Results and problems for 3-tensors

Shmuel Friedland
Univ. Illinois at Chicago

NSF TENSOR COMPUTATION WORKSHOP,
Arlington, February 20-21, 2009

Overview

1. The conjectured value of the generic rank of 3-tensor and its numerical verification over $\mathbb{C}$. Some results for $\mathbb{R}$.
2. Best rank one approximation.
   1. ℓ_2 case.
   2. $\ell_p, p \in (1, \infty)$ case.
   3. Perron-Frobenius theorem for *irreducible* nonnegative tensors for $p = 3$.
3. Analogs of SVD decomposition of 3-tensors.
   1. The maximal number of zero entries in 3-tensor under the orthogonal conjugation in each of 3-modes.
   2. The expected limit form of tensor under the iteration an analog of *QR* algorithm.
   3. An analog of Kogbetliantz's algorithm.
4. CUR decompositions for tensors
5. Scaling of nonnegative tensors to balanced tensors. (The analog of scaling to doubly stochastic matrices.)

Generic rank I

Generic rank I

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

Generic rank I

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

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

Generic rank I

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

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

Generic rank I

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

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

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

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

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

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 $(m-1)(n-1) \le l$.

Generic rank I

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

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

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 $(m-1)(n-1)\le l$.

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

Generic rank I

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

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

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 $(m-1)(n-1)\le l$.

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

Generic rank I

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

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

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 $(m-1)(n-1)\le l$.

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

Generic rank 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 $(3,n,l) \neq (3,2p+1,2p+1)$

Generic rank II

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

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

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

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

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

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

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 distinct open connected semi-algebraic sets s.t.

$\mathrm{Closure}(\cup_{i=1}^{c(m,n,l)}) = \mathbb{R}^{m \times n \times l}$

$\mathrm{rank}\ \mathcal{T} = \mathrm{grank}_{\mathbb{C}}(m, n, l)$ 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 $i = 2, \ldots, c(m, n, l)$

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 distinct open connected semi-algebraic sets s.t.

$\mathrm{Closure}(\cup_{i=1}^{c(m,n,l)}) = \mathbb{R}^{m\times n\times l}$

$\mathrm{rank}\ \mathcal{T} = \mathrm{grank}_{\mathbb{C}}(m,n,l)$ 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 $l = (m-1)(n-1)$ $\exists m, n$:

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

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 distinct open connected semi-algebraic sets s.t.
Closure$(\cup_{i=1}^{c(m,n,l)}) = \mathbb{R}^{m\times n\times l}$
$\mathrm{rank}\ \mathcal{T} = \mathrm{grank}_{\mathbb{C}}(m,n,l)$ 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 $l = (m-1)(n-1)$ $\exists m, n$:
$c(m,n,l) > 1, \rho_{c(m,n,l)} \ge \mathrm{grank}_{\mathbb{C}}(m,n,l) + 1$

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

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 distinct open connected semi-algebraic sets s.t.
Closure$(\cup_{i=1}^{c(m,n,l)}) = \mathbb{R}^{m\times n\times l}$
$\mathrm{rank}\ \mathcal{T} = \mathrm{grank}_{\mathbb{C}}(m,n,l)$ 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 $l = (m-1)(n-1)$ $\exists m,n$:
$c(m,n,l) > 1, \rho_{c(m,n,l)} \ge \mathrm{grank}_{\mathbb{C}}(m,n,l) + 1$

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

More results?

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

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

Lagrange multipliers: $\mathcal{T}\times\mathbf{y}\otimes\mathbf{z} := \sum_{j=k=1} t_{i,j,k} y_j z_k = \lambda\mathbf{x}$
$\mathcal{T}\times\mathbf{x}\otimes\mathbf{z} = \lambda\mathbf{y}$, $\mathcal{T}\times\mathbf{x}\otimes\mathbf{y} = \lambda\mathbf{z}$
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: $\max_{\|\mathbf{x}\|_p=\|\mathbf{y}\|_p=\|\mathbf{z}\|_p=1} \sum_{i=j=k}^{m,n,l} t_{i,j,k}x_iy_jz_k$

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

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

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

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 $p = 3$, and probably for $p > 3$
No, for $p = 2$, and probably for $p < 3$

Analogs of SVD decomposition: I

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

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 $\mathcal{T}_1$: $t_{j,i,i} = t_{i,j,i} =_{i,i,j}$ for $i < j$.

Analogs of SVD decomposition: I

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 $\mathcal{T}_1$: $t_{j,i,i} = t_{i,j,i} =_{i,i,j}$ 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 $k = 1,2,3,1,2,3,...$

Analogs of SVD decomposition: I

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 $\mathcal{T}_1$: $t_{j,i,i} = t_{i,j,i} =_{i,i,j}$ for $i < j$.

