Results and problems for 3-tensors

Shmuel Friedland
Univ. Illinois at Chicago

NIU LA'09 , 14 August, 2009

Overview

Overview

1. The rank of 3-tensors.

Overview

1. The rank of 3-tensors.
2. Best rank one approximation.

Overview

1. The rank of 3-tensors.
2. Best rank one approximation.
3. Perron-Frobenius theorem for *irreducible* nonnegative tensors.

Overview

1. The rank of 3-tensors.
2. Best rank one approximation.
3. Perron-Frobenius theorem for *irreducible* nonnegative tensors.
4. Diagonal scaling of nonnegative tensors to tensors with given rows, columns and depth sums.

Overview

1. The rank of 3-tensors.
2. Best rank one approximation.
3. Perron-Frobenius theorem for *irreducible* nonnegative tensors.
4. Diagonal scaling of nonnegative tensors to tensors with given rows, columns and depth sums.
5. Analogs of SVD decomposition of 3-tensors.

Overview

1. The rank of 3-tensors.
2. Best rank one approximation.
3. Perron-Frobenius theorem for *irreducible* nonnegative tensors.
4. Diagonal scaling of nonnegative tensors to tensors with given rows, columns and depth sums.
5. Analogs of SVD decomposition of 3-tensors.
6. CUR decompositions for tensors

Rank of tensor

Rank of tensor

$$\mathbb{F}^{m\times n\times l} := \{\mathcal{T} = [t_{i,j,k}]_{i=j=k}^{m,n,l}, t_{i,j,k} \in \mathbb{F}\}.$$

Rank of tensor

$\mathbb{F}^{m\times n\times l} := \{\mathcal{T} = [t_{i,j,k}]_{i=j=k}^{m,n,l}, t_{i,j,k} \in \mathbb{F}\}$.

rank one tensor: $\mathbf{x} \otimes \mathbf{y} \otimes \mathbf{z} = [x_i y_j z_k] = (a\mathbf{x}) \otimes (b\mathbf{y})((ab)^{-1}\mathbf{z})$

Rank of tensor

$\mathbb{F}^{m\times n\times l} := \{\mathcal{T} = [t_{i,j,k}]_{i=j=k}^{m,n,l}, t_{i,j,k} \in \mathbb{F}\}$.

rank one tensor: $\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z} = [x_i y_j z_k] = (a\mathbf{x})\otimes(b\mathbf{y})((ab)^{-1}\mathbf{z})$

rank $\mathcal{T}$ minimal r:
$\mathcal{T} = f_r(\mathbf{x}_1,\mathbf{y}_1,\mathbf{z}_1,\ldots,\mathbf{x}_r,\mathbf{y}_r,\mathbf{z}_r) := \sum_{i=1}^r \mathbf{x}_i\otimes\mathbf{y}_i\otimes\mathbf{z}_i$,
$\mathbf{x}_i \in \mathbb{F}^m, \mathbf{y}_i \in \mathbb{F}^n, \mathbf{z}_i \in \mathbb{F}^l$

Rank of tensor

$\mathbb{F}^{m\times n\times l} := \{\mathcal{T} = [t_{i,j,k}]_{i=j=k}^{m,n,l}, t_{i,j,k} \in \mathbb{F}\}$.

rank one tensor: $\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z} = [x_i y_j z_k] = (a\mathbf{x})\otimes(b\mathbf{y})((ab)^{-1}\mathbf{z})$

$\mathrm{rank}\ \mathcal{T}$ minimal r:
$\mathcal{T} = f_r(\mathbf{x}_1, \mathbf{y}_1, \mathbf{z}_1, \ldots, \mathbf{x}_r, \mathbf{y}_r, \mathbf{z}_r) := \sum_{i=1}^{r} \mathbf{x}_i \otimes \mathbf{y}_i \otimes \mathbf{z}_i$,
$\mathbf{x}_i \in \mathbb{F}^m, \mathbf{y}_i \in \mathbb{F}^n, \mathbf{z}_i \in \mathbb{F}^l$

THM Let $\mathcal{T} = [t_{i,j,k}] \in \mathbb{F}^{m\times n\times l}$. $T_k := [t_{i,j,k}]_{i,j=1}^{m,n}, k = 1, \ldots, l$. Then $\mathrm{rank}\ \mathcal{T}$ is the minimal dimension of subspace spanned by rank one matrices containing span $(T_1, \ldots, T_l)$.

Rank of tensor

$\mathbb{F}^{m\times n\times l} := \{\mathcal{T} = [t_{i,j,k}]_{i=j=k}^{m,n,l}, t_{i,j,k} \in \mathbb{F}\}$.

rank one tensor: $\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z} = [x_i y_j z_k] = (a\mathbf{x})\otimes(b\mathbf{y})((ab)^{-1}\mathbf{z})$

$\mathrm{rank}\ \mathcal{T}$ minimal r:
$\mathcal{T} = f_r(\mathbf{x}_1,\mathbf{y}_1,\mathbf{z}_1,\ldots,\mathbf{x}_r,\mathbf{y}_r,\mathbf{z}_r) := \sum_{i=1}^{r}\mathbf{x}_i\otimes\mathbf{y}_i\otimes\mathbf{z}_i$,
$\mathbf{x}_i \in \mathbb{F}^m, \mathbf{y}_i \in \mathbb{F}^n, \mathbf{z}_i \in \mathbb{F}^l$

THM Let $\mathcal{T} = [t_{i,j,k}] \in \mathbb{F}^{m\times n\times l}$. $T_k := [t_{i,j,k}]_{i,j=1}^{m,n}, k = 1,\ldots,l$. Then $\mathrm{rank}\ \mathcal{T}$ is the minimal dimension of subspace spanned by rank one matrices containing span $(T_1,\ldots,T_l)$.

Cor For $\mathcal{T} \in \mathbb{F}^{m\times n\times l}$ and $l \geq mn$ $\mathrm{rank}_{\mathbb{F}}\mathcal{T} \leq mn$ (sharp).

Rank of tensor

$\mathbb{F}^{m\times n\times l} := \{\mathcal{T} = [t_{i,j,k}]_{i=j=k}^{m,n,l}, t_{i,j,k} \in \mathbb{F}\}$.

rank one tensor: $\mathbf{x} \otimes \mathbf{y} \otimes \mathbf{z} = [x_i y_j z_k] = (a\mathbf{x}) \otimes (b\mathbf{y})((ab)^{-1}\mathbf{z})$

$\mathrm{rank}\ \mathcal{T}$ minimal r:
$\mathcal{T} = f_r(\mathbf{x}_1, \mathbf{y}_1, \mathbf{z}_1, \ldots, \mathbf{x}_r, \mathbf{y}_r, \mathbf{z}_r) := \sum_{i=1}^{r} \mathbf{x}_i \otimes \mathbf{y}_i \otimes \mathbf{z}_i$,
$\mathbf{x}_i \in \mathbb{F}^m, \mathbf{y}_i \in \mathbb{F}^n, \mathbf{z}_i \in \mathbb{F}^l$

THM Let $\mathcal{T} = [t_{i,j,k}] \in \mathbb{F}^{m\times n\times l}$. $T_k := [t_{i,j,k}]_{i,j=1}^{m,n}, k = 1, \ldots, l$. Then $\mathrm{rank}\ \mathcal{T}$ is the minimal dimension of subspace spanned by rank one matrices containing span $(T_1, \ldots, T_l)$.

Cor For $\mathcal{T} \in \mathbb{F}^{m\times n\times l}$ and $l \ge mn$ $\mathrm{rank}\ _{\mathbb{F}}\mathcal{T} \le mn$ (sharp).

Normalization: $2 \le m \le n \le l \le mn$

Generic rank I

Generic rank I

generic rank: $\mathrm{grank}_{\mathbb{F}}(m,n,l)$ - the rank of a random tensor $\mathcal{T} \in \mathbb{F}^{m\times n\times l}$

Generic rank I

generic rank: $\mathrm{grank}_{\mathbb{F}}(m,n,l)$ - the rank of a random tensor $\mathcal{T} \in \mathbb{F}^{m\times n\times l}$

Thm: $\mathrm{grank}_{\mathbb{C}}(m,n,l) = \min(l, mn)$ for $(m-1)(n-1)+1 \le l$.

Generic rank I

generic rank: $\mathrm{grank}_{\mathbb{F}}(m,n,l)$ - the rank of a random tensor $\mathcal{T} \in \mathbb{F}^{m\times n\times l}$

Thm: $\mathrm{grank}_{\mathbb{C}}(m,n,l) = \min(l, mn)$ for $(m-1)(n-1)+1 \le l$.

PRF: Generic subspace $L \subset \mathbb{C}^{m\times n}$ of dimension $l = (m-1)(n-1)+1$ intersects $\mathcal{R}(n,m,1,\mathbb{C})$ at $\binom{m+n-2}{m-1}$, $(\ge l)$, s.t. any l are linearly independent

Generic rank I

generic rank: $\mathrm{grank}_{\mathbb{F}}(m, n, l)$ - the rank of a random tensor $\mathcal{T} \in \mathbb{F}^{m\times n\times l}$

Thm: $\mathrm{grank}_{\mathbb{C}}(m, n, l) = \min(l, mn)$ for $(m-1)(n-1)+1 \le l$.

PRF: Generic subspace $L \subset \mathbb{C}^{m\times n}$ of dimension $l = (m-1)(n-1)+1$ intersects $\mathcal{R}(n, m, 1, \mathbb{C})$ at $\binom{m+n-2}{m-1}$, $(\ge l)$, s.t. any l are linearly independent

Any generic subspace $L \subset \mathbb{C}^{m\times n}$ of dimension $l \ge (m-1)(n-1)+1$ spanned by rank one matrices

Generic rank I

generic rank: $\mathrm{grank}_{\mathbb{F}}(m,n,l)$ - the rank of a random tensor $\mathcal{T} \in \mathbb{F}^{m\times n\times l}$

Thm: $\mathrm{grank}_{\mathbb{C}}(m,n,l) = \min(l, mn)$ for $(m-1)(n-1)+1 \le l$.

PRF: Generic subspace $L \subset \mathbb{C}^{m\times n}$ of dimension $l = (m-1)(n-1)+1$ intersects $\mathcal{R}(n,m,1,\mathbb{C})$ at $\binom{m+n-2}{m-1}$, $(\ge l)$, s.t. any l are linearly independent

Any generic subspace $L \subset \mathbb{C}^{m\times n}$ of dimension $l \ge (m-1)(n-1)+1$ spanned by rank one matrices

Dimension count for $\mathbb{F} = \mathbb{C}$ and $2 \le m \le n \le l \le (m-1)(n-1)$:

Generic rank I

generic rank: $\mathrm{grank}_{\mathbb{F}}(m,n,l)$ - the rank of a random tensor $\mathcal{T} \in \mathbb{F}^{m\times n\times l}$

Thm: $\mathrm{grank}_{\mathbb{C}}(m,n,l) = \min(l, mn)$ for $(m-1)(n-1)+1 \le l$.

