Fixed points theorems for nonnegative tensors and Newton method

Shmuel Friedland
Univ. Illinois at Chicago

October 29, 2009
SIAM LA09, Monterey, CA

Overview

Overview

1. Perron-Frobenius theorem for *irreducible* nonnegative tensors.

Overview

1. Perron-Frobenius theorem for *irreducible* nonnegative tensors.
2. Diagonal scaling of nonnegative tensors to tensors with given rows, columns and depth sums.

SVD of nonnegative matrices

SVD of nonnegative matrices

$A \in \mathbb{R}^{m \times n}$, $\sigma_1(A) \geq \ldots \geq 0$ singular values
$A\mathbf{y}_i = \sigma_i(A)\mathbf{x}_i, \; A^\top \mathbf{x}_i = \sigma_i(A)\mathbf{y}_i$

SVD of nonnegative matrices

$A \in \mathbb{R}^{m \times n}$, $\sigma_1(A) \geq \ldots \geq 0$ singular values
$A\mathbf{y}_i = \sigma_i(A)\mathbf{x}_i, \; A^\top \mathbf{x}_i = \sigma_i(A)\mathbf{y}_i$
$\pm\sigma_i(A), i = 1, \ldots$ are critical values of $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top A\mathbf{y}$
restricted to $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$

SVD of nonnegative matrices

$A \in \mathbb{R}^{m \times n}$, $\sigma_1(A) \ge \ldots \ge 0$ singular values
$A\mathbf{y}_i = \sigma_i(A)\mathbf{x}_i,\ A^\top \mathbf{x}_i = \sigma_i(A)\mathbf{y}_i$
$\pm\sigma_i(A), i = 1, \ldots$ are critical values of $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top A\mathbf{y}$
restricted to $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$

Perron-Frobenius for $A = [a_{ij}] \in \mathbb{R}_+^{m \times n}$:

SVD of nonnegative matrices

$A \in \mathbb{R}^{m\times n}$, $\sigma_1(A) \ge \ldots \ge 0$ singular values
$A\mathbf{y}_i = \sigma_i(A)\mathbf{x}_i,\ A^\top\mathbf{x}_i = \sigma_i(A)\mathbf{y}_i$
$\pm\sigma_i(A), i = 1, \ldots$ are critical values of $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top A\mathbf{y}$
restricted to $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$

Perron-Frobenius for $A = [a_{ij}] \in \mathbb{R}_+^{m\times n}$:

$\mathbf{u} \in \mathbb{R}_+^m, \mathbf{v} \in \mathbb{R}_+^n,\ \mathbf{u}^\top\mathbf{u} = \mathbf{v}^\top\mathbf{v} = 1\ A\mathbf{v} = \sigma_1(A)\mathbf{u},\ A^\top\mathbf{u} = \sigma_1(A)\mathbf{v}$

SVD of nonnegative matrices

$A \in \mathbb{R}^{m\times n}$, $\sigma_1(A) \geq \ldots \geq 0$ singular values
$A\mathbf{y}_i = \sigma_i(A)\mathbf{x}_i,\ A^\top\mathbf{x}_i = \sigma_i(A)\mathbf{y}_i$
$\pm\sigma_i(A), i = 1, \ldots$ are critical values of $f(\mathbf{x},\mathbf{y}) = \mathbf{x}^\top A\mathbf{y}$
restricted to $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$

Perron-Frobenius for $A = [a_{ij}] \in \mathbb{R}_+^{m\times n}$:

$\mathbf{u} \in \mathbb{R}_+^m, \mathbf{v} \in \mathbb{R}_+^n,\ \mathbf{u}^\top\mathbf{u} = \mathbf{v}^\top\mathbf{v} = 1\ A\mathbf{v} = \sigma_1(A)\mathbf{u},\ A^\top\mathbf{u} = \sigma_1(A)\mathbf{v}$

$\max_{\mathbf{x}\in\mathbb{R}^m,\mathbf{y}\in\mathbb{R}^n,\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=1} \mathbf{x}^\top A\mathbf{y} = \mathbf{u}^\top A\mathbf{v}$.

