



Non-negative Tensor Approximations

Yang Qi, Pierre Comon

► To cite this version:

| Yang Qi, Pierre Comon. Non-negative Tensor Approximations. 2014. hal-01015519v1

HAL Id: hal-01015519

<https://hal.science/hal-01015519v1>

Preprint submitted on 26 Jun 2014 (v1), last revised 12 Apr 2016 (v5)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NON-NEGATIVE TENSOR APPROXIMATIONS

YANG QI, PIERRE COMON

ABSTRACT. In this note, we study the relations between approximations and singular vector tuples.

Nonnegative tensors are widely used in several fields, including *e.g.*, machine learning, hyperspectral imaging, spectrography, and chemometrics, among others. There exist known sufficient conditions ensuring uniqueness of tensor decomposition into a sum of rank-1 terms. Yet, these conditions do not apply neither to nonnegative decompositions nor to low-rank approximations. The purpose of this note is to provide some first results in the latter direction.

1. EXISTENCE AND UNIQUENESS OF APPROXIMATIONS

Let V_1, \dots, V_n be real vector spaces, and V_i^+ denote the set of non-negative vectors in V_i . Let V^+ be the set of non-negative tensors in $V_1 \otimes \dots \otimes V_n$. Given a $T \in V^+$, we consider the best rank- r approximations of T , where r is less than the non-negative rank of T .

Definition 1.1. For a fixed positive integer r , let $D_r = \{X \in V^+ | \text{rank}(X) \leq r\}$, where $\text{rank}(X)$ means the non-negative rank of X , and let $d(T) = \inf_{X \in D_r} \|T - X\|$, where $\|\cdot\|$ is the l^2 -norm.

According to [1], D_r is a closed set, thus for any $T \notin D_r$, there is some $T^* \in D_r$ such that $\|T - T^*\| = d(T)$.

Proposition 1.2. For any nonnegative $T \notin D_r$, $\min_{X \in D_r} \|T - X\| = \min_{\text{rank}(X)=r} \|T - X\|$.

Proof. Let $\|T - T^*\| = \min_{X \in D_r} \|T - X\|$, and assume $\text{rank}(T^*) < r$. Since $T \notin D_r$, $T - T^* \neq 0$. If $T - T^* \leq 0$, pick i_1, \dots, i_n such that $(T - T^*)_{i_1 \dots i_n} < 0$, let $T'_{i_1 \dots i_n} = \lambda(T - T^*)_{i_1 \dots i_n}$ for some $0 < \lambda < 1$ and $T'_{j_1 \dots j_n} = 0$ for other indices, then $T^* + T' \geq 0$, $\text{rank}(T^* + T') \leq r$ and $\|T - T^* - T'\| < \|T - T^*\|$ which is a contradiction. So $(T - T^*)_{i_1 \dots i_n} > 0$ for some i_1, \dots, i_n , let $T'_{i_1 \dots i_n} = (T - T^*)_{i_1 \dots i_n}$ and $T'_{j_1 \dots j_n} = 0$ for other indices, then $\text{rank}(T^* + T') \leq r$ and $\|T - T^* - T'\| < \|T - T^*\|$ which is a contradiction. Therefore $\text{rank}(T^*) = r$. \square

The following observation is based on some fundamental facts of approximation analysis [2], and we give it a proof for completeness.

Proposition 1.3. Almost every $T \in V^+$ has a unique best rank- r approximation.

Proof. For any $T, T' \in V_1 \otimes \dots \otimes V_n$, $|d(T) - d(T')| \leq \|T - T'\|$, i.e. d is Lipschitz, thus differentiable a.e. in $V_1 \otimes \dots \otimes V_n$ by Rademacher's theorem.

The project is funded by the European Research Council under the European Community's Seventh Framework Program FP7/2007-2013 Grant Agreement no. 320594.