PRF: Generic subspace $L \subset \mathbb{C}^{m\times n}$ of dimension $l = (m-1)(n-1)+1$ intersects $\mathcal{R}(n,m,1,\mathbb{C})$ at $\binom{m+n-2}{m-1}$, $(\ge l)$, s.t. any l are linearly independent

Any generic subspace $L \subset \mathbb{C}^{m\times n}$ of dimension $l \ge (m-1)(n-1)+1$ spanned by rank one matrices

Dimension count for $\mathbb{F} = \mathbb{C}$ and $2 \le m \le n \le l \le (m-1)(n-1)$:
$f_r : (\mathbb{C}^m \times \mathbb{C}^n \times \mathbb{C}^l)^r \to \mathbb{C}^{m\times n\times l}$

Generic rank I

generic rank: $\mathrm{grank}_{\mathbb{F}}(m,n,l)$ - the rank of a random tensor $\mathcal{T} \in \mathbb{F}^{m\times n\times l}$

Thm: $\mathrm{grank}_{\mathbb{C}}(m,n,l) = \min(l,mn)$ for $(m-1)(n-1)+1 \le l$.

PRF: Generic subspace $L \subset \mathbb{C}^{m\times n}$ of dimension $l = (m-1)(n-1)+1$ intersects $\mathcal{R}(n,m,1,\mathbb{C})$ at $\binom{m+n-2}{m-1}$, $(\ge l)$, s.t. any l are linearly independent

Any generic subspace $L \subset \mathbb{C}^{m\times n}$ of dimension $l \ge (m-1)(n-1)+1$ spanned by rank one matrices

Dimension count for $\mathbb{F} = \mathbb{C}$ and $2 \le m \le n \le l \le (m-1)(n-1)$:
$f_r : (\mathbb{C}^m \times \mathbb{C}^n \times \mathbb{C}^l)^r \to \mathbb{C}^{m\times n\times l}$
$\mathrm{grank}_{\mathbb{C}}(m,n,l)(m+n+l-2) \ge mnl \Rightarrow \mathrm{grank}_{\mathbb{C}}(m,n,l) \ge \lceil \frac{mnl}{(m+n+l-2)} \rceil$

Generic rank I

generic rank: $\mathrm{grank}_{\mathbb{F}}(m,n,l)$ - the rank of a random tensor $\mathcal{T} \in \mathbb{F}^{m\times n\times l}$

Thm: $\mathrm{grank}_{\mathbb{C}}(m,n,l) = \min(l,mn)$ for $(m-1)(n-1)+1 \le l$.

PRF: Generic subspace $L \subset \mathbb{C}^{m\times n}$ of dimension $l = (m-1)(n-1)+1$ intersects $\mathcal{R}(n,m,1,\mathbb{C})$ at $\binom{m+n-2}{m-1}$, $(\ge l)$, s.t. any l are linearly independent

Any generic subspace $L \subset \mathbb{C}^{m\times n}$ of dimension $l \ge (m-1)(n-1)+1$ spanned by rank one matrices

Dimension count for $\mathbb{F} = \mathbb{C}$ and $2 \le m \le n \le l \le (m-1)(n-1)$:
$f_r : (\mathbb{C}^m \times \mathbb{C}^n \times \mathbb{C}^l)^r \to \mathbb{C}^{m\times n\times l}$
$\mathrm{grank}_{\mathbb{C}}(m,n,l)(m+n+l-2) \ge mnl \Rightarrow \mathrm{grank}_{\mathbb{C}}(m,n,l) \ge \lceil \frac{mnl}{(m+n+l-2)} \rceil$
$\frac{mnl}{(m+n+l-2)} = l$ for $l = (m-1)(n-1)+1$

Generic rank I

generic rank: $\mathrm{grank}_{\mathbb{F}}(m,n,l)$ - the rank of a random tensor $\mathcal{T} \in \mathbb{F}^{m\times n\times l}$

Thm: $\mathrm{grank}_{\mathbb{C}}(m,n,l) = \min(l,mn)$ for $(m-1)(n-1)+1 \le l$.

PRF: Generic subspace $L \subset \mathbb{C}^{m\times n}$ of dimension $l = (m-1)(n-1)+1$ intersects $\mathcal{R}(n,m,1,\mathbb{C})$ at $\binom{m+n-2}{m-1}$, $(\ge l)$, s.t. any l are linearly independent

Any generic subspace $L \subset \mathbb{C}^{m\times n}$ of dimension $l \ge (m-1)(n-1)+1$ spanned by rank one matrices

Dimension count for $\mathbb{F} = \mathbb{C}$ and $2 \le m \le n \le l \le (m-1)(n-1)$:
$f_r : (\mathbb{C}^m \times \mathbb{C}^n \times \mathbb{C}^l)^r \to \mathbb{C}^{m\times n\times l}$
$\mathrm{grank}_{\mathbb{C}}(m,n,l)(m+n+l-2) \ge mnl \Rightarrow \mathrm{grank}_{\mathbb{C}}(m,n,l) \ge \lceil \frac{mnl}{(m+n+l-2)} \rceil$
$\frac{mnl}{(m+n+l-2)} = l$ for $l = (m-1)(n-1)+1$
COR $\mathrm{grank}_{\mathbb{C}}(m,n,(m-1)(n-1)) = (m-1)(n-1)+1$.

Generic rank II

Conjecture $\mathrm{grank}_{\mathbb{C}}(m,n,l) = \lceil \frac{mnl}{(m+n+l-2)} \rceil$
for $2 \le m \le n \le l < (m-1)(n-1)$ and $(3,n,l) \ne (3,2p+1,2p+1)$

Generic rank II

Conjecture $\mathrm{grank}_{\mathbb{C}}(m,n,l) = \lceil \frac{mnl}{(m+n+l-2)} \rceil$
for $2 \le m \le n \le l < (m-1)(n-1)$ and $(3,n,l) \neq (3,2p+1,2p+1)$

Fact: $\mathrm{grank}_{\mathbb{C}}(3,2p+1,2p+1) = \lceil \frac{3(2p+1)^2}{4p+3} \rceil + 1$

Generic rank II

Conjecture $\mathrm{grank}_{\mathbb{C}}(m,n,l) = \lceil \frac{mnl}{(m+n+l-2)} \rceil$
for $2 \le m \le n \le l < (m-1)(n-1)$ and $(3,n,l) \ne (3,2p+1,2p+1)$

Fact: $\mathrm{grank}_{\mathbb{C}}(3,2p+1,2p+1) = \lceil \frac{3(2p+1)^2}{4p+3} \rceil + 1$

Conjecture is known in some cases

Generic rank II

Conjecture $\mathrm{grank}_{\mathbb{C}}(m, n, l) = \lceil \frac{mnl}{(m+n+l-2)} \rceil$
for $2 \le m \le n \le l < (m-1)(n-1)$ and $(3, n, l) \ne (3, 2p+1, 2p+1)$

Fact: $\mathrm{grank}_{\mathbb{C}}(3, 2p+1, 2p+1) = \lceil \frac{3(2p+1)^2}{4p+3} \rceil + 1$

Conjecture is known in some cases
For symmetric 3-tensors: Alexander-Hirschowitz Theorem

Generic rank II

Conjecture $\mathrm{grank}_{\mathbb{C}}(m,n,l) = \lceil \frac{mnl}{(m+n+l-2)} \rceil$
for $2 \le m \le n \le l < (m-1)(n-1)$ and $(3,n,l) \ne (3,2p+1,2p+1)$

Fact: $\mathrm{grank}_{\mathbb{C}}(3,2p+1,2p+1) = \lceil \frac{3(2p+1)^2}{4p+3} \rceil + 1$

Conjecture is known in some cases
For symmetric 3-tensors: Alexander-Hirschowitz Theorem

Easy to compute $\mathrm{grank}_{\mathbb{C}}(m,n,l)$:
Pick at random $\mathbf{w}_r := (\mathbf{x}_1,\mathbf{y}_1,\mathbf{z}_1,\ldots,\mathbf{x}_r,\mathbf{y}_r,\mathbf{z}_r) \in (\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^l)^r$
The minimal $r \ge \lceil \frac{mnl}{(m+n+l-2)} \rceil$ s.t. $\mathrm{rank}\, J(f_r)(\mathbf{w}_r) = mnl$
is $\mathrm{grank}_{\mathbb{C}}(m,n,l)$

Generic rank II

Conjecture $\mathrm{grank}_{\mathbb{C}}(m,n,l) = \lceil \frac{mnl}{(m+n+l-2)} \rceil$
for $2 \le m \le n \le l < (m-1)(n-1)$ and $(3,n,l) \ne (3,2p+1,2p+1)$

Fact: $\mathrm{grank}_{\mathbb{C}}(3,2p+1,2p+1) = \lceil \frac{3(2p+1)^2}{4p+3} \rceil + 1$

Conjecture is known in some cases
For symmetric 3-tensors: Alexander-Hirschowitz Theorem

Easy to compute $\mathrm{grank}_{\mathbb{C}}(m,n,l)$:
Pick at random $\mathbf{w}_r := (\mathbf{x}_1,\mathbf{y}_1,\mathbf{z}_1,\ldots,\mathbf{x}_r,\mathbf{y}_r,\mathbf{z}_r) \in (\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^l)^r$
The minimal $r \ge \lceil \frac{mnl}{(m+n+l-2)} \rceil$ s.t. $\mathrm{rank}\, J(f_r)(\mathbf{w}_r) = mnl$
is $\mathrm{grank}_{\mathbb{C}}(m,n,l)$

Avoid round-off error:
$\mathbf{w}_r \in (\mathbb{Z}^m \times \mathbb{Z}^n \times \mathbb{Z}^l)^r$ find $\mathrm{rank}\, J(f_r)(\mathbf{w}_r)$ exact arithmetic

Generic rank II

Conjecture $\mathrm{grank}_{\mathbb{C}}(m,n,l) = \lceil \frac{mnl}{(m+n+l-2)} \rceil$
for $2 \le m \le n \le l < (m-1)(n-1)$ and $(3,n,l) \ne (3,2p+1,2p+1)$

Fact: $\mathrm{grank}_{\mathbb{C}}(3,2p+1,2p+1) = \lceil \frac{3(2p+1)^2}{4p+3} \rceil + 1$

Conjecture is known in some cases
For symmetric 3-tensors: Alexander-Hirschowitz Theorem

Easy to compute $\mathrm{grank}_{\mathbb{C}}(m,n,l)$:
Pick at random $\mathbf{w}_r := (\mathbf{x}_1,\mathbf{y}_1,\mathbf{z}_1,\ldots,\mathbf{x}_r,\mathbf{y}_r,\mathbf{z}_r) \in (\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^l)^r$
The minimal $r \ge \lceil \frac{mnl}{(m+n+l-2)} \rceil$ s.t. $\mathrm{rank}\, J(f_r)(\mathbf{w}_r) = mnl$
is $\mathrm{grank}_{\mathbb{C}}(m,n,l)$

Avoid round-off error:
$\mathbf{w}_r \in (\mathbb{Z}^m \times \mathbb{Z}^n \times \mathbb{Z}^l)^r$ find $\mathrm{rank}\, J(f_r)(\mathbf{w}_r)$ exact arithmetic
I checked the conjecture up to $m,n,l \le 14$

Generic rank III - the real case

For $mn \leq l$ $\mathrm{grank}_{\mathbb{R}}(m,n,l) = mn$.