SVD of nonnegative matrices

$A \in \mathbb{R}^{m\times n}$, $\sigma_1(A) \ge \ldots \ge 0$ singular values
$A\mathbf{y}_i = \sigma_i(A)\mathbf{x}_i,\ A^\top \mathbf{x}_i = \sigma_i(A)\mathbf{y}_i$
$\pm\sigma_i(A), i = 1, \ldots$ are critical values of $f(\mathbf{x},\mathbf{y}) = \mathbf{x}^\top A\mathbf{y}$
restricted to $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$

Perron-Frobenius for $A = [a_{ij}] \in \mathbb{R}_+^{m\times n}$:

$\mathbf{u} \in \mathbb{R}_+^m, \mathbf{v} \in \mathbb{R}_+^n,\ \mathbf{u}^\top\mathbf{u} = \mathbf{v}^\top\mathbf{v} = 1\ A\mathbf{v} = \sigma_1(A)\mathbf{u},\ A^\top\mathbf{u} = \sigma_1(A)\mathbf{v}$

$\max_{\mathbf{x}\in\mathbb{R}^m,\mathbf{y}\in\mathbb{R}^n,\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=1} \mathbf{x}^\top A\mathbf{y} = \mathbf{u}^\top A\mathbf{v}$.

$G(A) = (V_1 \cup V_2, E)$ bipartite graph on
$V_1 = \langle m\rangle := \{1, \ldots, m\}, V_2 := \langle n\rangle, (i,j) \in E \iff a_{ij} > 0$.

SVD of nonnegative matrices

$A \in \mathbb{R}^{m\times n}$, $\sigma_1(A) \geq \ldots \geq 0$ singular values
$A\mathbf{y}_i = \sigma_i(A)\mathbf{x}_i,\ A^\top\mathbf{x}_i = \sigma_i(A)\mathbf{y}_i$
$\pm\sigma_i(A), i = 1, \ldots$ are critical values of $f(\mathbf{x},\mathbf{y}) = \mathbf{x}^\top A\mathbf{y}$
restricted to $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$

Perron-Frobenius for $A = [a_{ij}] \in \mathbb{R}_+^{m\times n}$:

$\mathbf{u} \in \mathbb{R}_+^m, \mathbf{v} \in \mathbb{R}_+^n,\ \mathbf{u}^\top\mathbf{u} = \mathbf{v}^\top\mathbf{v} = 1\ A\mathbf{v} = \sigma_1(A)\mathbf{u},\ A^\top\mathbf{u} = \sigma_1(A)\mathbf{v}$

$\max_{\mathbf{x}\in\mathbb{R}^m,\mathbf{y}\in\mathbb{R}^n,\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=1} \mathbf{x}^\top A\mathbf{y} = \mathbf{u}^\top A\mathbf{v}$.

$G(A) = (V_1 \cup V_2, E)$ bipartite graph on
$V_1 = \langle m\rangle := \{1,\ldots,m\}, V_2 := \langle n\rangle, (i,j) \in E \iff a_{ij} > 0$.

If $G(A)$ connected. Then $\mathbf{u},\mathbf{v}$ unique.

SVD of nonnegative matrices

$A \in \mathbb{R}^{m\times n}$, $\sigma_1(A) \ge \ldots \ge 0$ singular values
$A\mathbf{y}_i = \sigma_i(A)\mathbf{x}_i,\ A^\top \mathbf{x}_i = \sigma_i(A)\mathbf{y}_i$
$\pm\sigma_i(A), i = 1, \ldots$ are critical values of $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top A\mathbf{y}$
restricted to $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$

Perron-Frobenius for $A = [a_{ij}] \in \mathbb{R}_+^{m\times n}$:

$\mathbf{u} \in \mathbb{R}_+^m, \mathbf{v} \in \mathbb{R}_+^n,\ \mathbf{u}^\top\mathbf{u} = \mathbf{v}^\top\mathbf{v} = 1\ A\mathbf{v} = \sigma_1(A)\mathbf{u},\ A^\top\mathbf{u} = \sigma_1(A)\mathbf{v}$

$\max_{\mathbf{x}\in\mathbb{R}^m, \mathbf{y}\in\mathbb{R}^n, \|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1} \mathbf{x}^\top A\mathbf{y} = \mathbf{u}^\top A\mathbf{v}$.