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

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

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 $\mathcal{T}_1$: $t_{j,i,i} = t_{i,j,i} =_{i,i,j}$ 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 $k = 1, 2, 3, 1, 2, 3, ...$

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

An analog of Kogbetliantz's algorithm

$\mathcal{T} \in \mathbb{R}^{m\times n\times l}, 2 \leq m, n, l$

An analog of Kogbetliantz's algorithm

$\mathcal{T} \in \mathbb{R}^{m \times n \times l}, 2 \leq m, n, l$

Choose $2 \times 2 \times 2$ subtensor $\mathcal{A} \in \mathbb{R}^{2 \times 2 \times 2}$

An analog of Kogbetliantz's algorithm

$\mathcal{T} \in \mathbb{R}^{m \times n \times l}, 2 \leq m, n, l$

Choose $2 \times 2 \times 2$ subtensor $\mathcal{A} \in \mathbb{R}^{2 \times 2 \times 2}$

Find best rank approximation $\mathcal{B} = [b_{i,j,k}] = \mathcal{A} \times_1 Q_1 \times_2 Q_2 \times_3 Q_3$
such that $b_{1,1,1}$ is maximal possible

An analog of Kogbetliantz's algorithm

$\mathcal{T} \in \mathbb{R}^{m \times n \times l}, 2 \le m, n, l$

Choose $2 \times 2 \times 2$ subtensor $\mathcal{A} \in \mathbb{R}^{2 \times 2 \times 2}$

Find best rank approximation $\mathcal{B} = [b_{i,j,k}] = \mathcal{A} \times_1 Q_1 \times_2 Q_2 \times_3 Q_3$
such that $b_{1,1,1}$ is maximal possible

Obtain the corresponding $\mathcal{T}'$

An analog of Kogbetliantz's algorithm

$\mathcal{T} \in \mathbb{R}^{m\times n\times l}, 2 \leq m, n, l$

Choose $2 \times 2 \times 2$ subtensor $\mathcal{A} \in \mathbb{R}^{2\times 2\times 2}$

Find best rank approximation $\mathcal{B} = [b_{i,j,k}] = \mathcal{A} \times_1 Q_1 \times_2 Q_2 \times_3 Q_3$ such that $b_{1,1,1}$ is maximal possible

Obtain the corresponding $\mathcal{T}'$

Repeat with all choices of subtensors

An analog of Kogbetliantz's algorithm

$\mathcal{T} \in \mathbb{R}^{m \times n \times l}, 2 \leq m, n, l$

Choose $2 \times 2 \times 2$ subtensor $\mathcal{A} \in \mathbb{R}^{2 \times 2 \times 2}$

Find best rank approximation $\mathcal{B} = [b_{i,j,k}] = \mathcal{A} \times_1 Q_1 \times_2 Q_2 \times_3 Q_3$
such that $b_{1,1,1}$ is maximal possible

Obtain the corresponding $\mathcal{T}'$

Repeat with all choices of subtensors

Can be parallelized

An analog of Kogbetliantz's algorithm

$\mathcal{T} \in \mathbb{R}^{m \times n \times l}, 2 \leq m, n, l$

Choose $2 \times 2 \times 2$ subtensor $\mathcal{A} \in \mathbb{R}^{2 \times 2 \times 2}$

Find best rank approximation $\mathcal{B} = [b_{i,j,k}] = \mathcal{A} \times_1 Q_1 \times_2 Q_2 \times_3 Q_3$
such that $b_{1,1,1}$ is maximal possible

Obtain the corresponding $\mathcal{T}'$

Repeat with all choices of subtensors

Can be parallelized

Does it converge in generic case and to what

CUR approximation of matrices

CUR approximation of matrices

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

CUR approximation of matrices

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

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 $A \in \mathbb{R}^{m\times n}, F \in \mathbb{R}^{m\times p}, E \in \mathbb{R}^{q\times n}$
$\min_{U\in\mathbb{C}^{p\times q}} \|A - EUF\|_F$ achieved for $U = E^\dagger A F^\dagger$

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.

CUR approximation of matrices

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

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 $A \in \mathbb{R}^{m\times n}, F \in \mathbb{R}^{m\times p}, E \in \mathbb{R}^{q\times n}$
$\min_{U\in\mathbb{C}^{p\times q}} \|A - EUF\|_F$ achieved for $U = E^\dagger A F^\dagger$

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 $A \in \mathbb{R}^{m\times n}, F \in \mathbb{R}^{m\times p}, E \in \mathbb{R}^{q\times n}$
$\min_{U\in\mathbb{C}^{p\times q}} \|A - EUF\|_F$ achieved for $U = E^\dagger A F^\dagger$

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 $A \in \mathbb{R}^{m\times n}, F \in \mathbb{R}^{m\times p}, E \in \mathbb{R}^{q\times n}$
$\min_{U\in\mathbb{C}^{p\times q} } \|A - EUF\|_F$ achieved for $U = E^{\dagger}AF^{\dagger}$

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

$C(C^{\dagger}AR^{\dagger})R$ 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 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:

Extension to 3-tensors I:

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 $\langle p \rangle \subset \langle n \rangle \times \langle l \rangle, \langle q \rangle \subset \langle m \rangle \times \langle l \rangle, \langle r \rangle \subset \langle m \rangle \times \langle l \rangle$

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

Extension to 3-tensors I:

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 $\langle p\rangle \subset \langle n\rangle \times \langle l\rangle, \langle q\rangle \subset \langle m\rangle \times \langle l\rangle, \langle r\rangle \subset \langle m\rangle \times \langle l\rangle$

$\min_{\mathcal{U}\in\mathbb{C}^{p\times q\times r}} \|\mathcal{A} - \mathcal{U} \times F \times E \times G\|_F$ 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
$R := \alpha_{\langle m\rangle,J,K} = [a_{i,j,k}]_{\langle m\rangle,J,K} \in \mathbb{R}^{m\times(\#J\cdot\#K)}$,
$C := \alpha_{I,\langle n\rangle,K} \in \mathbb{R}^{\langle n\rangle\times(\#I\cdot\#K)}$,
$D := \alpha_{I,J,\langle \ell\rangle} \in \mathbb{R}^{l\times(\#I\cdot\#J)}$
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

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

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

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
$\mathcal{A}_{\langle m \rangle, J, \{k\}} E^{\dagger}_{I,J,\{k\}} \mathcal{A}_{I, \langle n \rangle, \{k\}}$

Scaling of nonnegative tensors to balanced tensors

$0 \le \mathcal{T} = [t_{i,j,k}] \in \mathbb{R}^{m \times n \times l}$ balanced if each unfolding has fixed row sum:
$\sum_{j,k} t_{i,j,k} = \alpha > 0, \ \sum_{i,k} t_{i,j,k} = \beta > 0, \sum_{i,j} t_{i,j,k} = \gamma > 0$

Scaling of nonnegative tensors to balanced tensors

$0 \le \mathcal{T} = [t_{i,j,k}] \in \mathbb{R}^{m\times n\times l}$ balanced if each unfolding has fixed row sum:
$\sum_{j,k} t_{i,j,k} = \alpha > 0,\ \sum_{i,k} t_{i,j,k} = \beta > 0, \sum_{i,j} t_{i,j,k} = \gamma > 0$

Find nec. and suf. conditions for scaling:
$\mathcal{T}' = [x_i y_j z_k t_{i,j,k}], \mathbf{x}, \mathbf{y}, \mathbf{z} > \mathbf{0}$ such that $\mathcal{T}'$ balanced

Scaling of nonnegative tensors to balanced tensors

$0 \le \mathcal{T} = [t_{i,j,k}] \in \mathbb{R}^{m\times n\times l}$ balanced if each unfolding has fixed row sum:
$\sum_{j,k} t_{i,j,k} = \alpha > 0, \ \sum_{i,k} t_{i,j,k} = \beta > 0, \sum_{i,j} t_{i,j,k} = \gamma > 0$

Find nec. and suf. conditions for scaling:
$\mathcal{T}' = [x_i y_j z_k t_{i,j,k}], \mathbf{x}, \mathbf{y}, \mathbf{z} > \mathbf{0}$ such that $\mathcal{T}'$ balanced

THM: $\mathcal{T} \in \mathbb{R}_+^{m\times n\times l}, 1 < m \le n \le,$ each $m \times m$ submatrix of the unfolded tensor $A \in \mathbb{R}^{m\times nl}$ in the first mode has positive permanent. Then there exists a "unique" scaling of $\mathcal{T}'$ of $\mathcal{T}$ such that $A' \in \mathbb{R}^{m\times nl}$ is stochastic.

Scaling of nonnegative tensors to balanced tensors

$0 \le \mathcal{T} = [t_{i,j,k}] \in \mathbb{R}^{m \times n \times l}$ balanced if each unfolding has fixed row sum:
$\sum_{j,k} t_{i,j,k} = \alpha > 0, \ \sum_{i,k} t_{i,j,k} = \beta > 0, \sum_{i,j} t_{i,j,k} = \gamma > 0$

Find nec. and suf. conditions for scaling:
$\mathcal{T}' = [x_i y_j z_k t_{i,j,k}], \mathbf{x}, \mathbf{y}, \mathbf{z} > \mathbf{0}$ such that $\mathcal{T}'$ balanced

THM: $\mathcal{T} \in \mathbb{R}_+^{m \times n \times l}, 1 < m \le n \le,$ each $m \times m$ submatrix of the unfolded tensor $A \in \mathbb{R}^{m \times nl}$ in the first mode has positive permanent. Then there exists a "unique" scaling of $\mathcal{T}'$ of $\mathcal{T}$ such that $A' \in \mathbb{R}^{m \times nl}$ is stochastic.

Problem: Is Sinkhorn scaling algorithm in this case working?

References I

- S. Friedland, On the generic rank of 3-tensors, arXiv: 0805.3777v2.
- 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.
- 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.

References II

M.W. Mahoney and P. Drineas, CUR matrix decompositions for improved data analysis, *PNAS* 106, (2009), 697-702.