On critical rank-k approximations to tensors

Jan Draisma
(jww Giorgio Ottaviani and Alicia Tocino)

Atlanta, August 2017

SVD

$A \in \mathbb{R}^{m\times n}$, $m \leq n \rightsquigarrow A = \sum_{i=1}^{m} \sigma_i u_i v_i^T$ with singular values $\sigma_1 \geq \cdots \geq \sigma_m \geq 0$ and $(u_i|u_j) = (v_i|v_j) = \delta_{ij}$.

Theorem

$\sum_{i=1}^{k} \sigma_i u_i v_i^T$ minimises $d_A(B) := ||A - B||^2 = \sum_{i,j}(a_{ij} - b_{ij})^2$ among rank $\leq k$-matrices.

The Eckart-Young Theorem revisited

SVD

$A \in \mathbb{R}^{m \times n}$, $m \leq n \rightsquigarrow A = \sum_{i=1}^{m} \sigma_i u_i v_i^T$ with singular values $\sigma_1 \geq \cdots \geq \sigma_m \geq 0$ and $(u_i|u_j) = (v_i|v_j) = \delta_{ij}$.

Theorem

$\sum_{i=1}^{k} \sigma_i u_i v_i^T$ minimises $d_A(B) := ||A - B||^2 = \sum_{i,j} (a_{ij} - b_{ij})^2$ among rank $\leq k$-matrices.

Refinement

If $\sigma_1 > \cdots > \sigma_m > 0$, then the critical points of d_A on the manifold of rank-k matrices are $\sum_{i \in I} \sigma_i u_i v_i^T$ for $|I| = k$.

These lie in the span of the critical rank-1 approximations.

Main result

Theorem (D-Ottaviani-Tocino)
Under mild conditions, the critical rank-k approximations to a *tensor* f also lie in the span of its critical rank-1 approximations.

Main result

Theorem (D-Ottaviani-Tocino)
Under mild conditions, the critical rank-k approximations to a *tensor* f also lie in the span of its critical rank-1 approximations.

Setting

- $V_1, \dots, V_p$ f.d. $\mathbb{C}$-spaces with symmetric bilinear forms $(.|.)$
- $d_1, \dots, d_p$ natural numbers ≥ 1
- $T := S^{d_1} V_1 \otimes \cdots \otimes S^{d_p} V_p$ equipped with $(.|.)$ satisfying

$$(v_1^{d_1} \otimes \cdots \otimes v_p^{d_p} | w_1^{d_1} \otimes \cdots \otimes w_p^{d_p}) = (v_1|w_1)^{d_1} \cdots (v_p|w_p)^{d_p}$$

Main result

Theorem (D-Ottaviani-Tocino)
Under mild conditions, the critical rank-k approximations to a *tensor* f also lie in the span of its critical rank-1 approximations.

Setting

- $V_1, \ldots, V_p$ f.d. $\mathbb{C}$-spaces with symmetric bilinear forms $(.|.)$
- $d_1, \ldots, d_p$ natural numbers ≥ 1
- $T := S^{d_1} V_1 \otimes \cdots \otimes S^{d_p} V_p$ equipped with $(.|.)$ satisfying

$(v_1^{d_1} \otimes \cdots \otimes v_p^{d_p} | w_1^{d_1} \otimes \cdots \otimes w_p^{d_p}) = (v_1|w_1)^{d_1} \cdots (v_p|w_p)^{d_p}$

Mild conditions

- $f \in T$ is sufficiently general
- for all i with $d_i = 1$: $(\dim V_i - 1) \leq \sum_{j \neq i} (\dim V_j - 1)$

Necessary?

Main result

Theorem (D-Ottaviani-Tocino)
Under mild conditions, the critical rank-k approximations to a *tensor* f also lie in the span of its critical rank-1 approximations.