$G(A) = (V_1 \cup V_2, E)$ bipartite graph on
$V_1 = \langle m\rangle := \{1, \ldots, m\}, V_2 := \langle n\rangle, (i, j) \in E \iff a_{ij} > 0$.

If $G(A)$ connected. Then $\mathbf{u}, \mathbf{v}$ unique.

Proof: $A^\top A, AA^\top$ are irreducible

Rank one approximations for 3-tensors

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A}, \mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k} b_{i,j,k}$, $\|\mathcal{T}\|_2 = \sqrt{\langle \mathcal{T}, \mathcal{T}\rangle}$

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

Rank one approximations for 3-tensors

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A},\mathcal{B}\rangle=\sum_{i=j=k}^{m,n,l} a_{i,j,k}b_{i,j,k},\ \|\mathcal{T}\|_2=\sqrt{\langle \mathcal{T},\mathcal{T}\rangle}$
$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z},\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle=(\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1,\ldots,\mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$
$\mathrm{P}_{\mathbf{X}}(\mathcal{T})=\sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle\mathcal{X}_i,\quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2=\sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle^2$

Rank one approximations for 3-tensors

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A},\mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k}b_{i,j,k},\ \|\mathcal{T}\|_2 = \sqrt{\langle \mathcal{T},\mathcal{T}\rangle}$
$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z},\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle = (\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1,\ldots,\mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$
$\mathrm{P}_{\mathbf{X}}(\mathcal{T}) = \sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle\mathcal{X}_i,\quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2 = \sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle^2$

$\|\mathcal{T}\|_2^2 = \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2 + \|\mathcal{T}-\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2$

Rank one approximations for 3-tensors

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle\mathcal{A},\mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k}b_{i,j,k}$, $\|\mathcal{T}\|_2 = \sqrt{\langle\mathcal{T},\mathcal{T}\rangle}$
$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}, \mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle = (\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1,\ldots,\mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$
$\mathrm{P}_{\mathbf{X}}(\mathcal{T}) = \sum_{i=1}^{d}\langle\mathcal{T},\mathcal{X}_i\rangle\mathcal{X}_i$, $\quad\|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2 = \sum_{i=1}^{d}\langle\mathcal{T},\mathcal{X}_i\rangle^2$

$\|\mathcal{T}\|_2^2 = \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2 + \|\mathcal{T}-\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2$

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

Rank one approximations for 3-tensors

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A}, \mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k} b_{i,j,k}, \ \|\mathcal{T}\|_2 = \sqrt{\langle \mathcal{T}, \mathcal{T}\rangle}$
$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}, \mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle = (\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1, \ldots, \mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$
$\mathrm{P}_{\mathbf{X}}(\mathcal{T}) = \sum_{i=1}^{d} \langle \mathcal{T}, \mathcal{X}_i\rangle \mathcal{X}_i, \quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2 = \sum_{i=1}^{d} \langle \mathcal{T}, \mathcal{X}_i\rangle^2$

$\|\mathcal{T}\|_2^2 = \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2 + \|\mathcal{T} - \mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2$

Best rank one approximation of $\mathcal{T}$:
$\min_{\mathbf{x},\mathbf{y},\mathbf{z}} \|\mathcal{T} - \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|_2 = \min_{\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=\|\mathbf{z}\|_2=1, a} \|\mathcal{T} - a\,\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|_2$

Rank one approximations for 3-tensors

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A},\mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k}b_{i,j,k},\ \|\mathcal{T}\|_2 = \sqrt{\langle \mathcal{T},\mathcal{T}\rangle}$

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

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

$\mathrm{P}_{\mathbf{X}}(\mathcal{T}) = \sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle\mathcal{X}_i, \quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2 = \sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle^2$