For a general $T \in V^+$, there is an open neighbourhood $B(T, \delta)$ of T contained in V^+ , so d is differentiable a.e. in V^+ . Assume that d is differentiable at $T \in V^+$, for any $U \in V_1 \otimes \cdots \otimes V_n$, let $\partial d_T^2(U)$ be the differential of d^2 at T along the direction U , let $\|T - T^*\| = d(T)$. Then,

$$d^2(T + tU) = d^2(T) + t\partial d_T^2(U) + O(t^2)$$

$$d^2(T + tU) \leq \|T + tU - T^*\|^2 = d^2(T) + 2t\langle U, T - T^* \rangle + t^2 \|U\|^2.$$

Therefore, for any t , $t\partial d_T^2(U) \leq 2t\langle U, T - T^* \rangle$, then

$$\partial d_T^2(U) = 2\langle U, T - T^* \rangle.$$

Assume T' is another best rank- r approximation of T , then

$$2\langle U, T - T^* \rangle = \partial d_T^2(U) = 2\langle U, T - T' \rangle,$$

which implies $\langle T' - T^*, U \rangle = 0$ for any U , i.e. $T' = T^*$. \square

Proposition 1.4. The nonnegative tensors which does not have a unique best rank- r approximation form a semi-algebraic set.

Proof. D_r is semi-algebraic by Tarski-Seidenberg theorem [5], by [6, Theorem 3.4], the nonnegative tensors without a unique best rank- r approximation form a semi-algebraic set. \square

After knowing the existence and uniqueness of best rank- r approximations, we want to find out these approximations explicitly. A natural idea is to compute these approximations inductively: for $k < r$, let $\sum_{i=1}^k T_i$ be the best rank- k approximation of T , then we hope to obtain the best rank- r approximation by computing the rank- $(r - k)$ approximation of $T - \sum_{i=1}^k T_i$. The following observation tells us this does not work generally, which extends the result of [10] to nonnegative tensors.

Proposition 1.5. A best approximation of a general $T \in V^+$ can not be obtained by a sequence of lower rank approximations.

Proof. Let $\sum_{i=1}^r T_i \in D_r$, and assume $\left\| \sum_{i=1}^r T_i \right\| = 1$, so $d(T) = \min \left\| T - \alpha \sum_{i=1}^r T_i \right\|$.

Let L denote the line in $V_1 \otimes \cdots \otimes V_n$ spanned by $\sum_{i=1}^r T_i$, and L^\perp denote the orthogonal complement of L . Denote the projection of T to L by $Proj_L(T)$, then

$$\|T\|^2 = \|Proj_L(T)\|^2 + \|Proj_{L^\perp}(T)\|^2,$$

$$\left\| T - \alpha \sum_{i=1}^r T_i \right\|^2 = \|T - Proj_L(T)\|^2 = \|Proj_{L^\perp}(T)\|^2 = \|T\|^2 - \|Proj_L(T)\|^2,$$

so to compute $\min \left\| T - \alpha \sum_{i=1}^r T_i \right\|$ is equivalent to compute $\max Proj_L(T)$.

For convenience of notation, we assume $n = 3$, and let $T = [T_{jkl}]$, $T_i = u_i \otimes v_i \otimes w_i$, and $u_i = [x_j^i]$, $v_i = [y_k^i]$, $w_i = [z_l^i]$ are coordinates. Consider the Lagrangian:

$$\begin{aligned}\phi &= \langle T, \sum_{i=1}^r T_i \rangle - \lambda \left(\left\| \sum_{i=1}^r T_i \right\|^2 - 1 \right) \\ &= \sum_{i=1}^r \sum_{j,k,l} T_{jkl} x_j^i y_k^i z_l^i - \lambda \left(\sum_{j,k,l} \left(\sum_{i=1}^r x_j^i y_k^i z_l^i \right)^2 - 1 \right),\end{aligned}$$

then $\frac{\partial \phi}{\partial x_j^i} = 0$ gives $\sum_{k,l} T_{jkl} y_k^i z_l^i = 2\lambda \sum_{i=1}^r x_j^i \sum_{k,l} (y_k^i z_l^i)^2$, i.e.

$$\langle T, v_i \otimes w_i \rangle = \sum_{m=1}^r \lambda_m^i u_m,$$

where $\lambda_m^i = 2\lambda \langle v_m \otimes w_m, v_i \otimes w_i \rangle$. Similarly, we have

$$\langle T, u_i \otimes w_i \rangle = \sum_{m=1}^r \mu_m^i v_m, \quad \langle T, u_i \otimes v_i \rangle = \sum_{m=1}^r \nu_m^i w_m,$$

where $\mu_m^i = 2\lambda \langle u_m \otimes w_m, u_i \otimes w_i \rangle$, and $\nu_m^i = 2\lambda \langle u_m \otimes v_m, u_i \otimes v_i \rangle$.