Generic rank III - the real case

For $mn \le l$ $\mathrm{grank}_{\mathbb{R}}(m,n,l) = mn$.

For $2 \le m \le n \le l < mn - 1$, there exist $V_1, \ldots, V_{c(m,n,l)} \subset \mathbb{R}^{m\times n\times l}$ pairwise disjoint open connected semi-algebraic sets s.t.
Closure$(\cup_{i=1}^{c(m,n,l)} V_i) = \mathbb{R}^{m\times n\times l}$
$\mathrm{rank}\ \mathcal{T} = \mathrm{grank}_{\mathbb{C}}(m,n,l)$ for each $\mathcal{T} \in V_1$
$\mathrm{rank}\ \mathcal{T} = \rho_i$ for each $\mathcal{T} \in V_i$
$\rho_i \ge \mathrm{grank}_{\mathbb{C}}(m,n,l)$ for $i = 2, \ldots, c(m,n,l)$

Generic rank III - the real case

For $mn \le l$ $\mathrm{grank}_{\mathbb{R}}(m,n,l) = mn$.

For $2 \le m \le n \le l < mn - 1$, there exist $V_1, \ldots, V_{c(m,n,l)} \subset \mathbb{R}^{m\times n\times l}$ pairwise disjoint open connected semi-algebraic sets s.t.
Closure$(\cup_{i=1}^{c(m,n,l)} V_i) = \mathbb{R}^{m\times n\times l}$
$\mathrm{rank}\ \mathcal{T} = \mathrm{grank}_{\mathbb{C}}(m,n,l)$ for each $\mathcal{T} \in V_1$
$\mathrm{rank}\ \mathcal{T} = \rho_i$ for each $\mathcal{T} \in V_i$
$\rho_i \ge \mathrm{grank}_{\mathbb{C}}(m,n,l)$ for $i = 2, \ldots, c(m,n,l)$

For $l = (m-1)(n-1) + 1$ $\exists m, n$:
$c(m,n,l) > 1, \rho_{c(m,n,l)} \ge \mathrm{grank}_{\mathbb{C}}(m,n,l) + 1$

Generic rank III - the real case

For $mn \leq l$ $\mathrm{grank}_{\mathbb{R}}(m,n,l) = mn$.

For $2 \leq m \leq n \leq l < mn - 1$, there exist $V_1, \ldots, V_{c(m,n,l)} \subset \mathbb{R}^{m\times n\times l}$ pairwise disjoint open connected semi-algebraic sets s.t.
Closure$(\cup_{i=1}^{c(m,n,l)} V_i) = \mathbb{R}^{m\times n\times l}$
$\mathrm{rank}\ \mathcal{T} = \mathrm{grank}_{\mathbb{C}}(m,n,l)$ for each $\mathcal{T} \in V_1$
$\mathrm{rank}\ \mathcal{T} = \rho_i$ for each $\mathcal{T} \in V_i$
$\rho_i \geq \mathrm{grank}_{\mathbb{C}}(m,n,l)$ for $i = 2, \ldots, c(m,n,l)$

For $l = (m-1)(n-1) + 1$ $\exists m,n$:
$c(m,n,l) > 1, \rho_{c(m,n,l)} \geq \mathrm{grank}_{\mathbb{C}}(m,n,l) + 1$

Examples [2]
$m = n \geq 2,\ l = (m-1)(n-1) + 1.$
$m = n = 4,\ l = 11, 12$

Generic rank III - the real case

For $mn \leq l$ $\mathrm{grank}_{\mathbb{R}}(m,n,l) = mn$.

For $2 \leq m \leq n \leq l < mn - 1$, there exist $V_1, \ldots, V_{c(m,n,l)} \subset \mathbb{R}^{m\times n\times l}$ pairwise disjoint open connected semi-algebraic sets s.t.
$\mathrm{Closure}(\cup_{i=1}^{c(m,n,l)} V_i) = \mathbb{R}^{m\times n\times l}$
$\mathrm{rank}\ \mathcal{T} = \mathrm{grank}_{\mathbb{C}}(m,n,l)$ for each $\mathcal{T} \in V_1$
$\mathrm{rank}\ \mathcal{T} = \rho_i$ for each $\mathcal{T} \in V_i$
$\rho_i \geq \mathrm{grank}_{\mathbb{C}}(m,n,l)$ for $i = 2, \ldots, c(m,n,l)$

For $l = (m-1)(n-1)+1$ $\exists m,n$:
$c(m,n,l) > 1, \rho_{c(m,n,l)} \geq \mathrm{grank}_{\mathbb{C}}(m,n,l) + 1$

Examples [2]
$m = n \geq 2,\ l = (m-1)(n-1)+1.$
$m = n = 4,\ l = 11, 12$
Problem: $\rho_i \leq \mathrm{grank}_{\mathbb{C}}(m,n,l) + 1$?

Upper bounds for generic and maximal rank

Upper bounds for generic and maximal rank

$L \cap \mathcal{R}(m,n,k,\mathbb{C}) \supsetneq \{0\}$ if codim $L <$ dim $\mathcal{R}(m,n,k,\mathbb{C}) = k(n+m-k)$.

Upper bounds for generic and maximal rank

$L \cap \mathcal{R}(m,n,k,\mathbb{C}) \supsetneq \{0\}$ if codim $L <$ dim $\mathcal{R}(m,n,k,\mathbb{C}) = k(n+m-k)$.

$\mathrm{grank}(n,m,m) \le \lfloor \frac{n}{2} \rfloor m + (n - 2\lfloor \frac{n}{2} \rfloor)(m - \lfloor \sqrt{n-1} \rfloor)$ if $m \ge 2\lfloor \sqrt{n-1} \rfloor$

Upper bounds for generic and maximal rank

$L \cap \mathcal{R}(m,n,k,\mathbb{C}) \supsetneq \{0\}$ if codim $L <$ dim $\mathcal{R}(m,n,k,\mathbb{C}) = k(n+m-k)$.

$\mathrm{grank}(n,m,m) \le \lfloor \frac{n}{2} \rfloor m + (n - 2\lfloor \frac{n}{2} \rfloor)(m - \lfloor \sqrt{n-1} \rfloor)$ if $m \ge 2\lfloor \sqrt{n-1} \rfloor$

$\mathrm{grank}(n,m,m) \le n(m - \lfloor \sqrt{n-1} \rfloor)$ if $m < 2\lfloor \sqrt{n-1} \rfloor < 2(m-1)$,

Upper bounds for generic and maximal rank

$L \cap \mathcal{R}(m,n,k,\mathbb{C}) \supsetneq \{0\}$ if codim $L < \dim \mathcal{R}(m,n,k,\mathbb{C}) = k(n+m-k)$.

$\mathrm{grank}(n,m,m) \le \lfloor \frac{n}{2} \rfloor m + (n - 2\lfloor \frac{n}{2} \rfloor)(m - \lfloor \sqrt{n-1} \rfloor)$ if $m \ge 2\lfloor \sqrt{n-1} \rfloor$

$\mathrm{grank}(n,m,m) \le n(m - \lfloor \sqrt{n-1} \rfloor)$ if $m < 2\lfloor \sqrt{n-1} \rfloor < 2(m-1)$,

$\mathrm{grank}(n,m,m) = \min(n, m^2)$ if $n \ge (m-1)^2 + 1$,

Upper bounds for generic and maximal rank

$L \cap \mathcal{R}(m,n,k,\mathbb{C}) \supsetneqq \{0\}$ if codim $L <$ dim $\mathcal{R}(m,n,k,\mathbb{C}) = k(n+m-k)$.

$\mathrm{grank}(n,m,m) \le \lfloor \frac{n}{2} \rfloor m + (n - 2\lfloor \frac{n}{2} \rfloor)(m - \lfloor \sqrt{n-1} \rfloor)$ if $m \ge 2\lfloor \sqrt{n-1} \rfloor$

$\mathrm{grank}(n,m,m) \le n(m - \lfloor \sqrt{n-1} \rfloor)$ if $m < 2\lfloor \sqrt{n-1} \rfloor < 2(m-1)$,

$\mathrm{grank}(n,m,m) = \min(n,m^2)$ if $n \ge (m-1)^2 + 1$,

$\mathrm{mrank}(n,m,m) \le$
$\sum_{i=1}^{\lfloor \sqrt{n-1} \rfloor} (2i-1)(m-i+1) + (m - \lfloor \sqrt{n-1} \rfloor^2)(m - \lfloor \sqrt{n-1} \rfloor)$

Rank one approximations

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A},\mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k}b_{i,j,k}, \ \|\mathcal{T}\| = \sqrt{\langle \mathcal{T},\mathcal{T}\rangle}$

$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z},\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle = (\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

Rank one approximations

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A},\mathcal{B}\rangle=\sum_{i=j=k}^{m,n,l} a_{i,j,k}b_{i,j,k}$, $\|\mathcal{T}\|=\sqrt{\langle\mathcal{T},\mathcal{T}\rangle}$

$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z},\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle=(\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1,\ldots,\mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$

$\mathrm{P}_{\mathbf{X}}(\mathcal{T})=\sum_{i=1}^{d}\langle\mathcal{T},\mathcal{X}_i\rangle\mathcal{X}_i,\quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2=\sum_{i=1}^{d}\langle\mathcal{T},\mathcal{X}_i\rangle^2$

Rank one approximations

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A},\mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k}b_{i,j,k},\ \ \|\mathcal{T}\| = \sqrt{\langle \mathcal{T},\mathcal{T}\rangle}$

$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z},\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle = (\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1,\ldots,\mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$

$\mathrm{P}_{\mathbf{X}}(\mathcal{T}) = \sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle\mathcal{X}_i,\quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2 = \sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle^2$

$\|\mathcal{T}\|^2 = \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2 + \|\mathcal{T} - \mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2$

Rank one approximations

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A}, \mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k} b_{i,j,k}, \ \|\mathcal{T}\| = \sqrt{\langle \mathcal{T}, \mathcal{T}\rangle}$