$\|\mathcal{T}\|_2^2 = \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2 + \|\mathcal{T}-\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2$

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

$\min_{\mathbf{x},\mathbf{y},\mathbf{z}}\|\mathcal{T}-\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|_2 = \min_{\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=\|\mathbf{z}\|_2=1,a}\|\mathcal{T}-a\,\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|_2$

Equivalent: $\max_{\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=\|\mathbf{z}\|_2=1}\sum_{i=j=k}^{m,n,l} t_{i,j,k}x_iy_jz_k$

Rank one approximations for 3-tensors

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A}, \mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k} b_{i,j,k}, \ \|\mathcal{T}\|_2 = \sqrt{\langle \mathcal{T}, \mathcal{T}\rangle}$
$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}, \mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle = (\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1,\ldots,\mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$
$\mathrm{P}_{\mathbf{X}}(\mathcal{T}) = \sum_{i=1}^d \langle \mathcal{T}, \mathcal{X}_i\rangle \mathcal{X}_i, \quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2 = \sum_{i=1}^d \langle \mathcal{T}, \mathcal{X}_i\rangle^2$

$\|\mathcal{T}\|_2^2 = \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2 + \|\mathcal{T} - \mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2$

Best rank one approximation of $\mathcal{T}$:
$\min_{\mathbf{x},\mathbf{y},\mathbf{z}} \|\mathcal{T} - \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|_2 = \min_{\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=\|\mathbf{z}\|_2=1,a} \|\mathcal{T} - a\,\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|_2$

Equivalent: $\max_{\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=\|\mathbf{z}\|_2=1} \sum_{i=j=k}^{m,n,l} t_{i,j,k} x_i y_j z_k$

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

Rank one approximations for 3-tensors

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A},\mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k}b_{i,j,k},\ \|\mathcal{T}\|_2 = \sqrt{\langle \mathcal{T},\mathcal{T}\rangle}$
$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z},\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle = (\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1,\ldots,\mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$
$\mathrm{P}_{\mathbf{X}}(\mathcal{T}) = \sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle\mathcal{X}_i, \quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2 = \sum_{i=1}^{d}\langle \mathcal{T},\mathcal{X}_i\rangle^2$

$\|\mathcal{T}\|_2^2 = \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2 + \|\mathcal{T}-\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2$

Best rank one approximation of $\mathcal{T}$:
$\min_{\mathbf{x},\mathbf{y},\mathbf{z}}\|\mathcal{T}-\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|_2 = \min_{\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=\|\mathbf{z}\|_2=1,a}\|\mathcal{T}-a\,\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|_2$

Equivalent: $\max_{\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=\|\mathbf{z}\|_2=1}\sum_{i=j=k}^{m,n,l} t_{i,j,k}x_iy_jz_k$

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 for 3-tensors

$\mathbb{R}^{m\times n\times l}$ IPS: $\langle \mathcal{A},\mathcal{B}\rangle = \sum_{i=j=k}^{m,n,l} a_{i,j,k}b_{i,j,k},\ \|\mathcal{T}\|_2 = \sqrt{\langle \mathcal{T},\mathcal{T}\rangle}$
$\langle \mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z},\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\rangle = (\mathbf{u}^\top\mathbf{x})(\mathbf{v}^\top\mathbf{y})(\mathbf{w}^\top\mathbf{z})$

$\mathbf{X}$ subspace of $\mathbb{R}^{m\times n\times l}$, $\mathcal{X}_1,\ldots,\mathcal{X}_d$ an orthonormal basis of $\mathbf{X}$
$\mathrm{P}_{\mathbf{X}}(\mathcal{T}) = \sum_{i=1}^d \langle \mathcal{T},\mathcal{X}_i\rangle\mathcal{X}_i,\quad \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2 = \sum_{i=1}^d \langle \mathcal{T},\mathcal{X}_i\rangle^2$

$\|\mathcal{T}\|_2^2 = \|\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2 + \|\mathcal{T}-\mathrm{P}_{\mathbf{X}}(\mathcal{T})\|_2^2$

