$ can be written as

$$A = \sum_{i=1}^{r} a^1_i \otimes a^2_i \otimes \cdots \otimes a^d_i,$$

i.e., as a linear combination of rank-1 tensors.

We assume that $r$ is minimal.
The set of tensors of rank 1 in $\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d}$ forms an algebraic variety called the Segre variety $S$.

For $p_i \in S$ with

$$p_i = a_1^i \otimes a_2^i \otimes \cdots \otimes a_d^i \in \mathbb{C}^{n_1 n_2 \cdots n_d},$$

the tangent space is given by the column span of

$$T_i = T_{p_i}S = \begin{bmatrix} l_{n_1} \otimes a_2^i \otimes \cdots \otimes a_d^i & \cdots & a_1^i \otimes \cdots \otimes a_{d-1}^i \otimes l_{n_d} \end{bmatrix}$$
Let $S$ be a nondefective Segre variety that is not $r$-identifiable. For $r$ generic points $p_1, p_2, \ldots, p_r \in S$, the set of points $p \in S$ obeying

$$T_p S \subset H := \langle T_{p_1} S, T_{p_2} S, \ldots, T_{p_r} S \rangle$$

contains at least an algebraic curve $C_r$ passing through $p_1, p_2, \ldots, p_r$.
Consider the tangentially $r$-contact variety:

$$C_r = \{ p \in S \mid T_p S \subset H = \langle T_{p_1} S, T_{p_2} S, \ldots, T_{p_r} S \rangle \}$$

for some general fixed $p_i \in S$.

$C_r$ is zero-dimensional $\rightarrow$ the Segre variety $S$ is $r$-identifiable.
Sample $r$ random points on the Segre variety. That is, choose

$$p_i = a_i^1 \otimes a_i^2 \otimes \cdots \otimes a_i^d.$$  

Then,

$$H = \begin{bmatrix} T_1 & T_2 & \cdots & T_r \end{bmatrix}$$

with $T_i$ the tangent space matrix as before.
Let

\[ \Pi := \prod_{k=1}^{d} n_k, \quad \text{i.e., the dimension of the embedding space} \]

\[ \Sigma := \sum_{k=1}^{d} (n_k - 1), \quad \text{i.e., the dimension of } S \]

\[ \ell := \Pi - r\Sigma, \quad \text{i.e., the expected codimension of } \sigma_r(S). \]
Let $K = [k_i]$ be a matrix representation of $(\text{Im } T)^\perp$.

If $\text{rank}(K) \neq \ell$, then $S$ is (likely) $r$-defective.

We can consider $\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d} \cong \mathbb{C}^{n_1 \times \cdots \times n_d}$, so

$$q_l(x_1, x_2, \ldots, x_\Pi) = \sum_{i=1}^{\Pi} k_i, l x_i$$

$$= \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_d=1}^{n_d} k(i_1, i_2, \ldots, i_d), l x(i_1, i_2, \ldots, i_d) = 0$$

are the $\ell$ Cartesian equations of $H$. 
Next, plug in the equations for the Segre variety:

\[ x(i_1, i_2, \ldots, i_d) = a^{1}_{i_1} a^{2}_{i_2} \cdots a^{d}_{i_d}; \]

then, we can write

\[
q_l(a^{1}, a^{2}, \ldots, a^{d}) = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_d=1}^{n_d} k(i_1, i_2, \ldots, i_d), l a^{1}_{i_1} a^{2}_{i_2} \cdots a^{d}_{i_d} = 0.
\]

This gives Cartesian equations for \( H \cap S \).
An algorithm proving generic identifiability: Step 3

Deriving with respect to the parameterization of the Segre variety yields

$$
\begin{bmatrix}
\frac{\partial}{\partial a} q_1(a) \\
\frac{\partial}{\partial a} q_2(a) \\
\vdots \\
\frac{\partial}{\partial a} q_\ell(a)
\end{bmatrix} = 0,
$$

which are equations of $C_r = H \cap TS$.

Deriving again, we find the tangent space from which the dimension of $C_r$ can be computed using linear algebra.
In summary, for computing $\dim C_r$, we build the stacked Hessian:

$$H = \begin{bmatrix}
\frac{\partial^2}{\partial a \partial a} q_1(a) & \frac{\partial^2}{\partial a \partial a} q_2(a) & \cdots & \frac{\partial^2}{\partial a \partial a} q_\ell(a)
\end{bmatrix}.$$ 

$$\text{rank}(H) = \Sigma \rightarrow \text{The Segre variety } S \text{ is } r\text{-identifiable.}$$
Let 
\[ r_E = \left\lceil \frac{n_1 \cdots n_d}{n_1 + \cdots + n_d - d + 1} \right\rceil \]
be the expected rank of a generic tensor \( \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d} \).

**Proposition**
If the Segre variety \( S \) is \( r \)-defective, it is also not \( r \)-identifiable.

**Proposition**
The Segre variety is not \( r \)-identifiable if \( r > r_E \).
**Conjecture:** Generic rank-$r$ tensors in $\mathbb{C}^{n_1 \times \cdots \times n_d}$ have a unique rank decomposition if $r \leq r_E - 1$, unless

<table>
<thead>
<tr>
<th>$(n_1, \ldots, n_d)$</th>
<th>type</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(m, n)$</td>
<td>$r \geq 2$ defective [AOP09]</td>
</tr>
<tr>
<td>$(4, 4, 3)$</td>
<td>$r \geq r_E - 1$ defective [AOP09]</td>
</tr>
<tr>
<td>$(4, 4, 4)$</td>
<td>$r \geq r_E - 1$ sporadic [CO12]</td>
</tr>
<tr>
<td>$(6, 6, 3)$</td>
<td>$r \geq r_E - 1$ sporadic [CMO13]</td>
</tr>
<tr>
<td>$(2n + 1, 2n + 1, 3)$</td>
<td>$r \geq r_E - 0$ defective [S83]</td>
</tr>
<tr>
<td>$(n, n, 2, 2)$</td>
<td>$r \geq r_E - 1$ defective [AOP09]</td>
</tr>
<tr>
<td>$(2, 2, 2, 2, 2)$</td>
<td>$r \geq r_E - 1$ sporadic [BC13]</td>
</tr>
<tr>
<td>$n_1 \geq \eta$</td>
<td>$r \geq \eta$ defective [BCO13]</td>
</tr>
</tbody>
</table>

with $\eta = \prod_{k=2}^{d} n_k - \sum_{k=2}^{d} (n_k - 1)$. 
The conjecture is true when $n_1 \cdots n_d \leq 17500!$

The symmetric version of this conjecture was proved [COV15].
Overview

1. An application
2. Identifiability
3. Computations
4. References
Overview

1. An application
2. Identifiability
3. Computations
4. References


Further reading