$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}, \mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle = (\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1, \ldots, \mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$

$\mathrm{P}_{\mathbf{X}}(\mathcal{T}) = \sum_{i=1}^{d} \langle \mathcal{T}, \mathcal{X}_i\rangle \mathcal{X}_i, \quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2 = \sum_{i=1}^{d} \langle \mathcal{T}, \mathcal{X}_i\rangle^2$

$\|\mathcal{T}\|^2 = \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2 + \|\mathcal{T} - \mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2$

Best rank one approximation of $\mathcal{T}$:

Rank one approximations

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A},\mathcal{B}\rangle=\sum_{i=j=k}^{m,n,l} a_{i,j,k}b_{i,j,k},\ \|\mathcal{T}\|=\sqrt{\langle\mathcal{T},\mathcal{T}\rangle}$

$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z},\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle=(\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1,\ldots,\mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$

$\mathrm{P}_{\mathbf{X}}(\mathcal{T})=\sum_{i=1}^{d}\langle\mathcal{T},\mathcal{X}_i\rangle\mathcal{X}_i,\quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2=\sum_{i=1}^{d}\langle\mathcal{T},\mathcal{X}_i\rangle^2$

$\|\mathcal{T}\|^2=\|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2+\|\mathcal{T}-\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2$

Best rank one approximation of $\mathcal{T}$:

$\min_{\mathbf{x},\mathbf{y},\mathbf{z}}\|\mathcal{T}-\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|=\min_{\|\mathbf{x}\|=\|\mathbf{y}\|=\|\mathbf{z}\|=1,a}\|\mathcal{T}-a\,\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|$

Rank one approximations

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A},\mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k}b_{i,j,k}, \ \|\mathcal{T}\| = \sqrt{\langle \mathcal{T},\mathcal{T}\rangle}$

$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z},\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle = (\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1,\ldots,\mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$

$\mathrm{P}_{\mathbf{X}}(\mathcal{T}) = \sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle\mathcal{X}_i, \quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2 = \sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle^2$

$\|\mathcal{T}\|^2 = \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2 + \|\mathcal{T}-\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2$

Best rank one approximation of $\mathcal{T}$:

$\min_{\mathbf{x},\mathbf{y},\mathbf{z}} \|\mathcal{T}-\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\| = \min_{\|\mathbf{x}\|=\|\mathbf{y}\|=\|\mathbf{z}\|=1,a} \|\mathcal{T}-a\,\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|$

Equivalent: $\max_{\|\mathbf{x}\|=\|\mathbf{y}\|=\|\mathbf{z}\|=1} \sum_{i=j=k}^{m,n,l} t_{i,j,k}x_iy_jz_k$

Rank one approximations

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A},\mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k}b_{i,j,k},\ \|\mathcal{T}\| = \sqrt{\langle \mathcal{T},\mathcal{T}\rangle}$
$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z},\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle = (\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1,\dots,\mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$
$\mathrm{P}_{\mathbf{X}}(\mathcal{T}) = \sum_{i=1}^d \langle \mathcal{T},\mathcal{X}_i\rangle\mathcal{X}_i, \quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2 = \sum_{i=1}^d \langle \mathcal{T},\mathcal{X}_i\rangle^2$

$\|\mathcal{T}\|^2 = \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2 + \|\mathcal{T}-\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2$

Best rank one approximation of $\mathcal{T}$:
$\min_{\mathbf{x},\mathbf{y},\mathbf{z}} \|\mathcal{T}-\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\| = \min_{\|\mathbf{x}\|=\|\mathbf{y}\|=\|\mathbf{z}\|=1,a} \|\mathcal{T}-a\,\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|$

Equivalent: $\max_{\|\mathbf{x}\|=\|\mathbf{y}\|=\|\mathbf{z}\|=1} \sum_{i=j=k}^{m,n,l} t_{i,j,k}x_iy_jz_k$

Lagrange multipliers: $\mathcal{T}\times\mathbf{y}\otimes\mathbf{z} := \sum_{j=k=1} t_{i,j,k}y_jz_k = \lambda\mathbf{x}$
$\mathcal{T}\times\mathbf{x}\otimes\mathbf{z} = \lambda\mathbf{y},\ \mathcal{T}\times\mathbf{x}\otimes\mathbf{y} = \lambda\mathbf{z}$

Rank one approximations

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A},\mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k}b_{i,j,k}, \ \|\mathcal{T}\| = \sqrt{\langle \mathcal{T},\mathcal{T}\rangle}$
$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z},\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle = (\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1,\ldots,\mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$
$\mathrm{P}_{\mathbf{X}}(\mathcal{T}) = \sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle\mathcal{X}_i, \quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2 = \sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle^2$

$\|\mathcal{T}\|^2 = \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2 + \|\mathcal{T}-\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2$

Best rank one approximation of $\mathcal{T}$:
$\min_{\mathbf{x},\mathbf{y},\mathbf{z}}\|\mathcal{T}-\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\| = \min_{\|\mathbf{x}\|=\|\mathbf{y}\|=\|\mathbf{z}\|=1,a}\|\mathcal{T}-a\,\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|$

Equivalent: $\max_{\|\mathbf{x}\|=\|\mathbf{y}\|=\|\mathbf{z}\|=1}\sum_{i=j=k}^{m,n,l} t_{i,j,k}x_iy_jz_k$

Lagrange multipliers: $\mathcal{T}\times\mathbf{y}\otimes\mathbf{z} := \sum_{j=k=1} t_{i,j,k}y_jz_k = \lambda\mathbf{x}$
$\mathcal{T}\times\mathbf{x}\otimes\mathbf{z} = \lambda\mathbf{y},\ \mathcal{T}\times\mathbf{x}\otimes\mathbf{y} = \lambda\mathbf{z}$
λ singular value, $\mathbf{x},\mathbf{y},\mathbf{z}$ singular vectors

Rank one approximations

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A},\mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k}b_{i,j,k},\ \|\mathcal{T}\| = \sqrt{\langle \mathcal{T},\mathcal{T}\rangle}$
$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z},\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle = (\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1,\ldots,\mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$
$\mathrm{P}_{\mathbf{X}}(\mathcal{T}) = \sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle\mathcal{X}_i, \quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2 = \sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle^2$

$\|\mathcal{T}\|^2 = \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2 + \|\mathcal{T}-\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|^2$

Best rank one approximation of $\mathcal{T}$:
$\min_{\mathbf{x},\mathbf{y},\mathbf{z}} \|\mathcal{T}-\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\| = \min_{\|\mathbf{x}\|=\|\mathbf{y}\|=\|\mathbf{z}\|=1,a} \|\mathcal{T}-a\,\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|$

Equivalent: $\max_{\|\mathbf{x}\|=\|\mathbf{y}\|=\|\mathbf{z}\|=1} \sum_{i=j=k}^{m,n,l} t_{i,j,k}x_iy_jz_k$

Lagrange multipliers: $\mathcal{T}\times\mathbf{y}\otimes\mathbf{z} := \sum_{j=k=1} t_{i,j,k}y_jz_k = \lambda\mathbf{x}$
$\mathcal{T}\times\mathbf{x}\otimes\mathbf{z} = \lambda\mathbf{y},\ \mathcal{T}\times\mathbf{x}\otimes\mathbf{y} = \lambda\mathbf{z}$
λ singular value, $\mathbf{x},\mathbf{y},\mathbf{z}$ singular vectors
How many distinct singular values are for a generic tensor?

ℓ_p maximal problem and Perron-Frobenius

ℓ_p maximal problem and Perron-Frobenius

$\|(x_1, \ldots, x_n)^\top\|_p := (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$

Problem: $\max_{\|\mathbf{x}\|_p=\|\mathbf{y}\|_p=\|\mathbf{z}\|_p=1} \sum_{i=j=k}^{m,n,l} t_{i,j,k} x_i y_j z_k$

Lagrange multipliers: $\mathcal{T} \times \mathbf{y} \otimes \mathbf{z} := \sum_{j=k=1} t_{i,j,k} y_j z_k = \lambda \mathbf{x}^{p-1}$
$\mathcal{T} \times \mathbf{x} \otimes \mathbf{z} = \lambda \mathbf{y}^{p-1}$, $\mathcal{T} \times \mathbf{x} \otimes \mathbf{y} = \lambda \mathbf{z}^{p-1}$ ($p = \frac{2t}{2s-1}, t, s \in \mathbb{N}$)

ℓ_p maximal problem and Perron-Frobenius

$\|(x_1,\ldots,x_n)^\top\|_p := (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$

Problem: $\max_{\|\mathbf{x}\|_p=\|\mathbf{y}\|_p=\|\mathbf{z}\|_p=1} \sum_{i=j=k}^{m,n,l} t_{i,j,k} x_i y_j z_k$

Lagrange multipliers: $\mathcal{T} \times \mathbf{y} \otimes \mathbf{z} := \sum_{j=k=1} t_{i,j,k} y_j z_k = \lambda \mathbf{x}^{p-1}$
$\mathcal{T} \times \mathbf{x} \otimes \mathbf{z} = \lambda \mathbf{y}^{p-1}$, $\mathcal{T} \times \mathbf{x} \otimes \mathbf{y} = \lambda \mathbf{z}^{p-1}$ $(p = \frac{2t}{2s-1}, t, s \in \mathbb{N})$

$p = 3$ is most natural in view of homogeneity

ℓ_p maximal problem and Perron-Frobenius

$\|(x_1,\ldots,x_n)^\top\|_p := (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$

Problem: $\max_{\|\mathbf{x}\|_p=\|\mathbf{y}\|_p=\|\mathbf{z}\|_p=1} \sum_{i=j=k}^{m,n,l} t_{i,j,k}x_iy_jz_k$

Lagrange multipliers: $\mathcal{T}\times\mathbf{y}\otimes\mathbf{z} := \sum_{j=k=1} t_{i,j,k}y_jz_k = \lambda\mathbf{x}^{p-1}$
$\mathcal{T}\times\mathbf{x}\otimes\mathbf{z} = \lambda\mathbf{y}^{p-1}$, $\mathcal{T}\times\mathbf{x}\otimes\mathbf{y} = \lambda\mathbf{z}^{p-1}$ $(p = \frac{2t}{2s-1}, t, s \in \mathbb{N})$

$p = 3$ is most natural in view of homogeneity

Assume that $\mathcal{T} \geq 0$. Then $\mathbf{x},\mathbf{y},\mathbf{z} \geq 0$

For which values of p we have an analog of Perron-Frobenius theorem?