Assume $\left\| T - \sum_{i=1}^s T_i \right\| = \min_{X \in D_s} \|T - X\|$, and $\left\| T - \sum_{i=1}^r T_i \right\| = \min_{X \in D_r} \|T - X\|$ for some $s < r$, then $\left\| T - \sum_{i=1}^r T_i \right\| = \min_{X \in D_{r-s}} \left\| (T - \sum_{i=1}^s T_i) - X \right\|$, so

$$\langle T, v_i \otimes w_i \rangle = \sum_{m=1}^s \lambda_m^i u_m, \text{ where } \lambda_m^i = 2\lambda \langle v_m \otimes w_m, v_i \otimes w_i \rangle, 1 \leq i \leq s,$$

$$\langle T, v_j \otimes w_j \rangle = \sum_{m=1}^r \beta_m^j u_m, \text{ where } \beta_m^j = 2\lambda \langle v_m \otimes w_m, v_j \otimes w_j \rangle, 1 \leq j \leq r,$$

$$\langle T - \sum_{i=1}^s T_i, v_k \otimes w_k \rangle = \sum_{m=s+1}^r \gamma_m^k u_m, \text{ where } \gamma_m^k = 2\lambda \langle v_m \otimes w_m, v_k \otimes w_k \rangle, s+1 \leq k \leq r,$$

which implies $\sum_{m=s+1}^r \gamma_m^k u_m$ is parallel to $\sum_{m=1}^s \lambda_m^i u_m$. By eliminating the parameters,

we can obtain the algebraic conditions that T has to satisfy. Since it is easy to find a T which does not satisfy these conditions, then a general T does not have this property. Therefore for a general T , a best rank- r approximation is not obtained from an approximation of a best rank- s approximation. \square

2. RANK ONE APPROXIMATION

Definition 2.1. For $T \in V_1 \otimes \cdots \otimes V_n$, $(\lambda, u_1 \otimes \cdots \otimes u_n)$ is called a nonnegative singular pair of T if $\lambda \geq 0$, and for all $i \in \{1, \dots, n\}$, $0 \neq u_i \geq 0$, and

$$(2.1) \quad \langle T, \otimes_{j \neq i} u_j \rangle = \lambda u_i.$$

Lemma 2.2. A nonnegative tensor T has a nonnegative singular pair.

Proof. Let $u_i = (u_{i,1}, \dots, u_{i,d_i})$ be the coordinate of u_i . Let $D = \{(u_1, \dots, u_n) | u_{i,j} \geq 0, \sum_{i,j} u_{i,j} = 1\}$, then D is a compact convex set. Define

$$\phi: D \rightarrow D$$

$$(u_1, \dots, u_n) \mapsto \left(\frac{\langle T, u_2 \otimes \dots \otimes u_n \rangle}{\sum_{i,l} \langle T, \otimes_{j \neq i} u_j \rangle_l}, \dots, \frac{\langle T, u_1 \otimes \dots \otimes u_{n-1} \rangle}{\sum_{i,l} \langle T, \otimes_{j \neq i} u_j \rangle_l} \right)$$

If $\sum_{i,l} \langle T, \otimes_{j \neq i} u_j \rangle_l = 0$, then $\langle T, \otimes_{j \neq i} u_j \rangle = 0$ for all i , i.e. $\lambda = 0$.