Best rank one approximation of $\mathcal{T}$:
$\min_{\mathbf{x},\mathbf{y},\mathbf{z}} \|\mathcal{T}-\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|_2 = \min_{\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=\|\mathbf{z}\|_2=1,a} \|\mathcal{T}-a\,\mathbf{x}\otimes\mathbf{y}\otimes\mathbf{z}\|_2$

Equivalent: $\max_{\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=\|\mathbf{z}\|_2=1} \sum_{i=j=k}^{m,n,l} t_{i,j,k}x_iy_jz_k$

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

Lagrange multipliers: $\mathcal{T}\times\mathbf{y}\otimes\mathbf{z} := \sum_{j=k=1} t_{i,j,k}y_jz_k = \lambda\mathbf{x}^{p-1}$
$\mathcal{T}\times\mathbf{x}\otimes\mathbf{z} = \lambda\mathbf{y}^{p-1}$, $\mathcal{T}\times\mathbf{x}\otimes\mathbf{y} = \lambda\mathbf{z}^{p-1}$ $(p = \frac{2t}{2s-1}, t, s \in \mathbb{N})$
See L.-H. Lim 2005 for more general results

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: $\mathcal{T} \times \mathbf{y} \otimes \mathbf{z} := \sum_{j=k=1} t_{i,j,k} y_j z_k = \lambda \mathbf{x}^{p-1}$
$\mathcal{T} \times \mathbf{x} \otimes \mathbf{z} = \lambda \mathbf{y}^{p-1}, \ \mathcal{T} \times \mathbf{x} \otimes \mathbf{y} = \lambda \mathbf{z}^{p-1}$ $(p = \frac{2t}{2s-1}, t, s \in \mathbb{N})$
See L.-H. Lim 2005 for more general results
$p = 3$ 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: $\mathcal{T} \times \mathbf{y}\otimes\mathbf{z} := \sum_{j=k=1} t_{i,j,k}y_jz_k = \lambda \mathbf{x}^{p-1}$
$\mathcal{T} \times \mathbf{x}\otimes\mathbf{z} = \lambda \mathbf{y}^{p-1}$, $\mathcal{T} \times \mathbf{x}\otimes\mathbf{y} = \lambda \mathbf{z}^{p-1}$ $(p = \frac{2t}{2s-1}, t, s \in \mathbb{N})$
See L.-H. Lim 2005 for more general results
$p = 3$ 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?, UNIQUENESS

Yes, for $p \geq 3$, No, for $p < 3$,
Friedland-Gauber-Han [5]

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

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

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

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

Outline of the proof

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

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

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

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

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

Outline of the proof

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

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

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

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

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

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

$\mathcal{T} = [t_{i,j,k}]$ 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 $t_{i,j,k} > 0$, sim. for j, k

Outline of the proof

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

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

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

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

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

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

$\mathcal{T} = [t_{i,j,k}]$ 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 $t_{i,j,k} > 0$, sim. for j, k

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

Outline of the proof

Define: $F : \mathbb{R}^m_+ \times \mathbb{R}^n_+ \times \mathbb{R}^l_+ \to \mathbb{R}^m_+ \times \mathbb{R}^n_+ \times \mathbb{R}^l_+$:

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

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

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

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

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

$\mathcal{T} = [t_{i,j,k}]$ 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 $t_{i,j,k} > 0$, 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

Numerical counterexamples

Numerical counterexamples

$\mathcal{F} := [f_{i,j,k}] \in \mathbb{R}_+^{2\times 2\times 2}$: $f_{1,1,1} = f_{2,2,2} = a > 0$ otherwise, $f_{i,j,k} = b > 0$.

$$f(\mathbf{x}, \mathbf{y}, \mathbf{z}) = b(x_1 + x_2)(y_1 + y_2)(z_1 + z_2) + (a - b)(x_1 y_1 z_1 + x_2 y_2 z_2).$$

Numerical counterexamples

$\mathcal{F} := [f_{i,j,k}] \in \mathbb{R}_+^{2\times 2\times 2}$: $f_{1,1,1} = f_{2,2,2} = a > 0$ otherwise, $f_{i,j,k} = b > 0$.

