Rank Decompositions of Tensors

Tamara G. Kolda

Sandia National Labs

Funding support from the Applied Mathematical Sciences Program at the U.S. Department of Energy.

SAND No. 99-8630C

Outline

- Rank Decomposition of a Matrix
- Tensor Terminology
- Motivations for Tensor Decompositions
- Decomposed Tensors — Building Blocks
- Notions of Orthogonality and Rank for Tensors
- Examples
- Can the Eckart-Young Theorem be Extended to Tensors?

Research motivated by Leibovici & Sabatier, *LAA*, 1998.

Decomposing a Matrix

Let A be an $m_1 \times m_2$ matrix.

The *rank* of A is defined to be the minimum r such that A can be written as

$$A = \sum_{i=1}^{r} \sigma_i U_i, \qquad (*)$$

where each σ_i is a positive scalar and each U_i is a rank-1 matrix defined by the outer product

$$U_i \equiv u_i^{(1)} \otimes u_i^{(2)}.$$

Equation $(*)$ is called the *rank decomposition* of A.

Theorem. (Eckart-Young, 1936) The *best rank-k approximation* of A is given by

$$\sum_{i=1}^{k} \sigma_i U_i = \arg\min \|A - A_k\| \text{ s.t. } \operatorname{rank}(A_k) = k.$$

Terminology

Let A be an $m_1 \times m_2 \times \cdots \times m_n$ array.

- A is a *tensor*.
- The *order* of A is n.
- The *jth dimension* of A is m_j.
- If B is the same size as A, then the *inner product* of A and B is defined as

$$A \cdot B \equiv \sum_{i_1=1}^{m_1} \sum_{i_2=1}^{m_2} \cdots \sum_{i_n=1}^{m_n} A(i_1, i_2, \cdots, i_n) B(i_1, i_2, \cdots, i_n).$$

- The *norm* of A is defined as

$$\|A\|^2 \equiv A \cdot A = \sum_{i_1=1}^{m_1} \sum_{i_2=1}^{m_2} \cdots \sum_{i_n=1}^{m_n} A(i_1, i_2, \cdots, i_n)^2.$$

Goals & Motivations for Tensor Decompositions

Our goal is to use tensor decompositions to generate best low-rank approximations.

- **Image Collection Compression** - A series of related images can be compressed simultaneously exploiting commonalities between them yet still allowing for differences (e.g., NMR images).
- **Image Retrieval** - Latent Semantic Indexing for images.
- **Multimode Statistical Analysis & Clustering** - Similar to Principal Components Analysis.
- Etc.

Decomposed Tensors – Building Blocks

A *decomposed tensor* is a tensor that can be written as

$$U = u^{(1)} \otimes u^{(2)} \otimes \cdots \otimes u^{(n)},$$

where $u^{(j)} \in \Re^{m_j}$ for $j = 1, \ldots, n$. In this case,

$$U(i_1, i_2, \ldots, i_n) = u^{(1)}(i_1) \cdot u^{(2)}(i_2) \cdots u^{(n)}(i_n).$$

Given decomposed tensors

$$U = u^{(1)} \otimes u^{(2)} \otimes \cdots \otimes u^{(n)}, \text{ and}$$

$$V = v^{(1)} \otimes v^{(2)} \otimes \cdots \otimes v^{(n)},$$

then

$$U \cdot V = \prod_{j=1}^{n} u^{(j)} \cdot v^{(j)} \quad \text{and} \quad \|U\| = \prod_{j=1}^{n} \|u^{(j)}\|_2.$$

Lemma: The tensor $W = U + V$ is itself a decomposed tensor iff all but at most one of the components of U and V are equal.