If $\sum_{i,l} \langle T, \otimes_{j \neq i} u_j \rangle_l > 0$, by Brouwer's Fixed Point Theorem, there is some $u_1 \otimes \dots \otimes u_n$ such that $\langle T, \otimes_{j \neq i} u_j \rangle = \lambda u_i$, where $\lambda = \sum_{i,l} \langle T, \otimes_{j \neq i} u_j \rangle_l$. \square

Lemma 2.3. *If T is positive, T has a nonnegative pair $(\lambda, u_1 \otimes \dots \otimes u_n)$ with $\lambda > 0$. If $u_1 \otimes \dots \otimes u_n$ has unit length, then $u_1 \otimes \dots \otimes u_n$ is unique and every $u_i > 0$.*

Proof. Let $I_i = \{j | u_{i,j} \neq 0\}$, and $\alpha = \min\{u_{i,j} | 1 \leq i \leq n, j \in I_i\}$. For any i and k , $\lambda u_{i,k} = \langle T, \otimes_{j \neq i} u_j \rangle_k \geq \alpha^{n-1} \sum_{l_j \in I_j} T_{l_1 \dots l_{i-1} k l_{i+1} \dots l_n} > 0$.

Assume T had two positive singular vector tuples $u_1 \otimes \dots \otimes u_n$ and $v_1 \otimes \dots \otimes v_n$ corresponding to λ ,

$$\langle T, u_1 \otimes \dots \otimes u_{i-1} \otimes \hat{u}_i \otimes u_{i+1} \otimes \dots \otimes u_n \rangle = \lambda u_i,$$

$$\langle T, v_1 \otimes \dots \otimes v_{i-1} \otimes \hat{v}_i \otimes v_{i+1} \otimes \dots \otimes v_n \rangle = \lambda v_i.$$

Let $\alpha_i = \max\{\alpha \in \mathbb{R}_{\geq 0} | u_i - \alpha v_i \geq 0\}$ and $\beta_i = \max\{\beta \in \mathbb{R}_{\geq 0} | v_i - \beta u_i \geq 0\}$. Since $\|u_i\| = \|v_i\| = 1$ and $u_i, v_i > 0$, then $0 < \alpha_i, \beta_i \leq 1$. Therefore

$$\lambda u_i = \langle T, \otimes_{j \neq i} u_j \rangle \geq \langle T, \otimes_{j \neq i} \alpha_j v_j \rangle = \lambda \prod_{j \neq i} \alpha_j \cdot v_i,$$

$$\lambda v_i = \langle T, \otimes_{j \neq i} v_j \rangle \geq \langle T, \otimes_{j \neq i} \beta_j u_j \rangle = \lambda \prod_{j \neq i} \beta_j \cdot u_i.$$

By the maximality of α_i , $\frac{\prod_{j \neq i} \alpha_j}{\alpha_i} \leq 1$ for each i , thus $\alpha_i = 1$, and similarly, $\beta_i = 1$. \square

Proposition 2.4. A positive tensor T with rank > 1 has a unique best rank one non-negative approximation.

Proof. Since the smooth function

$$\varphi: \mathbb{S}^{d_1-1} \times \dots \times \mathbb{S}^{d_n-1} \rightarrow \mathbb{R}_{\geq 0}$$

$$(u_1, \dots, u_n) \mapsto \langle T, u_1 \otimes \dots \otimes u_n \rangle$$

reaches its maximal value at some $(u_1, \dots, u_n) \geq 0$, where \mathbb{S}^{d_i-1} is unit sphere in V_i , then the critical points of the Lagrangian

$$\langle T, u_1 \otimes \dots \otimes u_n \rangle - \sum_{i=1}^n \lambda_i (\|u_i\| - 1)$$

give us

$$(2.2) \quad \langle T, u_1 \otimes \dots \otimes u_{i-1} \otimes \hat{u}_i \otimes u_{i+1} \otimes \dots \otimes u_n \rangle = \lambda_i u_i$$

and $\langle T, u_1 \otimes \cdots \otimes u_n \rangle = \lambda_i$ gives us $\lambda_1 = \cdots = \lambda_n$, denoted by λ .

Since λ is maximal, $\lambda > 0$ and $u_1 \otimes \cdots \otimes u_n$ is unique by 2.3. Hence this unique critical point $u_1 \otimes \cdots \otimes u_n$ yields the best rank one approximation. \square

In [2], Friedland and Ottaviani compute the number of singular vector tuples of a generic complex tensor. Now we study the real case in the smooth category based on their method.