$f(\mathbf{x}, \mathbf{y}, \mathbf{z}) = b(x_1 + x_2)(y_1 + y_2)(z_1 + z_2) + (a - b)(x_1 y_1 z_1 + x_2 y_2 z_2)$.

For $p_1 = p_2 = p_3 = p > 1$ positive singular vectors:
$\mathbf{x} = \mathbf{y} = \mathbf{z} = (0.5^{1/p}, 0.5^{1/p})^\top$.

Numerical counterexamples

$\mathcal{F} := [f_{i,j,k}] \in \mathbb{R}_+^{2\times 2\times 2}$: $f_{1,1,1} = f_{2,2,2} = a > 0$ otherwise, $f_{i,j,k} = b > 0$.

$f(\mathbf{x}, \mathbf{y}, \mathbf{z}) = b(x_1 + x_2)(y_1 + y_2)(z_1 + z_2) + (a - b)(x_1 y_1 z_1 + x_2 y_2 z_2)$.

For $p_1 = p_2 = p_3 = p > 1$ positive singular vectors:
$\mathbf{x} = \mathbf{y} = \mathbf{z} = (0.5^{1/p}, 0.5^{1/p})^\top$.

For $a = 1.2, b = 0.2$ and $p = 2$ additional positive singular vectors:
$\mathbf{x} = \mathbf{y} = \mathbf{z} \approx (0.9342, 0.3568)^\top$,
$\mathbf{x} = \mathbf{y} = \mathbf{z} \approx (0.3568, 0.9342)^\top$.

Numerical counterexamples

$\mathcal{F} := [f_{i,j,k}] \in \mathbb{R}_+^{2\times 2\times 2}$: $f_{1,1,1} = f_{2,2,2} = a > 0$ otherwise, $f_{i,j,k} = b > 0$.

$f(\mathbf{x}, \mathbf{y}, \mathbf{z}) = b(x_1 + x_2)(y_1 + y_2)(z_1 + z_2) + (a - b)(x_1 y_1 z_1 + x_2 y_2 z_2)$.

For $p_1 = p_2 = p_3 = p > 1$ positive singular vectors:
$\mathbf{x} = \mathbf{y} = \mathbf{z} = (0.5^{1/p}, 0.5^{1/p})^\top$.

For $a = 1.2, b = 0.2$ and $p = 2$ additional positive singular vectors:
$\mathbf{x} = \mathbf{y} = \mathbf{z} \approx (0.9342, 0.3568)^\top$,
$\mathbf{x} = \mathbf{y} = \mathbf{z} \approx (0.3568, 0.9342)^\top$.

For $a = 1.001, b = 0.001$ and $p = 2.99$ additional positive singular vectors:
$\mathbf{x} = \mathbf{y} = \mathbf{z} \approx (0.9667, 0.4570)^\top$,
$\mathbf{x} = \mathbf{y} = \mathbf{z} \approx (0.4570, 0.9667)^\top$

Assume F completely irreducible

Fixed point iterations for $p \geq 3$

Assume F completely irreducible
To find the fixed point iterate:

$$(\mathbf{x}_k, \mathbf{y}_k, \mathbf{z}_k) := F(\mathbf{x}_{k-1}, \mathbf{y}_{k-1}, \mathbf{z}_{k-1})$$

and renormalize $\mathbf{x}_k, \mathbf{y}_k, \mathbf{z}_k$ to have ℓ_p norm to be one

Fixed point iterations for $p \geq 3$

Assume F completely irreducible
To find the fixed point iterate:

$$(\mathbf{x}_k, \mathbf{y}_k, \mathbf{z}_k) := F(\mathbf{x}_{k-1}, \mathbf{y}_{k-1}, \mathbf{z}_{k-1})$$

and renormalize $\mathbf{x}_k, \mathbf{y}_k, \mathbf{z}_k$ to have ℓ_p norm to be one

This iterations should converge geometrically to unique positive singular vectors