Notions of Orthogonality

$$U = u^{(1)} \otimes u^{(2)} \otimes \cdots \otimes u^{(n)}.$$

$$V = v^{(1)} \otimes v^{(2)} \otimes \cdots \otimes v^{(n)}.$$

We say that U and V are *orthogonal* ($U \perp V$) if

$$U \cdot V = \prod_{j=1}^{n} u^{(j)} \cdot v^{(j)} = 0.$$

We say that U and V are *completely orthogonal* if for every $j = 1, 2, \ldots, n$,

$$u^{(j)} \cdot v^{(j)} = 0$$

We say that U and V are *strongly orthogonal* ($U \perp_s V$) if $U \perp V$ and for every $j = 1, 2, \ldots, n$,

$$u^{(j)} \cdot v^{(j)} = 0 \quad \underline{\text{or}} \quad u^{(j)} = v^{(j)},$$

Completely Orthogonal $\Rightarrow$ Strongly Orthogonal $\Rightarrow$ Orthogonal

Rank Decompositions

Our goal is to express a A as a weighted sum of decomposed tensors:

$$A = \sum_{i=1}^{r} \sigma_i U_i \qquad (\dagger)$$

where $\sigma_i > 0$ and $||U_i|| = 1$ for all i.

- The *rank* of A is defined to be the minimum r such that A can be written as ($\dagger$), and the decomposition is called the *rank decomposition*.
- The *orthogonal rank* of A is defined to be the minimum r such that A can be written as ($\dagger$) **and** $U_i \perp U_j$ for all $i \neq j$, and the decomposition is called the *orthogonal rank decomposition*.
- The *strongly orthogonal rank* of A is defined to be the minimum r such that A can be written as ($\dagger$) **and** $U_i \perp_s U_j$ for all $i \neq j$, and the decomposition is called the *strongly orthogonal rank decomposition*.

Example

Let $a, b \in \Re^m$ with $a^T b = 0$. Let $\sigma_1, \sigma_2, \sigma_3 > 0$.

Let A be an $m \times m \times m$ tensor defined by

$$A = \sigma_1 \, a \otimes b \otimes b + \sigma_2 \, b \otimes b \otimes b + \sigma_3 \, a \otimes a \otimes b$$

So the strong orthogonal rank of A is 3.

The first two decomposed tensors in A can be combined to yield

$$A = \sqrt{\sigma_1^2 + \sigma_2^2} \; \frac{\sigma_1 a + \sigma_2 b}{\sqrt{\sigma_1^2 + \sigma_2^2}} \otimes b \otimes b + \sigma_3 \, a \otimes a \otimes b,$$

And the orthogonal rank is 2.

Alternatively, we can combine the first and third decomposed tensors in A to yield

$$A = \sqrt{\sigma_1^2 + \sigma_3^2} \;\; a \otimes \frac{\sigma_1 b + \sigma_3 a}{\sqrt{\sigma_1^2 + \sigma_3^2}} \otimes b + \sigma_2 \, b \otimes b \otimes b,$$

Observe that the orthogonal rank decomposition is <u>not unique</u>.

Tensor Rank

Theorem. (L & S) For a given tensor A,

$$\text{rank}(A) \;\leq\; \text{orthog rank}(A) \;\leq\; \text{strong orthog rank}(A). \qquad (\ddagger)$$

Furthermore, equality holds if the order of A is 2.

Corollary. For any order $(n > 2)$, there exists tensor of order n such that strict inequality holds in $(\ddagger)$.

Corollary. For any order $n > 2$ there exists a tensor that cannot be decomposed as the weighted sum of completely orthogonal decomposed tensors.

Corollary. (L & S) If a tensor can be decomposed as the weighted sum of completely orthogonal decomposed tensors, then equality holds in $(\ddagger)$.

Matrices (tensors of order 2) are special cases where we can always find a completely orthogonal decomposition.

Orthogonalizing a Tensor (1/3)

Let A be an $m_1 \times m_2 \times m_3$ tensor:

$$A = \sigma_1 U + \sigma_2 V,$$

where $\sigma_1 \geq \sigma_2$ and,

$$\begin{aligned} U &= u^{(1)} \otimes u^{(2)} \otimes u^{(3)}, \\ V &= v^{(1)} \otimes v^{(2)} \otimes v^{(3)}, \end{aligned}$$

with $u^{(i)}, v^{(i)}$ unequal, non-orthogonal unit vectors in $\Re^{m_i}$ for $i = 1, 2, 3$. We can decompose each $v^{(i)}$ as

$$v^{(i)} = \alpha^{(i)} u^{(i)} + \hat{\alpha}^{(i)} \hat{u}^{(i)},$$

where

$$\begin{aligned} \alpha^{(i)} &= v^{(i)} \cdot u^{(i)}, \\ \hat{\alpha}^{(i)} &= \|v^{(i)} - \alpha^{(i)} u^{(i)}\|, \text{ and} \\ \hat{u}^{(i)} &= (v^{(i)} - \alpha^{(i)} u^{(i)}) / \hat{\alpha}^{(i)}. \end{aligned}$$