Proposition 2.5. Almost every real tensor $A \in V_1 \otimes \cdots \otimes V_n$, $n \geq 3$, has an even number of singular vector tuples.

Proof. Let $X = \mathbb{S}^{d_1-1} \times \cdots \times \mathbb{S}^{d_n-1}$, and $\pi_i : X \rightarrow \mathbb{S}^{d_i-1}$ be the projection. Let N_i be the normal bundle over \mathbb{S}^{d_i-1} , and T_i be the tangent bundle of \mathbb{S}^{d_i-1} , which is isomorphic to the quotient bundle V_i/N_i . Let $E = \bigoplus_{i=1}^n \text{Hom}(\otimes_{j \neq i} \pi_j^* N_j, \pi_i^* V_i/N_i)$.

For any $[x_i] \in V_i/\langle u_i \rangle$, there is some A such that $[\text{Con}(A, \otimes_{j \neq i} u_j)] = [x_i]$ if and only if $u_1^\top(x_1 + t_1 u_1) = \cdots = u_n^\top(x_n + t_n u_n)$ for some t_i . Since $u_i^\top u_i = 1$, A always exists. Let $S = \{s \in H^0(X, E) \mid s = ([\text{Con}(A, \otimes_{j \geq 2} u_j)], \dots, [\text{Con}(A, \otimes_{j < n} u_j)])\}$, then S generates E . Then a general $s \in S$ has finite number of zeros, which is Poincaré dual to $e(E)$, the Euler class of E .

Assume all d_i are odd, then $e(E) = 2^n \alpha$ where α is a generator of $H^m(X, \mathbb{Z})$ and $m = \sum_i d_i - n$. If some d_i is even, then $e(E) = 0$ in the class. Therefore, a general A has even number of singular vector tuples. \square

3. r -SINGULAR VECTOR

Definition 3.1. A vector tuple $(u_{1,1}, \dots, u_{n,1}, \dots, u_{1,r}, \dots, u_{n,r}) \in (V_1 \times \cdots \times V_n)^r$ with $u_{i,j} \neq 0$ for all i, j is called a r -singular tuple of T if

$$(3.1) \quad \text{Con}(T, \otimes_{k \neq i} u_{k,j}) = \lambda \text{Con}(\sum_{j=1}^r u_{1,j} \otimes \cdots \otimes u_{n,j}, \otimes_{k \neq i} u_{k,j})$$

for all i, j and some λ , where Con denotes the contraction.

Similar to [2], we will show a generic T has finite number of r -singular vector tuples.

Let M be a nonsingular complex variety, and $E \xrightarrow{\pi} M$ be a holomorphic vector bundle on M with $\dim M = \text{rank} E$. Let $S \subset H^0(M, E)$ be a finite dimensional subspace, and $M \times S \xrightarrow{\nu} E$ be the evaluation map $(p, s) \mapsto s(p)$.

Lemma 3.2. *If there exists an open subset $U \subset M$ such that S_U generates $H^0(U, E)$, ν has constant rank on $U \times S_U$, and for a generic $\sigma \in S$, the zero locus of σ , Z_σ , is noetherian and contained in U , then Z_σ consists of a finite number of simple points.*

Proof. Let τ be the zero section of E , and $Z_U := \{(p, s_U) \in U \times S_U \mid s_U(p) = 0\} \subset \nu^{-1}(\tau)$. Since ν is dominant and has constant rank on $U \times S_U$, Z_U has dimension $\dim M + \dim S - \text{rank} E$. Let $p : \nu^{-1}(\tau) \rightarrow S$ be the projection, and $p_U : Z_U \rightarrow S_U$ the restriction on U . Since for a generic $\sigma \in S$, Z_σ is isomorphic to $p_U^{-1}(\sigma)$. By the generic smoothness theorem [4], Z_σ is a smooth 0-dimensional subvariety of M . Since Z_σ is noetherian, then Z_σ is of a finite number of simple points. \square