Yes, for $p \geq 3$, No, for $p < 3$,
Friedland-Gauber-Han [3]

Outline of the proof

Outline of the proof

Define: $F : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^l$:

Outline of the proof

Define: $F : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^l$:

$$F((\mathbf{x}, \mathbf{y}, \mathbf{z}))_{i,1} = \left(\|\mathbf{x}\|_p^{p-3} \sum_{j=k=1}^{n,l} f_{i,j,k} y_j z_k \right)^{\frac{1}{p-1}}, i = 1, \ldots, m$$

$$F((\mathbf{x}, \mathbf{y}, \mathbf{z}))_{j,2} = \left(\|\mathbf{y}\|_p^{p-3} \sum_{i=k=1}^{m,l} f_{i,j,k} x_i z_k \right)^{\frac{1}{p-1}}, j = 1, \ldots, n$$

$$F((\mathbf{x}, \mathbf{y}, \mathbf{z}))_{k,3} = \left(\|\mathbf{z}\|_p^{p-3} \sum_{i=j=1}^{m,n} f_{i,j,k} x_i y_j \right)^{\frac{1}{p-1}}, k = 1, \ldots, l$$

Outline of the proof

Define: $F : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^l$:

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{i,1} = \left(\|\mathbf{x}\|_p^{p-3} \sum_{j=k=1}^{n,l} f_{i,j,k} y_j z_k\right)^{\frac{1}{p-1}}, i = 1,\dots,m$$

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{j,2} = \left(\|\mathbf{y}\|_p^{p-3} \sum_{i=k=1}^{m,l} f_{i,j,k} x_i z_k\right)^{\frac{1}{p-1}}, j = 1,\dots,n$$

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{k,3} = \left(\|\mathbf{z}\|_p^{p-3} \sum_{i=j=1}^{m,n} f_{i,j,k} x_i y_j\right)^{\frac{1}{p-1}}, k = 1,\dots,l$$

Assume $\sum_{j=k=1}^{n,l} f_{i,j,k} > 0, i = 1,\dots,m,$
$\sum_{i=k=1}^{m,l} f_{i,j,k} > 0, j = 1,\dots,n,$ $\sum_{i=j=1}^{m,n} f_{i,j,k} > 0, k = 1,\dots,l$

Outline of the proof

Define: $F : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^l$:

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{i,1} = \left(\|\mathbf{x}\|_p^{p-3} \sum_{j=k=1}^{n,l} f_{i,j,k} y_j z_k\right)^{\frac{1}{p-1}}, i = 1,\dots,m$$

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{j,2} = \left(\|\mathbf{y}\|_p^{p-3} \sum_{i=k=1}^{m,l} f_{i,j,k} x_i z_k\right)^{\frac{1}{p-1}}, j = 1,\dots,n$$

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{k,3} = \left(\|\mathbf{z}\|_p^{p-3} \sum_{i=j=1}^{m,n} f_{i,j,k} x_i y_j\right)^{\frac{1}{p-1}}, k = 1,\dots,l$$

Assume $\sum_{j=k=1}^{n,l} f_{i,j,k} > 0, i = 1,\dots,m$,
$\sum_{i=k=1}^{m,l} f_{i,j,k} > 0, j = 1,\dots,n$, $\sum_{i=j=1}^{m,n} f_{i,j,k} > 0, k = 1,\dots,l$

F 1-homogeneous monotone, maps open positive cone $\mathbb{R}_+^m \times \mathbb{R}_+^n \times \mathbb{R}_+^l$ to itself.

$\mathcal{F} = [f_{i,j,k}]$ induces tri-partite graph on $\langle m\rangle, \langle n\rangle, \langle l\rangle$:
$i \in \langle m\rangle$ connected to $j \in \langle n\rangle$ and $k \in \langle l\rangle$ iff $f_{i,j,k} > 0$, sim. for j, k

Outline of the proof

Define: $F:\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}^l\to\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}^l$:

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{i,1}=\left(\|\mathbf{x}\|_p^{p-3}\sum_{j=k=1}^{n,l}f_{i,j,k}y_jz_k\right)^{\frac{1}{p-1}},i=1,\ldots,m$$

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{j,2}=\left(\|\mathbf{y}\|_p^{p-3}\sum_{i=k=1}^{m,l}f_{i,j,k}x_iz_k\right)^{\frac{1}{p-1}},j=1,\ldots,n$$

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{k,3}=\left(\|\mathbf{z}\|_p^{p-3}\sum_{i=j=1}^{m,n}f_{i,j,k}x_iy_j\right)^{\frac{1}{p-1}},k=1,\ldots,l$$

Assume $\sum_{j=k=1}^{n,l}f_{i,j,k}>0,i=1,\ldots,m$,
$\sum_{i=k=1}^{m,l}f_{i,j,k}>0,j=1,\ldots,n$, $\sum_{i=j=1}^{m,n}f_{i,j,k}>0,k=1,\ldots,l$

F 1-homogeneous monotone, maps open positive cone $\mathbb{R}_+^m\times\mathbb{R}_+^n\times\mathbb{R}_+^l$ to itself.

$\mathcal{F}=[f_{i,j,k}]$ induces tri-partite graph on $\langle m\rangle,\langle n\rangle,\langle l\rangle$:
$i\in\langle m\rangle$ connected to $j\in\langle n\rangle$ and $k\in\langle l\rangle$ iff $f_{i,j,k}>0$, sim. for j,k

If tri-partite graph is connected then F has unique positive eigenvector

Outline of the proof

Define: $F:\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}^l\to\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}^l$:

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{i,1}=\left(\|\mathbf{x}\|_p^{p-3}\sum_{j=k=1}^{n,l} f_{i,j,k}y_jz_k\right)^{\frac{1}{p-1}},i=1,\ldots,m$$

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{j,2}=\left(\|\mathbf{y}\|_p^{p-3}\sum_{i=k=1}^{m,l} f_{i,j,k}x_iz_k\right)^{\frac{1}{p-1}},j=1,\ldots,n$$

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{k,3}=\left(\|\mathbf{z}\|_p^{p-3}\sum_{i=j=1}^{m,n} f_{i,j,k}x_iy_j\right)^{\frac{1}{p-1}},k=1,\ldots,l$$

Assume $\sum_{j=k=1}^{n,l} f_{i,j,k}>0, i=1,\ldots,m$,
$\sum_{i=k=1}^{m,l} f_{i,j,k}>0, j=1,\ldots,n$, $\sum_{i=j=1}^{m,n} f_{i,j,k}>0, k=1,\ldots,l$

F 1-homogeneous monotone, maps open positive cone $\mathbb{R}_+^m\times\mathbb{R}_+^n\times\mathbb{R}_+^l$ to itself.

$\mathcal{F}=[f_{i,j,k}]$ induces tri-partite graph on $\langle m\rangle,\langle n\rangle,\langle l\rangle$:
$i\in\langle m\rangle$ connected to $j\in\langle n\rangle$ and $k\in\langle l\rangle$ iff $f_{i,j,k}>0$, sim. for j,k

If tri-partite graph is connected then F has unique positive eigenvector

If F completely irreducible, i.e. F^N maps nonzero nonnegative vectors to positive, nonnegative eigenvector is unique and positive

Outline of the proof

Define: $F : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^l$:

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{i,1} = \left(\|\mathbf{x}\|_p^{p-3} \sum_{j=k=1}^{n,l} f_{i,j,k} y_j z_k\right)^{\frac{1}{p-1}}, i = 1,\ldots,m$$

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{j,2} = \left(\|\mathbf{y}\|_p^{p-3} \sum_{i=k=1}^{m,l} f_{i,j,k} x_i z_k\right)^{\frac{1}{p-1}}, j = 1,\ldots,n$$

$$F((\mathbf{x},\mathbf{y},\mathbf{z}))_{k,3} = \left(\|\mathbf{z}\|_p^{p-3} \sum_{i=j=1}^{m,n} f_{i,j,k} x_i y_j\right)^{\frac{1}{p-1}}, k = 1,\ldots,l$$

Assume $\sum_{j=k=1}^{n,l} f_{i,j,k} > 0, i = 1,\ldots,m$,
$\sum_{i=k=1}^{m,l} f_{i,j,k} > 0, j = 1,\ldots,n$, $\sum_{i=j=1}^{m,n} f_{i,j,k} > 0, k = 1,\ldots,l$

F 1-homogeneous monotone, maps open positive cone $\mathbb{R}^m_+ \times \mathbb{R}^n_+ \times \mathbb{R}^l_+$ to itself.

$\mathcal{F} = [f_{i,j,k}]$ induces tri-partite graph on $\langle m\rangle, \langle n\rangle, \langle l\rangle$:
$i \in \langle m\rangle$ connected to $j \in \langle n\rangle$ and $k \in \langle l\rangle$ iff $f_{i,j,k} > 0$, sim. for j, k

If tri-partite graph is connected then F has unique positive eigenvector

If F completely irreducible, i.e. F^N maps nonzero nonnegative vectors to positive, nonnegative eigenvector is unique and positive

$p < 3$ numerical counterexamples $m = n = l = 2$

Scaling of nonnegative tensors to tensors with given row, column and depth sums

$0 \le \mathcal{T} = [t_{i,j,k}] \in \mathbb{R}^{m\times n\times l}$ has given row, column and depth sums:
$\mathbf{r} = (r_1, \dots, r_m)^\top, \mathbf{c} = (c_1, \dots, c_n)^\top, \mathbf{d} = (d_1, \dots, d_l)^\top > \mathbf{0}$:

Scaling of nonnegative tensors to tensors with given row, column and depth sums

$0 \leq \mathcal{T} = [t_{i,j,k}] \in \mathbb{R}^{m\times n\times l}$ has given row, column and depth sums:
$\mathbf{r} = (r_1, \dots, r_m)^\top, \mathbf{c} = (c_1, \dots, c_n)^\top, \mathbf{d} = (d_1, \dots, d_l)^\top > \mathbf{0}$:
$\sum_{j,k} t_{i,j,k} = r_i > 0,\ \sum_{i,k} t_{i,j,k} = c_j > 0, \sum_{i,j} t_{i,j,k} = d_k > 0$
$\sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j = \sum_{k=1}^{l} d_k$

Scaling of nonnegative tensors to tensors with given row, column and depth sums

$0 \le \mathcal{T} = [t_{i,j,k}] \in \mathbb{R}^{m\times n\times l}$ has given row, column and depth sums:
$\mathbf{r} = (r_1, \ldots, r_m)^\top, \mathbf{c} = (c_1, \ldots, c_n)^\top, \mathbf{d} = (d_1, \ldots, d_l)^\top > \mathbf{0}$:
$\sum_{j,k} t_{i,j,k} = r_i > 0,\ \sum_{i,k} t_{i,j,k} = c_j > 0, \sum_{i,j} t_{i,j,k} = d_k > 0$
$\sum_{i=1}^m r_i = \sum_{j=1}^n c_j = \sum_{k=1}^l d_k$

Find nec. and suf. conditions for scaling:
$\mathcal{T}' = [t_{i,j,k} e^{x_i+y_j+z_k}], \mathbf{x}, \mathbf{y}, \mathbf{z}$ such that $\mathcal{T}'$ has given row, column and depth sum

