Orthogonal Tensor Decompositions

Tamara G. Kolda

Sandia National Labs

Funding for this work has been provided by
DOE's Applied Mathematical Sciences Research Program.

Outline

- Orthogonal Decomposition of a Matrix
- Notation & Terminology for Tensors
- Definition of a Rank-1 Tensor
- Definitions of Orthogonality for Tensors
  - Orthogonal, Strongly Orthogonal, Completely Orthogonal
- Tensor Rank Decompositions
  - Orthogonal Rank, Strong Orthogonal Rank
- Examples
- Best Rank-k Approximation?
  - Orthogonal Rank, Strong Orthogonal Rank

Orthogonal Decomposition of a Matrix

Singular Value Decomposition

$$A = U\Sigma V^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

Eckart-Young Theorem (1939) – The solution to

$$\min \|A - B\| \ \text{ s.t. } \operatorname{rank}(B) = k$$

is given by

$$B \equiv \sum_{i=1}^{k} \sigma_i u_i v_i^T$$

Can we say something similar about the best rank-k approximation of a tensor?

Matrix ⇒ Tensor Notation

$$\begin{aligned} A &= \sum_{i=1}^{r} \sigma_i \, u_i v_i^T \\ &= \sum_{i=1}^{r} \sigma_i \, u_i \circ v_i \\ &= \sum_{i=1}^{r} \sigma_i \, u_i^{(1)} \circ u_i^{(2)} \\ &= \sum_{i=1}^{r} \sigma_i \, U_i \end{aligned}$$

Tensor Terminology

Tensor - an n-way array

$$A \in \mathbb{R}^{m_1 \times m_2 \times \cdots \times m_n}$$

The order of A is n. The jth dimension of A is m_j.

Inner Product of two Tensors

$$\langle A, B \rangle = \sum_{i_1=1}^{m_1} \sum_{i_2=1}^{m_2} \cdots \sum_{i_n=1}^{m_n} A_{i_1,i_2,\cdots,i_n} B_{i_1,i_2,\cdots,i_n}$$

Norm of a Tensor (Frobenius norm for matrices)

$$\|A\|^2 \equiv \langle A, A \rangle = \sum_{i_1=1}^{m_1} \sum_{i_2=1}^{m_2} \cdots \sum_{i_n=1}^{m_n} A^2_{i_1,i_2,\cdots,i_n}.$$

Rank-1 = Decomposed

Decomposed tensor

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

$$u^{(j)} \in \mathbb{R}^{m_j} \text{ for each } j$$

Each $u^{(j)}$ is called a component of U.

The $(i_1, i_2, \ldots, i_n)$-entry of U is given by

$$U_{i_1,i_2,\ldots,i_n} = u^{(1)}_{i_1} u^{(2)}_{i_2} \cdots u^{(n)}_{i_n}$$

Decomposed tensors are the building blocks of tensor decompositions.

Products of Decomposed Tensors

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

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

Inner Product - Product of products of components

$$\langle U, V \rangle = \prod_{j=1}^{n} \langle u^{(j)}, v^{(j)} \rangle$$

Norm - Product of norms of components

$$\|U\| = \prod_{j=1}^{n} \|u^{(j)}\|_2.$$

Sums of Decomposed Tensors

$$U = u^{(1)} \circ u^{(2)} \circ u^{(3)} \circ \cdots \circ u^{(n)}$$

$$V = u^{(1)} \circ v^{(2)} \circ u^{(3)} \circ \cdots \circ u^{(n)}$$

Lemma (K) The sum of U and V is a decomposed tensor if and only if *all but at most one of the components of U and V are equal* (within a scalar multiple).

$$W = \sqrt{2}\, u^{(1)} \circ \frac{u^{(2)} + v^{(2)}}{\sqrt{2}} \circ u^{(3)} \circ \cdots \circ u^{(n)}$$

$\rightarrow$ Not necessarily true for *three* or more tensors $\leftarrow$

Orthogonality

$$U = u^{(1)} \circ u^{(2)} \circ \cdots \circ u^{(n)}$$
$$V = v^{(1)} \circ v^{(2)} \circ \cdots \circ v^{(n)}$$

Orthogonal ($U \perp V$)

$$\langle U, V \rangle = \prod_{j=1}^{n} \langle u^{(j)}, v^{(j)} \rangle = 0$$

Strongly Orthogonal ($U \perp_s V$)

$$U \perp V \text{ and } \langle u^{(j)}, v^{(j)} \rangle = 0 \ \underline{\text{or}}\ u^{(j)} = v^{(j)} \text{ for each } j$$

Completely Orthogonal ($U \perp_c V$)

$$\langle u^{(j)}, v^{(j)} \rangle = 0 \text{ for each } j$$

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

Orthogonal Rank Tensor Decompositions

Find the minimal r such that A can be expressed as

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

- **Orthogonal Rank Decomposition**

$$U_i \perp U_j \text{ for all } i \neq j$$

- **Strong Orthogonal Rank Decomposition**

$$U_i \perp_s U_j \text{ for all } i \neq j$$

No completely orthogonal decomposition in general!

Tensor Rank

Theorem (Leibovici & Sabatier)

$$\text{orthog rank}(A) \leq \text{strong orthog rank}(A)$$

Furthermore, equality holds if A has a completely orthogonal decomposition.