Orthogonalizing a Tensor (2/3)

Then, we can rewrite A as

$$\begin{aligned} A = & (\sigma_1 + \sigma_2\, \alpha^{(1)}\alpha^{(2)}\alpha^{(3)}) && u^{(1)} \otimes u^{(2)} \otimes u^{(3)} \\ + & \sigma_2\, \alpha^{(1)}\alpha^{(2)}\hat{\alpha}^{(3)} && u^{(1)} \otimes u^{(2)} \otimes \hat{u}^{(3)} \\ + & \sigma_2\, \alpha^{(1)}\hat{\alpha}^{(2)}\alpha^{(3)} && u^{(1)} \otimes \hat{u}^{(2)} \otimes u^{(3)} \\ + & \sigma_2\, \alpha^{(1)}\hat{\alpha}^{(2)}\hat{\alpha}^{(3)} && u^{(1)} \otimes \hat{u}^{(2)} \otimes \hat{u}^{(3)} \\ + & \sigma_2\, \hat{\alpha}^{(1)}\alpha^{(2)}\alpha^{(3)} && \hat{u}^{(1)} \otimes u^{(2)} \otimes u^{(3)} \\ + & \sigma_2\, \hat{\alpha}^{(1)}\alpha^{(2)}\hat{\alpha}^{(3)} && \hat{u}^{(1)} \otimes u^{(2)} \otimes \hat{u}^{(3)} \\ + & \sigma_2\, \hat{\alpha}^{(1)}\hat{\alpha}^{(2)}\alpha^{(3)} && \hat{u}^{(1)} \otimes \hat{u}^{(2)} \otimes u^{(3)} \\ + & \sigma_2\, \hat{\alpha}^{(1)}\hat{\alpha}^{(2)}\hat{\alpha}^{(3)} && \hat{u}^{(1)} \otimes \hat{u}^{(2)} \otimes \hat{u}^{(3)}. \end{aligned}$$

This is a strong orthogonal rank decomposition of A, and so the strong orthogonal rank is 8.

Orthogonalizing a Tensor (3/3)

Combining each pair of lines in the previous equation, we get

$$\begin{aligned} A &= \sqrt{\gamma^2 + \hat{\gamma}^2} \quad u^{(1)} \otimes u^{(2)} \otimes w^{(3)} \\ &+ \sigma_2\, \alpha^{(1)}\hat{\alpha}^{(2)} \quad u^{(1)} \otimes \hat{u}^{(2)} \otimes v^{(3)} \\ &+ \sigma_2\, \hat{\alpha}^{(1)}\alpha^{(2)} \quad \hat{u}^{(1)} \otimes u^{(2)} \otimes v^{(3)} \\ &+ \sigma_2\, \hat{\alpha}^{(1)}\hat{\alpha}^{(2)} \quad \hat{u}^{(1)} \otimes \hat{u}^{(2)} \otimes v^{(3)}. \end{aligned}$$

where $\gamma = \sigma_1 + \sigma_2\, \alpha^{(1)}\alpha^{(2)}\alpha^{(3)}$, $\hat{\gamma} = \sigma_2\, \alpha^{(1)}\alpha^{(2)}\hat{\alpha}^{(3)}$, and $w^{(3)} = (\gamma u^{(3)} + \hat{\gamma}\hat{u}^{(3)})/\sqrt{\gamma^2 + \hat{\gamma}^2}$.

Combining the last two lines of the previous line yields

$$\begin{aligned} A &= \sqrt{\gamma^2 + \hat{\gamma}^2} \quad u^{(1)} \otimes u^{(2)} \otimes w^{(3)} \\ &+ \sigma_2\, \alpha^{(1)}\hat{\alpha}^{(2)} \quad u^{(1)} \otimes \hat{u}^{(2)} \otimes v^{(3)} \\ &+ \sigma_2\, \hat{\alpha}^{(1)}\alpha^{(2)} \quad \hat{u}^{(1)} \otimes v^{(2)} \otimes v^{(3)}. \end{aligned}$$

So the orthogonal rank of A is 3.