For each $1 \leq i \leq n$, let V_i be a complex vector space with dimension d_i , and for each $1 \leq j \leq r$, $V_{i,j}$ be a complex vector space isomorphic to V_i . Let X_j denote $\text{Seg}(\mathbb{P}V_{1,j} \times \cdots \times \mathbb{P}V_{n,j})$, and $\alpha_{i,j} : X_j \rightarrow \mathbb{P}V_{i,j}$ be the natural projection. Let $X = X_1 \times \cdots \times X_r$, and $\beta_i : X \rightarrow X_i$ be the projection. Let $T_{i,j}$ be the tautological line bundle over $\mathbb{P}V_{i,j}$, and $M = \bigoplus_{j=1}^r \beta_j^* \left(\bigotimes_{i=1}^n \alpha_{i,j}^*(T_{i,j}) \right)$ be a rank- r vector bundle

$M \xrightarrow{\gamma} X$. Let $F_{i,j}$ be the trivial bundle over M with fibre $V_{i,j}$, and $Q_{i,j}$ denote the quotient bundle

$$0 \rightarrow \gamma^* \cdot \beta_j^* \cdot \alpha_{i,j}^*(T_{i,j}) \rightarrow F_{i,j} \rightarrow Q_{i,j} \rightarrow 0.$$

Let $H_{i,j} = \text{Hom}(\gamma^* \cdot \beta_j^* (\bigotimes_{k \neq i} \alpha_{k,j}^*(T_{k,j})), F_{i,j})$, $L_{i,j} = \text{Hom}(\gamma^* \cdot \beta_j^* (\bigotimes_{k \neq i} \alpha_{k,j}^*(T_{k,j}^\vee)), T_{i,j})$

and $E = \bigoplus_{j=1}^r (L_{1,j} \oplus H_{1,j} \oplus \cdots \oplus H_{n,j})$. So $\text{rank } E = \dim M = r \sum_{i=1}^n d_i - r(n-1)$.

Now we fix r such that $\text{Seg}(\mathbb{P}V_1 \times \cdots \times \mathbb{P}V_n)$ is not r -defective. Let $C_{i,j}$ be the quadric hypersurface in $V_{i,j}$ defined by $\{v \in V_{i,j} | v^\top v = 0\}$. Let U be the open subset of M consisting of $p = (\otimes_{i=1}^n u_{i,1}, \dots, \otimes_{i=1}^n u_{i,r})$ such that each $u_{i,j} \notin C_{i,j}$, and

$$(3.2) \quad \dim \left(\sum_{i,j} \left(\bigotimes_{k \neq i} u_{k,j} \right) \otimes V_{i,j} \right) = r \sum_{i=1}^n d_i - r(n-1).$$

Lemma 3.3. *For $p \in U$, any $x_{i,j} \in V_{i,j}$ and $[y_{i,j}] \in V_{i,j}/\langle u_{i,j} \rangle$,*

1. *There is some $A \in V_1 \otimes \cdots \otimes V_n$ such that $\text{Con}(A - \tilde{A}, \otimes_{k \neq i} u_{k,j}) = x_{i,j}$ if and only if $u_{1,j}^\top x_{1,j} = \cdots = u_{n,j}^\top x_{n,j}$ for all j .*

2. *There is some $A \in V_1 \otimes \cdots \otimes V_n$ such that $[\text{Con}(A - \tilde{A}, \otimes_{k \neq i} u_{k,j})] = [y_{i,j}]$.*

Proof. 1. If $\text{Con}(A - \tilde{A}, \otimes_{k \neq i} u_{k,j}) = x_{i,j}$, then $\text{Con}(A - \tilde{A}, \otimes_{k=1}^n u_{k,j}) = u_{i,j}^\top x_{i,j}$.