Setting

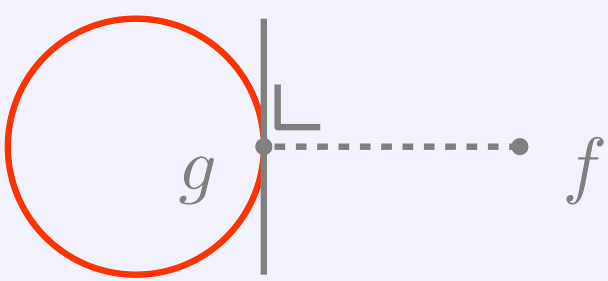
- $X := \{v_1^{d_1} \otimes \cdots \otimes v_p^{d_p}\} \leq T$ the variety of rank ≤ 1 tensors
- $\sigma_k X := \overline{\{x_1 + \cdots + x_k \mid x_i \in X\}}$ the k-th secant variety

Theorem (D-Ottaviani-Tocino)
Under mild conditions, the critical rank-k approximations to a *tensor* f also lie in the span of its critical rank-1 approximations.

Setting

- $X := \{v_1^{d_1} \otimes \cdots \otimes v_p^{d_p}\} \leq T$ the variety of rank ≤ 1 tensors
- $\sigma_k X := \overline{\{x_1 + \cdots + x_k \mid x_i \in X\}}$ the k-th secant variety

Definition
A *critical rank-k approximation* to $f \in T$ is a smooth point $g \in \sigma_k X$ such that $f - g \perp T_g \sigma_k X$.



For each $i \in \{1, \ldots, p\}$, there is a natural skew bilinear map $[.|.]_i : T \times T \to \bigwedge^2 V_i$ constructed from the bilinear forms.

Definition

The *critical space* for $f \in T$ is $H_f := \{g \in T \mid \forall i : [f|g]_i = 0\}$.

The critical space

For each $i \in \{1, \ldots, p\}$, there is a natural skew bilinear map $[.|.]_i : T \times T \to \bigwedge^2 V_i$ constructed from the bilinear forms.

Definition

The *critical space* for $f \in T$ is $H_f := \{g \in T \mid \forall i : [f|g]_i = 0\}$.

Example: matrices

For $A = \sum_{i=1}^{m} \sigma_i u_i v_i^T \in \mathbb{R}^{m \times n}$ this is the set of B such that AB^T and $A^T B$ are both symmetric; so each $u_j v_j^T \in H_A$.

Moreover, if the σ_i are positive and distinct, then H_A is the span of the $u_j v_j^T$.

Remark

H_f was called *singular space* by Ottaviani-Paoletti.

Proof sketch

Proposition 1

Under same mild conditions, $\operatorname{codim}_T H_f = \sum_{i=1}^{p} \dim \bigwedge^2 V_i$.

Proof sketch

Proposition 1

Under same mild conditions, $\operatorname{codim}_T H_f = \sum_{i=1}^{p} \dim \bigwedge^2 V_i$.

Proposition 2

The critical rank-k approximations to f lie in H_f.

Proof sketch

Proposition 1

Under same mild conditions, $\operatorname{codim}_T H_f = \sum_{i=1}^p \dim \bigwedge^2 V_i$.

Proposition 2

The critical rank-k approximations to f lie in H_f.

Proposition 3

The critical rank-*one* approximations to f span a space of the same codimension $\sum_i \dim \bigwedge^2 V_i$.

Proof sketch

Proposition 1
Under same mild conditions, $\operatorname{codim}_T H_f = \sum_{i=1}^p \dim \bigwedge^2 V_i$.

Proposition 2
The critical rank-k approximations to f lie in H_f.

Proposition 3
The critical rank-*one* approximations to f span a space of the same codimension $\sum_i \dim \bigwedge^2 V_i$.

Ad 3: Following Friedlander-Ottaviani, interpret the rank-one approximations as the zeroes of a section of a certain vector bundle on $\mathbb{P}V_1 \times \cdots \times \mathbb{P}V_p$, and we use vector bundle techniques.

Ad 1: Find an explicit (sparse) f for which this holds.

Proof sketch of Proposition 2

Definition

The *critical space* for $f \in T$ is $H_f := \{g \in T \mid \forall i : [f|g]_i = 0\}$.