Corollary (K) For any order $n > 2$, there exists a tensor of order n such that

$$\text{orthog rank}(A) < \text{strong orthog rank}(A)$$

Furthermore, that tensor cannot be decomposed as the weighted sum of completely orthogonal decomposed tensors.

What Makes Matrices Special?

$$A = \sum_{i=1}^{r} \sigma_i \;\; u_i \circ v_i$$

For any $i \neq j$, we have

$$(u_i \circ v_i) \perp_c (u_j \circ v_j)$$

since the SVD has the property that

$$u_i \perp u_j \quad \text{and} \quad v_i \perp v_j$$

Matrices always have a completely orthogonal decomposition

Example

Strong Orthogonal Decomposition

Assume $\|a\| = \|b\| = 1$, $a \perp b$, and $\sigma_1 > \sigma_2 > \sigma_3 > 0$, and

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

$\Rightarrow$ Strong Orthogonal Rank of $A = 3$.

Orthogonal Rank Decomposition

$$A = \sqrt{\sigma_1^2 + \sigma_2^2} \, \left(\frac{\sigma_1 a + \sigma_2 b}{\sqrt{\sigma_1^2 + \sigma_2^2}} \right) \circ b \circ b + \sigma_3 \, a \circ a \circ a,$$

$\Rightarrow$ Orthogonal Rank of $A = 2$.

Strong Orthogonal Decomposition is Not Unique

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

can also be expressed as ...

$$A = \hat{\sigma}_1 \, \hat{a} \circ b \circ b \; + \; \hat{\sigma}_2 \, \hat{a} \circ a \circ a \; + \; \hat{\sigma}_3 \, \hat{b} \circ a \circ a$$

where

$$\hat{\sigma}_1 = \sqrt{\sigma_1^2 + \sigma_2^2} \quad \hat{\sigma}_2 = \frac{\sigma_1 \, \sigma_3}{\hat{\sigma}_1} \quad \hat{\sigma}_3 = \frac{\sigma_2 \, \sigma_3}{\hat{\sigma}_1}$$

$$\hat{a} = \frac{\sigma_1 \, a + \sigma_2 \, b}{\hat{\sigma}_1} \quad \hat{b} = \frac{\sigma_2 \, a - \sigma_1 \, b}{\hat{\sigma}_1}$$

Note $\hat{a} \perp \hat{b}$.

Orthogonal Decomposition is Not Unique

$$A = \overbrace{\underbrace{\sigma_1\; a \circ b \circ b}} + \overbrace{\sigma_2\; b \circ b \circ b} + \underbrace{\sigma_3\; a \circ a \circ b}$$

Orthogonal Decomposition 1: (1st and 2nd terms)

$$A = \sqrt{\sigma_1^2 + \sigma_2^2}\;\left(\frac{\sigma_1 a + \sigma_2 b}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \circ b \circ b + \sigma_3\; a \circ a \circ a,$$

Orthogonal Decomposition 2: (1st and 3rd terms)

$$A = \sqrt{\sigma_1^2 + \sigma_3^2}\;\; a \circ \left(\frac{\sigma_1 b + \sigma_3 a}{\sqrt{\sigma_1^2 + \sigma_3^2}}\right) \circ b + \sigma_2\; b \circ b \circ b,$$

Problem of Uniqueness

Goal: Find the best approximation of A with orthogonal rank k.

Approach: Take the first k terms of the orthogonal rank decomposition.

Problem: Which of the orthogonal decompositions should we use?

Solution**?** Among all possible decompositions, choose that which has $\sigma_j \geq \max \sigma_j$ for each successive j.

[Ditto for strong orthogonal rank decomposition]

Eckart-Young Extension?

Question: Let the 'unique' orthogonal rank decomposition of a tensor A be given by

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

Is it true that the best orthogonal rank-k approximation to A is given by

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

What about the the strong orthogonal case?

Strong Orthog Counterexample

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.

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

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

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

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

So $A_1 = \sigma_1 U_1$.

On the other hand ...

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

So $A_2 = \gamma_1 V_1 + \gamma_2 V_2$!

Orthogonal Counterexample

Assume $a \perp b$, $c = \frac{1}{\sqrt{2}}(a+b)$, and $\sigma_1 > \sigma_2$

$$A = \sigma_1 \; a \circ a \circ a + \sigma_2 \; c \circ c \circ b$$

Best Rank-1 Approximation

$$A_1 \equiv \gamma \; x \circ y \circ z$$

$$x = \alpha_x a + \beta_x b \quad y = \alpha_y a + \beta_y b \quad z = \alpha_z a + \beta_z b$$

$$\gamma = \sigma_1 \alpha_x \alpha_y \alpha_z + \frac{\sigma_2}{2}(\alpha_x + \beta_x)(\alpha_x + \beta_y)\beta_z$$

$$\alpha_x \neq 0, \; \alpha_y \neq 0, \; \alpha_z \neq 0, \; \beta_x \neq 0, \; \beta_y \neq 0, \; \beta_z \neq 0$$

But cannot extend A_1 to a
2-term orthogonal rank decomposition.

Summary

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

⇒ Counterexample to Eckart-Young extension for orthogonal rank decomposition.

Larger Issues

- How can we efficiently calculate tensor decompositions?
- What are other applications of such decompositions?
- Can we exploit structure such as (partial) symmetry?

Contact Information

Tamara G. Kolda
Sandia National Labs, Livermore, CA, USA
tgkolda@sandia.gov
http://csmr.ca.sandia.gov/~tgkolda/