Scaling of nonnegative tensors to tensors with given row, column and depth sums

$0 \le \mathcal{T} = [t_{i,j,k}] \in \mathbb{R}^{m\times n\times l}$ has given row, column and depth sums:
$\mathbf{r} = (r_1, \ldots, r_m)^\top, \mathbf{c} = (c_1, \ldots, c_n)^\top, \mathbf{d} = (d_1, \ldots, d_l)^\top > \mathbf{0}$:
$\sum_{j,k} t_{i,j,k} = r_i > 0, \ \sum_{i,k} t_{i,j,k} = c_j > 0, \sum_{i,j} t_{i,j,k} = d_k > 0$
$\sum_{i=1}^m r_i = \sum_{j=1}^n c_j = \sum_{k=1}^l d_k$

Find nec. and suf. conditions for scaling:
$\mathcal{T}' = [t_{i,j,k} e^{x_i+y_j+z_k}], \mathbf{x}, \mathbf{y}, \mathbf{z}$ such that $\mathcal{T}'$ has given row, column and depth sum
Solution: Convert to the minimal problem:
$\min_{\mathbf{r}^\top \mathbf{x} = \mathbf{c}^\top \mathbf{y} = \mathbf{d}^\top \mathbf{z} = 0} f_{\mathcal{T}}(\mathbf{x}, \mathbf{y}, \mathbf{z}), \quad f_{\mathcal{T}}(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{i,j,k} t_{i,j,k} e^{x_i+y_j+z_k}$

Scaling of nonnegative tensors to tensors with given row, column and depth sums

$0 \le \mathcal{T} = [t_{i,j,k}] \in \mathbb{R}^{m\times n\times l}$ has given row, column and depth sums:
$\mathbf{r} = (r_1, \ldots, r_m)^\top, \mathbf{c} = (c_1, \ldots, c_n)^\top, \mathbf{d} = (d_1, \ldots, d_l)^\top > \mathbf{0}$:
$\sum_{j,k} t_{i,j,k} = r_i > 0,\ \sum_{i,k} t_{i,j,k} = c_j > 0, \sum_{i,j} t_{i,j,k} = d_k > 0$
$\sum_{i=1}^m r_i = \sum_{j=1}^n c_j = \sum_{k=1}^l d_k$

Find nec. and suf. conditions for scaling:
$\mathcal{T}' = [t_{i,j,k} e^{x_i+y_j+z_k}], \mathbf{x}, \mathbf{y}, \mathbf{z}$ such that $\mathcal{T}'$ has given row, column and depth sum
Solution: Convert to the minimal problem:
$\min_{\mathbf{r}^\top\mathbf{x}=\mathbf{c}^\top\mathbf{y}=\mathbf{d}^\top\mathbf{z}=0} f_{\mathcal{T}}(\mathbf{x}, \mathbf{y}, \mathbf{z}), \quad f_{\mathcal{T}}(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{i,j,k} t_{i,j,k} e^{x_i+y_j+z_k}$

Any critical point of $f_{\mathcal{T}}$ on $\mathcal{S} := \{\mathbf{r}^\top\mathbf{x} = \mathbf{c}^\top\mathbf{y} = \mathbf{d}^\top\mathbf{z} = 0\}$ gives rise to a solution of the scaling problem (Lagrange multipliers)

Scaling of nonnegative tensors to tensors with given row, column and depth sums

$0 \le \mathcal{T} = [t_{i,j,k}] \in \mathbb{R}^{m\times n\times l}$ has given row, column and depth sums:
$\mathbf{r} = (r_1, \ldots, r_m)^\top, \mathbf{c} = (c_1, \ldots, c_n)^\top, \mathbf{d} = (d_1, \ldots, d_l)^\top > \mathbf{0}$:
$\sum_{j,k} t_{i,j,k} = r_i > 0, \ \sum_{i,k} t_{i,j,k} = c_j > 0, \sum_{i,j} t_{i,j,k} = d_k > 0$
$\sum_{i=1}^m r_i = \sum_{j=1}^n c_j = \sum_{k=1}^l d_k$

Find nec. and suf. conditions for scaling:
$\mathcal{T}' = [t_{i,j,k} e^{x_i+y_j+z_k}], \mathbf{x}, \mathbf{y}, \mathbf{z}$ such that $\mathcal{T}'$ has given row, column and depth sum
Solution: Convert to the minimal problem:
$\min_{\mathbf{r}^\top \mathbf{x} = \mathbf{c}^\top \mathbf{y} = \mathbf{d}^\top \mathbf{z} = 0} f_{\mathcal{T}}(\mathbf{x}, \mathbf{y}, \mathbf{z}), \quad f_{\mathcal{T}}(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{i,j,k} t_{i,j,k} e^{x_i+y_j+z_k}$

Any critical point of $f_{\mathcal{T}}$ on $\mathcal{S} := \{\mathbf{r}^\top \mathbf{x} = \mathbf{c}^\top \mathbf{y} = \mathbf{d}^\top \mathbf{z} = 0\}$ gives rise to a solution of the scaling problem (Lagrange multipliers)
$f_{\mathcal{T}}$ is convex

Scaling of nonnegative tensors to tensors with given row, column and depth sums

$0 \le \mathcal{T} = [t_{i,j,k}] \in \mathbb{R}^{m\times n\times l}$ has given row, column and depth sums:
$\mathbf{r} = (r_1,\dots,r_m)^\top, \mathbf{c} = (c_1,\dots,c_n)^\top, \mathbf{d} = (d_1,\dots,d_l)^\top > \mathbf{0}$:
$\sum_{j,k} t_{i,j,k} = r_i > 0,\ \sum_{i,k} t_{i,j,k} = c_j > 0, \sum_{i,j} t_{i,j,k} = d_k > 0$
$\sum_{i=1}^m r_i = \sum_{j=1}^n c_j = \sum_{k=1}^l d_k$

Find nec. and suf. conditions for scaling:
$\mathcal{T}' = [t_{i,j,k} e^{x_i+y_j+z_k}], \mathbf{x}, \mathbf{y}, \mathbf{z}$ such that $\mathcal{T}'$ has given row, column and depth sum
Solution: Convert to the minimal problem:
$\min_{\mathbf{r}^\top\mathbf{x}=\mathbf{c}^\top\mathbf{y}=\mathbf{d}^\top\mathbf{z}=0} f_{\mathcal{T}}(\mathbf{x},\mathbf{y},\mathbf{z}), \quad f_{\mathcal{T}}(\mathbf{x},\mathbf{y},\mathbf{z}) = \sum_{i,j,k} t_{i,j,k} e^{x_i+y_j+z_k}$

Any critical point of $f_{\mathcal{T}}$ on $\mathcal{S} := \{\mathbf{r}^\top\mathbf{x} = \mathbf{c}^\top\mathbf{y} = \mathbf{d}^\top\mathbf{z} = 0\}$ gives rise to a solution of the scaling problem (Lagrange multipliers)
$f_{\mathcal{T}}$ is convex
$f_{\mathcal{T}}$ is strictly convex implies $\mathcal{T}$ is not decomposable: $\mathcal{T} \ne \mathcal{T}_1 \oplus \mathcal{T}_2$.

Scaling of nonnegative tensors II

Scaling of nonnegative tensors II

if $f_{\mathcal{T}}$ is strictly convex and is ∞ on $\partial\mathcal{S}$, $f_{\mathcal{T}}$ achieves its unique minimum

Scaling of nonnegative tensors II

if $f_{\mathcal{T}}$ is strictly convex and is ∞ on $\partial\mathcal{S}$, $f_{\mathcal{T}}$ achieves its unique minimum

Equivalent to: the inequalities $x_i + y_j + z_k \le 0$ if $t_{i,j,k} > 0$ and equalities $\mathbf{r}^\top \mathbf{x} = \mathbf{c}^\top \mathbf{y} = \mathbf{d}^\top \mathbf{z} = 0$ imply $\mathbf{x} = \mathbf{0}_m, \mathbf{y} = \mathbf{0}_n, \mathbf{z} = \mathbf{0}_l$.

Scaling of nonnegative tensors II

if $f_{\mathcal{T}}$ is strictly convex and is ∞ on $\partial\mathcal{S}$, $f_{\mathcal{T}}$ achieves its unique minimum

Equivalent to: the inequalities $x_i + y_j + z_k \leq 0$ if $t_{i,j,k} > 0$ and equalities $\mathbf{r}^\top \mathbf{x} = \mathbf{c}^\top \mathbf{y} = \mathbf{d}^\top \mathbf{z} = 0$ imply $\mathbf{x} = \mathbf{0}_m, \mathbf{y} = \mathbf{0}_n, \mathbf{z} = \mathbf{0}_l$.

Fact: For $\mathbf{r} = \mathbf{1}_m, \mathbf{c} = \mathbf{1}_n, \mathbf{d} = \mathbf{1}_l$ Sinkhorn scaling algorithm works.

Scaling of nonnegative tensors II

if $f_{\mathcal{T}}$ is strictly convex and is ∞ on $\partial\mathcal{S}$, $f_{\mathcal{T}}$ achieves its unique minimum

Equivalent to: the inequalities $x_i + y_j + z_k \le 0$ if $t_{i,j,k} > 0$ and equalities $\mathbf{r}^\top \mathbf{x} = \mathbf{c}^\top \mathbf{y} = \mathbf{d}^\top \mathbf{z} = 0$ imply $\mathbf{x} = \mathbf{0}_m, \mathbf{y} = \mathbf{0}_n, \mathbf{z} = \mathbf{0}_l$.

Fact: For $\mathbf{r} = \mathbf{1}_m, \mathbf{c} = \mathbf{1}_n, \mathbf{d} = \mathbf{1}_l$ Sinkhorn scaling algorithm works.

Newton method works, since the scaling problem is equivalent finding the unique minimum of strict convex function

Scaling of nonnegative tensors II

if $f_{\mathcal{T}}$ is strictly convex and is ∞ on $\partial\mathcal{S}$, $f_{\mathcal{T}}$ achieves its unique minimum

Equivalent to: the inequalities $x_i + y_j + z_k \le 0$ if $t_{i,j,k} > 0$ and equalities $\mathbf{r}^\top \mathbf{x} = \mathbf{c}^\top \mathbf{y} = \mathbf{d}^\top \mathbf{z} = 0$ imply $\mathbf{x} = \mathbf{0}_m, \mathbf{y} = \mathbf{0}_n, \mathbf{z} = \mathbf{0}_l$.

Fact: For $\mathbf{r} = \mathbf{1}_m, \mathbf{c} = \mathbf{1}_n, \mathbf{d} = \mathbf{1}_l$ Sinkhorn scaling algorithm works.