Conversely, the linear system $\text{Con}(A, \otimes_{k \neq i} u_{k,j}) = \text{Con}(\tilde{A}, \otimes_{k \neq i} u_{k,j}) + x_{i,j}$ is solvable if and only if the coefficient matrices $[\otimes_{k \neq i} u_{k,j}]$ and $[\otimes_{k \neq i} u_{k,j}, \text{Con}(\tilde{A}, \otimes_{k \neq i} u_{k,j}) + x_{i,j}]$ have the same rank. By (3.2), $\text{rank}[\otimes_{k \neq i} u_{k,j}] = r \sum_{i=1}^n d_i - r(n-1)$, and $u_{i,j} \otimes (\otimes_{k \neq i} u_{k,j}) = u_{i,j} \otimes (\otimes_{k \neq i} u_{k,j})$ are the only linear relations in $[\otimes_{k \neq i} u_{k,j}]$. So if $u_{1,j}^\top x_{1,j} = \cdots = u_{n,j}^\top x_{n,j}$, then $\text{rank}[\otimes_{k \neq i} u_{k,j}, \text{Con}(\tilde{A}, \otimes_{k \neq i} u_{k,j}) + x_{i,j}] \leq r \sum_{i=1}^n d_i - r(n-1)$.

2. $\text{Con}(A - \tilde{A}, \otimes_{k=2}^n u_{k,j}) = x_{1,j}$ and $[\text{Con}(A - \tilde{A}, \otimes_{k \neq i} u_{k,j})] = [y_{i,j}]$ are solvable if and only if $u_{1,j}^\top x_{1,j} = u_{2,j}^\top (y_{2,j} + t_{2,j} u_{2,j}) = \cdots = u_{n,j}^\top (y_{n,j} + t_{n,j} u_{n,j})$ for some $t_{i,j}$. Since $u_{i,j} \notin C_{i,j}$, let $t_{i,j} = \frac{u_{1,j}^\top x_{1,j} - u_{i,j}^\top y_{i,j}}{u_{i,j}^\top u_{i,j}}$.

□

$$\text{Let } S = \{s \in H^0(M, E) | s = (\bigoplus_j \text{Con}(\otimes_i u_{i,j}, \otimes_{k \neq 1} u_{k,j}), \bigoplus_{i,j} [\text{Con}(A - \tilde{A}, \otimes_{k \neq i} u_{k,j})])\}$$

Lemma 3.4. *The induced map $T_\nu : T_p U \times T_s S_U \rightarrow T_{s(p)} E$ is surjective.*

Lemma 3.5. *For $\alpha = \{(i_1, j_1), \dots, (i_l, j_l) | 1 \leq i_1, \dots, i_l \leq n, 1 \leq j_1, \dots, j_l \leq r\}$, let $F_\alpha = X_{1,1} \times \cdots \times X_{n,r}$, where $X_{i,j} = \mathbb{P}(C_{i,j})$ if $(i, j) \in \alpha$ and $X_{i,j} = \mathbb{P}V_{i,j}$ otherwise. For $p \in M|_{F_\alpha}$,*

1. If $\alpha \subset \{(i_1, 1), \dots, (i_r, r)\}$, there is some A such that $[Con(A - \tilde{A}, \otimes_{k \neq i} u_{k,j})] = [y_{i,j}]$
2. Otherwise, there is some A such that $[Con(A - \tilde{A}, \otimes_{k \neq i} u_{k,j})] = [y_{i,j}]$ if and only if $u_{i_1,j_1}^\top y_{i_1,j_1} = \dots = u_{i_l,j_l}^\top y_{i_l,j_l}$.

Let $R = r \sum_{i=1}^n d_i - r(n-1)$ and $D_q = \{p \in M \mid \dim \left(\sum_{i,j} (\otimes_{k \neq i} u_{k,j}) \otimes V_{i,j} \right) \leq R - q\}$, then D_q is defined by $(R - q) \times (R - q)$ minors. Each $p \in D_q$ satisfies $r(n-1) + q$ linear relations $\sum_{i=1}^r (\otimes_{k \neq i} u_{k,j}) \otimes v_{i,j} = 0$, where $v_{i,j} \in V_{i,j}$ is algebraic in $u_{i,j}$. $v_{i,j} = u_{i,j}$ consists of $r(n-1)$ of these linear equations, and there are q nontrivial linear relations for $p \in M|_{D_q \setminus D_{q-1}}$.