Proposition 2

The critical rank-k approximations to f lie in H_f.

Proof sketch of Proposition 2

Definition

The *critical space* for $f \in T$ is $H_f := \{g \in T \mid \forall i : [f|g]_i = 0\}$.

Proposition 2

The critical rank-k approximations to f lie in H_f.

- Let $g := x_1 + \cdots + x_k$ be critical for f, so $\forall i : f - g \perp T_{x_i} X$.

Proof sketch of Proposition 2

Definition

The *critical space* for $f \in T$ is $H_f := \{g \in T \mid \forall i : [f|g]_i = 0\}$.

Proposition 2

The critical rank-k approximations to f lie in H_f.

• Let $g := x_1 + \cdots + x_k$ be critical for f, so $\forall i : f - g \perp T_{x_i}X$.

• Write $x_1 = v_1^{d_1} \otimes \cdots \otimes v_p^{d_p}$, and extend each v_i to an orthogonal basis of V_i. This gives an x_1-adapted *monomial basis* of T.

Proof sketch of Proposition 2

Definition

The *critical space* for $f \in T$ is $H_f := \{g \in T \mid \forall i : [f|g]_i = 0\}$.

Proposition 2

The critical rank-k approximations to f lie in H_f.

- Let $g := x_1 + \cdots + x_k$ be critical for f, so $\forall i : f - g \perp T_{x_i}X$.
- Write $x_1 = v_1^{d_1} \otimes \cdots \otimes v_p^{d_p}$, and extend each v_i to an orthogonal basis of V_i. This gives an x_1-adapted *monomial basis* of T.
- Now $f - g$ is a linear combination of monomials that have gcd of degree $< -1 + \sum_i d_i$ with x_1. Hence $\forall i : [f - g|x_1]_i = 0$.

Proof sketch of Proposition 2

Definition
The *critical space* for $f \in T$ is $H_f := \{g \in T \mid \forall i : [f|g]_i = 0\}$.

Proposition 2
The critical rank-k approximations to f lie in H_f.

- Let $g := x_1 + \cdots + x_k$ be critical for f, so $\forall i : f - g \perp T_{x_i} X$.
- Write $x_1 = v_1^{d_1} \otimes \cdots \otimes v_p^{d_p}$, and extend each v_i to an orthogonal basis of V_i. This gives an x_1-adapted *monomial basis* of T.
- Now $f - g$ is a linear combination of monomials that have gcd of degree $< -1 + \sum_i d_i$ with x_1. Hence $\forall i : [f - g|x_1]_i = 0$.
- Similarly for x_2 etc, so $[f - g|g]_i = [f - g|x_1 + \cdots + x_k]_i = 0$.

Proof sketch of Proposition 2

Definition
The *critical space* for $f \in T$ is $H_f := \{g \in T \mid \forall i : [f|g]_i = 0\}$.

Proposition 2
The critical rank-k approximations to f lie in H_f.

- Let $g := x_1 + \cdots + x_k$ be critical for f, so $\forall i : f - g \perp T_{x_i} X$.
- Write $x_1 = v_1^{d_1} \otimes \cdots \otimes v_p^{d_p}$, and extend each v_i to an orthogonal basis of V_i. This gives an x_1-adapted *monomial basis* of T.
- Now $f - g$ is a linear combination of monomials that have gcd of degree $< -1 + \sum_i d_i$ with x_1. Hence $\forall i : [f - g|x_1]_i = 0$.
- Similarly for x_2 etc, so $[f - g|g]_i = [f - g|x_1 + \cdots + x_k]_i = 0$.
- Since $[g|g]_i = 0$, also $[f|g]_i = 0$. $\square$

Recap

Theorem (D-Ottaviani-Tocino)
Under mild conditions, the critical rank-k approximations to a *tensor* lie in the span of its critical rank-1 approximations.

Disclaimer
This does *not* mean that a best rank-k approximation can be found by iteratively subtracting best rank-1 approximations. This is true only seldomly (Vannieuwenhoven, Nicaise, Vandebril, and Meerbergen).

On the arXiv soon . . . comments welcome!