Newton method works, since the scaling problem is equivalent finding the unique minimum of strict convex function

Hence Newton method has a quadratic convergence versus linear convergence of Sinkhorn algorithm

Scaling of nonnegative tensors II

if $f_{\mathcal{T}}$ is strictly convex and is ∞ on $\partial\mathcal{S}$, $f_{\mathcal{T}}$ achieves its unique minimum

Equivalent to: the inequalities $x_i + y_j + z_k \le 0$ if $t_{i,j,k} > 0$ and equalities $\mathbf{r}^\top \mathbf{x} = \mathbf{c}^\top \mathbf{y} = \mathbf{d}^\top \mathbf{z} = 0$ imply $\mathbf{x} = \mathbf{0}_m, \mathbf{y} = \mathbf{0}_n, \mathbf{z} = \mathbf{0}_l$.

Fact: For $\mathbf{r} = \mathbf{1}_m, \mathbf{c} = \mathbf{1}_n, \mathbf{d} = \mathbf{1}_l$ Sinkhorn scaling algorithm works.

Newton method works, since the scaling problem is equivalent finding the unique minimum of strict convex function

Hence Newton method has a quadratic convergence versus linear convergence of Sinkhorn algorithm
True for matrices too

Scaling of nonnegative tensors II

if $f_{\mathcal{T}}$ is strictly convex and is ∞ on $\partial\mathcal{S}$, $f_{\mathcal{T}}$ achieves its unique minimum

Equivalent to: the inequalities $x_i + y_j + z_k \le 0$ if $t_{i,j,k} > 0$ and equalities $\mathbf{r}^\top \mathbf{x} = \mathbf{c}^\top \mathbf{y} = \mathbf{d}^\top \mathbf{z} = 0$ imply $\mathbf{x} = \mathbf{0}_m, \mathbf{y} = \mathbf{0}_n, \mathbf{z} = \mathbf{0}_l$.

Fact: For $\mathbf{r} = \mathbf{1}_m, \mathbf{c} = \mathbf{1}_n, \mathbf{d} = \mathbf{1}_l$ Sinkhorn scaling algorithm works.

Newton method works, since the scaling problem is equivalent finding the unique minimum of strict convex function

Hence Newton method has a quadratic convergence versus linear convergence of Sinkhorn algorithm
True for matrices too

Are variants of Menon and Brualdi theorems hold in the tensor case?

Scaling of nonnegative tensors II

if $f_{\mathcal{T}}$ is strictly convex and is ∞ on $\partial\mathcal{S}$, $f_{\mathcal{T}}$ achieves its unique minimum

Equivalent to: the inequalities $x_i + y_j + z_k \le 0$ if $t_{i,j,k} > 0$ and equalities $\mathbf{r}^\top \mathbf{x} = \mathbf{c}^\top \mathbf{y} = \mathbf{d}^\top \mathbf{z} = 0$ imply $\mathbf{x} = \mathbf{0}_m, \mathbf{y} = \mathbf{0}_n, \mathbf{z} = \mathbf{0}_l$.

Fact: For $\mathbf{r} = \mathbf{1}_m, \mathbf{c} = \mathbf{1}_n, \mathbf{d} = \mathbf{1}_l$ Sinkhorn scaling algorithm works.

Newton method works, since the scaling problem is equivalent finding the unique minimum of strict convex function

Hence Newton method has a quadratic convergence versus linear convergence of Sinkhorn algorithm
True for matrices too

Are variants of Menon and Brualdi theorems hold in the tensor case?
Yes for Menon, unknown for Brualdi

Analogs of SVD decomposition: I

Analogs of SVD decomposition: I

Unfolding tensor $\mathcal{T} = [t_{i,j,k}] \in \mathbb{F}^{m \times n \times l}$ in 1-index: viewing as a matrix $A_1 \in \mathbb{F}^{m \times (nl)}$.

Analogs of SVD decomposition: I

Unfolding tensor $\mathcal{T} = [t_{i,j,k}] \in \mathbb{F}^{m\times n\times l}$ in 1-index: viewing as a matrix $A_1 \in \mathbb{F}^{m\times(nl)}$.

For a cubic tensor $\mathcal{T} \in \mathbb{R}^{n\times n\times n}$

do orthonormal change of coordinates in each three components $\mathbb{R}^n$:

$\mathcal{T}_1 = \mathcal{T} \times_1 Q_1 \times_2 Q_2 \times_3 Q_3$

Analogs of SVD decomposition: I

Unfolding tensor $\mathcal{T} = [t_{i,j,k}] \in \mathbb{F}^{m\times n\times l}$ in 1-index: viewing as a matrix $A_1 \in \mathbb{F}^{m\times (nl)}$.
For a cubic tensor $\mathcal{T} \in \mathbb{R}^{n\times n\times n}$
do orthonormal change of coordinates in each three components $\mathbb{R}^n$:
$\mathcal{T}_1 = \mathcal{T} \times_1 Q_1 \times_2 Q_2 \times_3 Q_3$

One can have at most $\frac{3n(n-1)}{2}$ zero entries in $\mathcal{T}_1$: $t_{j,i,i} = t_{i,j,i} = t_{i,i,j}$ for $i < j$.

Analogs of SVD decomposition: I

Unfolding tensor $\mathcal{T} = [t_{i,j,k}] \in \mathbb{F}^{m \times n \times l}$ in 1-index: viewing as a matrix $A_1 \in \mathbb{F}^{m \times (nl)}$.
For a cubic tensor $\mathcal{T} \in \mathbb{R}^{n \times n \times n}$
do orthonormal change of coordinates in each three components $\mathbb{R}^n$:
$\mathcal{T}_1 = \mathcal{T} \times_1 Q_1 \times_2 Q_2 \times_3 Q_3$

One can have at most $\frac{3n(n-1)}{2}$ zero entries in $\mathcal{T}_1$: $t_{j,i,i} = t_{i,j,i} = t_{i,i,j}$ for $i < j$.

Apply QR algorithm to the n columns of the unfolded matrix in mode k indexed by $(j,j), j = 1, \ldots, n$ for $k = 1, 2, 3, 1, 2, 3, \ldots$

Analogs of SVD decomposition: I

Unfolding tensor $\mathcal{T} = [t_{i,j,k}] \in \mathbb{F}^{m\times n\times l}$ in 1-index: viewing as a matrix $A_1 \in \mathbb{F}^{m\times(nl)}$.
For a cubic tensor $\mathcal{T} \in \mathbb{R}^{n\times n\times n}$
do orthonormal change of coordinates in each three components $\mathbb{R}^n$:
$\mathcal{T}_1 = \mathcal{T} \times_1 Q_1 \times_2 Q_2 \times_3 Q_3$

One can have at most $\frac{3n(n-1)}{2}$ zero entries in $\mathcal{T}_1$: $t_{j,i,i} = t_{i,j,i} = t_{i,i,j}$ for $i < j$.

Apply QR algorithm to the n columns of the unfolded matrix in mode k indexed by $(j,j), j = 1,\ldots,n$ for $k = 1, 2, 3, 1, 2, 3, \ldots$

Example $2 \times 2 \times 2$: $\begin{bmatrix} t_{1,1,1} & t_{1,2,2} \\ t_{2,1,1} & t_{2,2,2} \end{bmatrix}; \begin{bmatrix} t_{1,1,1} & t_{2,1,2} \\ t_{1,2,1} & t_{2,2,2} \end{bmatrix}; \begin{bmatrix} t_{1,1,1} & t_{2,2,1} \\ t_{1,1,2} & t_{2,2,2} \end{bmatrix}$

Analogs of SVD decomposition: I

Unfolding tensor $\mathcal{T} = [t_{i,j,k}] \in \mathbb{F}^{m\times n\times l}$ in 1-index: viewing as a matrix $A_1 \in \mathbb{F}^{m\times(nl)}$.

For a cubic tensor $\mathcal{T} \in \mathbb{R}^{n\times n\times n}$

do orthonormal change of coordinates in each three components $\mathbb{R}^n$:

$\mathcal{T}_1 = \mathcal{T} \times_1 Q_1 \times_2 Q_2 \times_3 Q_3$

One can have at most $\frac{3n(n-1)}{2}$ zero entries in $\mathcal{T}_1$: $t_{j,i,i} = t_{i,j,i} = t_{i,i,j}$ for $i < j$.

Apply QR algorithm to the n columns of the unfolded matrix in mode k indexed by $(j,j), j = 1, \ldots, n$ for $k = 1, 2, 3, 1, 2, 3, \ldots$

Example $2 \times 2 \times 2$: $\begin{bmatrix} t_{1,1,1} & t_{1,2,2} \\ t_{2,1,1} & t_{2,2,2} \end{bmatrix}; \begin{bmatrix} t_{1,1,1} & t_{2,1,2} \\ t_{1,2,1} & t_{2,2,2} \end{bmatrix}; \begin{bmatrix} t_{1,1,1} & t_{2,2,1} \\ t_{1,1,2} & t_{2,2,2} \end{bmatrix}$

Do QR on each two columns successively to obtain:

$t_{2,1,1} = t_{1,2,1} = t_{1,1,2} = 0.$

CUR approximation of matrices

CUR approximation of matrices

For given $A \in \mathbb{R}^{m\times n}, F \in \mathbb{R}^{m\times p}, E \in \mathbb{R}^{q\times n}$
$\min_{U\in\mathbb{R}^{p\times q}} \|A - EUF\|_F$ achieved for $U = E^\dagger A F^\dagger$
($E^\dagger$ Moore-Penrose inverse)

CUR approximation of matrices

For given $A \in \mathbb{R}^{m \times n}, F \in \mathbb{R}^{m \times p}, E \in \mathbb{R}^{q \times n}$
$\min_{U \in \mathbb{R}^{p \times q}} \|A - EUF\|_F$ achieved for $U = E^\dagger A F^\dagger$
($E^\dagger$ Moore-Penrose inverse)

CUR approximation $C \in \mathbb{R}^{m \times p}, R \in \mathbb{R}^{q \times n}$ some submatrices of A.

CUR approximation of matrices

For given $A \in \mathbb{R}^{m\times n}, F \in \mathbb{R}^{m\times p}, E \in \mathbb{R}^{q\times n}$
$\min_{U\in\mathbb{R}^{p\times q}} \|A - EUF\|_F$ achieved for $U = E^\dagger A F^\dagger$
($E^\dagger$ Moore-Penrose inverse)

CUR approximation $C \in \mathbb{R}^{m\times p}, R \in \mathbb{R}^{q\times n}$ some submatrices of A.

$C(C^\dagger A R^\dagger)R$ best rank $\le min(p,q)$ approximation matrix based on C, R submatrices of A.