Uniqueness Fix

- Consider all rank r (strong) orthogonal decompositions.
- For $j = 1, 2, \ldots, r$,

  Eliminate all decompositions such that

$$\sigma_j \neq \max \sigma_j.$$

- Then we are at least guaranteed that the σ_j's are unique.

Greedy Tensor Decompositions

We define the *greedy orthogonal decomposition* as follows.

Let A be a tensor. Define

$$U_1 \equiv \arg\max A \cdot U \quad \text{s.t.} \quad U \in \mathcal{D},$$

where $\mathcal{D}$ is the set of all decomposed tensors with unit norm, and define $\sigma_1 = A \cdot U_1$.

Define

$$U_{k+1} \equiv \arg\max (A - A_k) \cdot U \quad \text{s.t.} \quad U \in \mathcal{D}, \quad U \perp \mathcal{U}_k,$$

where $A_k = \sum_{i=1}^{k} \sigma_i U_i$ and $\mathcal{U}_k = \{U_1, \ldots, U_k\}$, and define $\sigma_{k+1} = A \cdot U_{k+1}$.

A *greedy strong orthogonal decomposition* can be similarly described.

Lemma. The greedy (strong) orthogonal decomposition is finite.

Eckart-Young Extension?

Does the greedy (strong) orthogonal decomposition produce a (strong) orthogonal rank decomposition?

Theorem? (L&S) Let the 'unique' orthogonal rank decomposition of a tensor A be given by

$$A = \sum_{i=1}^{r} \sigma_i U_i.$$

Then the best orthogonal rank-k approximation to A satisfies

$$\sum_{i=k+1}^{r} \sigma_i^2 = \min \|A - A_k\|^2 \text{ s.t. orthog rank}(A_k) = k,$$

and

$$A_k \equiv \sum_{i=1}^{k} \sigma_i U_i.$$

The same can be said for the strong orthogonal case.

Counterexample for Strong Orthogonal Case

Let the m-vectors a, b, c, d be pairwise orthogonal, and define the $m \times m \times m$ tensor $A = \sum_{i=1}^{6} \sigma_i U_i$ as follows.

$$
\begin{aligned}
A &= 1.00\; a \otimes a \otimes a \\
&+ 0.75\; b \otimes b \otimes b \\
&+ 0.70\; a \otimes c \otimes d \\
&+ 0.70\; a \otimes d \otimes c \\
&+ 0.65\; b \otimes c \otimes d \\
&+ 0.65\; b \otimes d \otimes c
\end{aligned}
$$

$$\gamma_1 V_1 \equiv \sqrt{\sigma_3^2 + \sigma_5^2}\, \frac{\sigma_3 a + \sigma_5 b}{\sqrt{\sigma_3^2 + \sigma_5^2}} \otimes c \otimes d$$

$$\gamma_2 V_2 \equiv \sqrt{\sigma_4^2 + \sigma_6^2}\, \frac{\sigma_4 a + \sigma_6 b}{\sqrt{\sigma_4^2 + \sigma_6^2}} \otimes d \otimes c$$

$$\gamma_1 = \gamma_2 \approx 0.9552 < \sigma_1 = 1,$$

$$\text{So } A_1 = \sigma_1 U_1.$$

On the other hand ...

$$\gamma_1^2 + \gamma_2^2 = 1.825 > \sigma_1^2 + \sigma_2^2 = 1.5625.$$

$$\text{So } A_2 = \gamma_1 V_1 + \gamma_2 V_2!$$

Summary

- Counterexample to Eckart-Young extension for strong orthogonal decomposition.

Open Questions

- Is there an Eckart-Young extension for the orthogonal decomposition?
- How can we efficiently calculate the orthogonal decomposition?
- What are other applications of such decompositions?
- Eigendecomposition?
- Notions of symmetry? Partial symmetry?

Contact Information

Tamara G. Kolda
Sandia National Labs
Livermore, CA 94551-9214
tgkolda@sandia.gov
http://www.ca.sandia.gov/~tgkolda/
+1-925-560-5639