Lemma 3.6. For $p \in M|_{D_q \setminus \cup_\alpha F_\alpha}$,

1. $Con(A - \tilde{A}, \otimes_{k \neq i} u_{k,j}) = x_{i,j}$ if and only if $u_{1,j}^\top x_{1,j} = \dots = u_{n,j}^\top x_{n,j}$ and $\sum_{i,j} x_{i,j}^\top v_{i,j}^l = 0$ for $1 \leq l \leq q$.
2. $[Con(A - \tilde{A}, \otimes_{k \neq i} u_{k,j})] = [y_{i,j}]$ if and only if

$$(3.3) \quad u_{1,j}^\top (y_{1,j} + t_{1,j} u_{1,j}) = \dots = u_{n,j}^\top (y_{n,j} + t_{n,j} u_{n,j}),$$

$$(3.4) \quad \sum_{i,j} (y_{i,j} + t_{i,j} u_{i,j})^\top v_{i,j}^l = 0,$$

for $1 \leq l \leq q$. For each p , the linear subspace consisting of $\{[y_{i,j}]\}$ satisfying (3.4) is independent of the choice of $v_{i,j}$.

Construct E_α over $M|_{F_\alpha}$ to be all linear transformations $y_{i,j}$ such that $u_{i_1,j_1}^\top y_{i_1,j_1} = \dots = u_{i_l,j_l}^\top y_{i_l,j_l}$, then $\text{rank}(E_\alpha) > \dim M|_{F_\alpha}$.

Construct E_q over D_q to be all linear transformations $x_{i,j}$ such that (3.4) holds, then $\text{rank} E_q > \dim D_q$.

Construct $E_{\alpha,q}$ over $M|_{F_\alpha} \cap D_q$ to be all transformations such that $u_{i_1,j_1}^\top y_{i_1,j_1} = \dots = u_{i_l,j_l}^\top y_{i_l,j_l}$ and (3.4) holds, then $\text{rank} E_{\alpha,q} > \dim M|_{F_\alpha} \cap D_q$.

So for a generic σ , Z_σ is contained in U , since Z_σ is an affine algebraic variety, Z_σ consists of finite number of points.

REFERENCES

- [1] L.-H. Lim, P. Comon, *Nonnegative approximations of nonnegative tensors*, Journal of Chemometrics, 23 (2009), no. 7-8, pp. 432-441
- [2] S. Friedland, G. Ottaviani, *The Number of Singular Vector Tuples and Uniqueness of Best Rank-One Approximation of Tensors*, Foundations of Computational Mathematics, (2014)
- [3] L.-H. Lim, *Singular values and eigenvalues of tensors: a variational approach*, Proceedings of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP '05), 1 (2005), pp. 129-132
- [4] R. Hartshorne, *Algebraic Geometry*, Graduate Texts in Mathematics, vol. 52, Springer, New York (1977)
- [5] V. D. Silva, L.-H. Lim, *Tensor rank and the ill-posedness of the best low-rank approximation problem*, SIAM J. Matrix Anal. Appl., Vol. 30, No. 33, pp. 1084-1127 (2008)
- [6] S. Friedland, M. Stawiska, *Best approximation on semi-algebraic sets and k-border rank approximation of symmetric tensors*, arXiv:1311.1561
- [7] R. Bott, L. W. Tu, *Differential Forms in Algebraic Topology*, Graduate Texts in Mathematics, vol. 82, Springer, New York (1982)
- [8] K. C. Chang, K. Pearson, T. Zhang, *Perron-Frobenius theorem for nonnegative tensors*, Commun. Math. Sci., vol. 6, No. 2, pp. 507-520 (2008)

- [9] S. Friedland, S. Gaubert, L. Han, *Perron-Frobenius theorem for nonnegative multilinear forms and extensions*, Linear Algebra and its Applications, 438 (2013), pp. 738-749
- [10] A. Stegeman and P. Comon *Subtracting a best rank-1 approximation does not necessarily decrease tensor rank*, Linear Algebra Appl., pp. 1276-1300, No. 7, vol. 433 (2010)