CUR approximation of matrices

For given $A \in \mathbb{R}^{m\times n}, F \in \mathbb{R}^{m\times p}, E \in \mathbb{R}^{q\times n}$
$\min_{U\in\mathbb{R}^{p\times q}} \|A - EUF\|_F$ achieved for $U = E^\dagger A F^\dagger$
($E^\dagger$ Moore-Penrose inverse)

CUR approximation $C \in \mathbb{R}^{m\times p}, R \in \mathbb{R}^{q\times n}$ some submatrices of A.

$C(C^\dagger A R^\dagger)R$ best rank $\le min(p,q)$ approximation matrix based on C, R submatrices of A.

Goreinov-Tyrtyshnikov-Zmarashkin 95: for $p = q$ choose $U = A[I,J]^{-1}$

CUR approximation of matrices

For given $A \in \mathbb{R}^{m \times n}, F \in \mathbb{R}^{m \times p}, E \in \mathbb{R}^{q \times n}$
$\min_{U \in \mathbb{R}^{p \times q}} \|A - EUF\|_F$ achieved for $U = E^\dagger A F^\dagger$
($E^\dagger$ Moore-Penrose inverse)

CUR approximation $C \in \mathbb{R}^{m \times p}, R \in \mathbb{R}^{q \times n}$ some submatrices of A.

$C(C^\dagger A R^\dagger)R$ best rank $\le min(p, q)$ approximation matrix based on C, R submatrices of A.

Goreinov-Tyrtyshnikov-Zmarashkin 95: for $p = q$ choose $U = A[I, J]^{-1}$
(corresponds to best CUR approximation on the entries read)

CUR approximation of matrices

For given $A \in \mathbb{R}^{m\times n}, F \in \mathbb{R}^{m\times p}, E \in \mathbb{R}^{q\times n}$
$\min_{U\in\mathbb{R}^{p\times q}} \|A - EUF\|_F$ achieved for $U = E^\dagger A F^\dagger$
($E^\dagger$ Moore-Penrose inverse)

CUR approximation $C \in \mathbb{R}^{m\times p}, R \in \mathbb{R}^{q\times n}$ some submatrices of A.

$C(C^\dagger A R^\dagger)R$ best rank $\le min(p,q)$ approximation matrix based on C, R submatrices of A.

Goreinov-Tyrtyshnikov-Zmarashkin 95: for $p = q$ choose $U = A[I, J]^{-1}$
(corresponds to best CUR approximation on the entries read)
good approximation when the corresponding $\det A[I, J]$ is maximal

CUR approximation of matrices

For given $A \in \mathbb{R}^{m\times n}, F \in \mathbb{R}^{m\times p}, E \in \mathbb{R}^{q\times n}$
$\min_{U\in\mathbb{R}^{p\times q}} \|A - EUF\|_F$ achieved for $U = E^\dagger A F^\dagger$
($E^\dagger$ Moore-Penrose inverse)

CUR approximation $C \in \mathbb{R}^{m\times p}, R \in \mathbb{R}^{q\times n}$ some submatrices of A.

$C(C^\dagger A R^\dagger)R$ best rank $\leq min(p,q)$ approximation matrix based on C, R submatrices of A.

Goreinov-Tyrtyshnikov-Zmarashkin 95: for $p = q$ choose $U = A[I,J]^{-1}$
(corresponds to best CUR approximation on the entries read)
good approximation when the corresponding $\det A[I,J]$ is maximal

Friedland-Mehrmann-Miedlar-Nkengla 08: $U = A[I,J]^\dagger$-numerical

CUR approximation of matrices

For given $A \in \mathbb{R}^{m\times n}, F \in \mathbb{R}^{m\times p}, E \in \mathbb{R}^{q\times n}$
$\min_{U\in\mathbb{R}^{p\times q}} \|A - EUF\|_F$ achieved for $U = E^\dagger A F^\dagger$
($E^\dagger$ Moore-Penrose inverse)

CUR approximation $C \in \mathbb{R}^{m\times p}, R \in \mathbb{R}^{q\times n}$ some submatrices of A.

$C(C^\dagger A R^\dagger)R$ best rank $\le min(p, q)$ approximation matrix based on C, R submatrices of A.

Goreinov-Tyrtyshnikov-Zmarashkin 95: for $p = q$ choose $U = A[I, J]^{-1}$
(corresponds to best CUR approximation on the entries read)
good approximation when the corresponding $\det A[I, J]$ is maximal

Friedland-Mehrmann-Miedlar-Nkengla 08: $U = A[I, J]^\dagger$-numerical
choose several random choices of I, J set of rows and columns of A
such that $A[I, J]$ has maximal product of significant singular values

Extension to 3-tensors I:

$\langle n \rangle := \{1, \ldots, n\}$

Extension to 3-tensors I:

$\langle n \rangle := \{1, \ldots, n\}$

For given $\mathcal{A} \in \mathbb{R}^{m \times n \times l}, F \in \mathbb{R}^{m \times p}, E \in \mathbb{R}^{n \times q}, G \in \mathbb{R}^{l \times r}$,
where $\langle p \rangle \subset \langle n \rangle \times \langle l \rangle, \langle q \rangle \subset \langle m \rangle \times \langle l \rangle, \langle r \rangle \subset \langle m \rangle \times \langle l \rangle$

$\min_{\mathcal{U} \in \mathbb{R}^{p \times q \times r}} \|\mathcal{A} - \mathcal{U} \times F \times E \times G\|_F$ achieved for $\mathcal{U} = \mathcal{A} \times E^\dagger \times F^\dagger \times G^\dagger$

Extension to 3-tensors I:

$\langle n \rangle := \{1, \ldots, n\}$

For given $\mathcal{A} \in \mathbb{R}^{m\times n\times l}, F \in \mathbb{R}^{m\times p}, E \in \mathbb{R}^{n\times q}, G \in \mathbb{R}^{l\times r}$,
where $\langle p \rangle \subset \langle n \rangle \times \langle l \rangle, \langle q \rangle \subset \langle m \rangle \times \langle l \rangle, \langle r \rangle \subset \langle m \rangle \times \langle l \rangle$

$\min_{\mathcal{U}\in\mathbb{R}^{p\times q\times r}} \|\mathcal{A} - \mathcal{U} \times F \times E \times G\|_F$ achieved for $\mathcal{U} = \mathcal{A} \times E^\dagger \times F^\dagger \times G^\dagger$

CUR approximation of $\mathcal{A}$ obtained by choosing E, F, G submatrices of unfolded $\mathcal{A}$ in the mode $1, 2, 3$.

Extensions to 3-tensors: II

$\mathcal{A} = [a_{i,j,k}] \in \mathbb{R}^{m\times n\times \ell}$ - 3-tensor

given $I \subset \langle m\rangle, J \subset \langle n\rangle, K \subset \langle \ell\rangle$ define

$R := \mathcal{A}_{\langle m\rangle,J,K} = [a_{i,j,k}]_{\langle m\rangle,J,K} \in \mathbb{R}^{m\times(\#J\cdot\#K)}$,

$C := \mathcal{A}_{I,\langle n\rangle,K} \in \mathbb{R}^{\langle n\rangle\times(\#I\cdot\#K)}$,

$D := \mathcal{A}_{I,J,\langle \ell\rangle} \in \mathbb{R}^{l\times(\#I\cdot\#J)}$

Problem: Find 3-tensor $\mathcal{U} \in \mathbb{R}^{(\#J\cdot\#K)\times(\#I\cdot\#K)\times(\#I\cdot\#J)}$

such that $\mathcal{A}$ is approximated by the Tucker tensor

$\mathcal{V} = \mathcal{U} \times_1 C \times_2 R \times_3 D$

where $\mathcal{U}$ is the least squares solution

$$\mathcal{U}_{\mathrm{opt}} \in \arg \min_{\mathcal{U}\in\mathbb{R}^{\text{three tensor}}} \sum_{(i,j,k)\in\mathcal{S}} \left(a_{i,j,k} - (\mathcal{U} \times_1 C \times_2 R \times_3 D)_{i,j,k}\right)^2$$

$\mathcal{S} = (\langle m\rangle \times J \times K) \cup (I \times \langle n\rangle \times K) \cup (I \times J \times \langle \ell\rangle)$

Extension to 3-tensors: III

For $\#I = \#J = p,\ \#K = p^2,\ I \subset \langle m\rangle, J \subset \langle n\rangle, K \subset \langle \ell\rangle$
generically there is an exact solution to $\mathcal{U}_{\mathrm{opt}} \in \mathbb{R}^{p^3 \times p^3 \times p^2}$
obtained by unfolding in third direction
View $\mathcal{A}$ as $A \in \mathbb{R}^{(mn)\times\ell}$ by identifying
$\langle m\rangle \times \langle n\rangle \equiv \langle mn\rangle,\ I_1 = I \times J,\ J_1 = K$ and apply CUR again.

More generally, given $\#I = p,\ \#J = q,\ \#K = r$.
For $L = I \times J$ approximate $\mathcal{A}$ by $\mathcal{A}_{\langle m\rangle,\langle n\rangle,K} E^{\dagger}_{L,K} \mathcal{A}_{I,J,\langle \ell\rangle}$
Then for each $k \in K$ approximate each matrix $\mathcal{A}_{\langle m\rangle,\langle n\rangle,\{k\}}$ by
$\mathcal{A}_{\langle m\rangle,J,\{k\}} E^{\dagger}_{I,J,\{k\}} \mathcal{A}_{I,\langle n\rangle,\{k\}}$

References I

- R.A. Brualdi, Convex sets of nonnegative matrices, *Canad. J. Math* 20(1968), 144-157.
- S. Friedland, On the generic rank of 3-tensors, arXiv: 0805.3777v2.
- S. Friedland, S. Gauber and L. Han, Perron-Frobenius theorem for nonnegative multilinear forms, *arXiv:0905.1626*.
- S. Friedland, V. Mehrmann, A. Miedlar, and M. Nkengla, Fast low rank approximations of matrices and tensors, submitted, www.matheon.de/preprints/4903.
- S.A. Goreinov, E.E. Tyrtyshnikov, N.L. Zmarashkin, Pseudo-skeleton approximations of matrices, *Reports of the Russian Academy of Sciences* 343(2) (1995), 151-152.

References II

- S.A. Goreinov, E.E. Tyrtyshnikov, N.L. Zmarashkin, A theory of pseudo-skeleton approximations of matrices, *Linear Algebra Appl.* 261 (1997), 1-21.
- L.H. Lim, Singular values and eigenvalues of tensors: a variational approach, *CAMSAP* 05, 1 (2005), 129-132.
- M.W. Mahoney and P. Drineas, CUR matrix decompositions for improved data analysis, *PNAS* 106, (2009), 697-702.
- M.V. Menon, Matrix links, an extremisation problem and the reduction of a nonnegative matrix to one with with prescribed row and column sums.
- M.V. Menon and H. Schneider, The spectrum of a nonlinear operator associated with a matrix, *Linear Algebra Appl.* 321–334.