Scaling of nonnegative matrices to matrices with given row and column sums

Scaling of nonnegative matrices to matrices with given row and column sums

$0 \leq A = [a_{ij}] \in \mathbb{R}_+^{m\times n}$ has row and column and sums:
$\mathbf{r} = (r_1, \ldots, r_m)^\top, \mathbf{c} = (c_1, \ldots, c_n)^\top$,:

Scaling of nonnegative matrices to matrices with given row and column sums

$0 \le A = [a_{ij}] \in \mathbb{R}_+^{m\times n}$ has row and column and sums:
$\mathbf{r} = (r_1, \ldots, r_m)^\top, \mathbf{c} = (c_1, \ldots, c_n)^\top$,:
$\sum_j a_{ij} = r_i > 0,\ \sum_i a_{ij} = c_j > 0.$
$\sum_{i=1}^m r_i = \sum_{j=1}^n c_j$

Scaling of nonnegative matrices to matrices with given row and column sums

$0 \le A = [a_{ij}] \in \mathbb{R}_+^{m\times n}$ has row and column and sums:
$\mathbf{r} = (r_1, \dots, r_m)^\top, \mathbf{c} = (c_1, \dots, c_n)^\top$,:
$\sum_j a_{ij} = r_i > 0, \ \sum_i a_{ij} = c_j > 0.$
$\sum_{i=1}^m r_i = \sum_{j=1}^n c_j$

Find nec. and suf. conditions for scaling:
$A' = [a_{ij} e^{x_i + y_j}], \mathbf{x} \in \mathbb{R}^m, \mathbf{y} \in \mathbb{R}^n$ such that A' has given row, column

Scaling of nonnegative matrices to matrices with given row and column sums

$0 \le A = [a_{ij}] \in \mathbb{R}_+^{m\times n}$ has row and column and sums:
$\mathbf{r} = (r_1, \ldots, r_m)^\top, \mathbf{c} = (c_1, \ldots, c_n)^\top$,:
$\sum_j a_{ij} = r_i > 0, \ \sum_i a_{ij} = c_j > 0.$
$\sum_{i=1}^m r_i = \sum_{j=1}^n c_j$

Find nec. and suf. conditions for scaling:
$A' = [a_{ij}e^{x_i+y_j}], \mathbf{x} \in \mathbb{R}^m, \mathbf{y} \in \mathbb{R}^n$ such that A' has given row, column

Most known case: $m = n, \mathbf{r} = \mathbf{c} = \mathbf{1}$:
rescaling to doubly stochastic matrix

Scaling of nonnegative matrices to matrices with given row and column sums

$0 \le A = [a_{ij}] \in \mathbb{R}_+^{m\times n}$ has row and column and sums:
$\mathbf{r} = (r_1, \dots, r_m)^\top, \mathbf{c} = (c_1, \dots, c_n)^\top,$:
$\sum_j a_{ij} = r_i > 0, \ \sum_i a_{ij} = c_j > 0.$
$\sum_{i=1}^m r_i = \sum_{j=1}^n c_j$

Find nec. and suf. conditions for scaling:
$A' = [a_{ij} e^{x_i+y_j}], \mathbf{x} \in \mathbb{R}^m, \mathbf{y} \in \mathbb{R}^n$ such that A' has given row, column

Most known case: $m = n, \mathbf{r} = \mathbf{c} = \mathbf{1}$:
rescaling to doubly stochastic matrix

A: completely irreducible: $A \ne A_1 \oplus A_2, A_1 \in \mathbb{R}_+^{n_1\times n_2}$:

Scaling of nonnegative matrices to matrices with given row and column sums

$0 \le A = [a_{ij}] \in \mathbb{R}_+^{m\times n}$ has row and column and sums:
$\mathbf{r} = (r_1, \dots, r_m)^\top, \mathbf{c} = (c_1, \dots, c_n)^\top,$:
$\sum_j a_{ij} = r_i > 0,\ \sum_i a_{ij} = c_j > 0.$
$\sum_{i=1}^m r_i = \sum_{j=1}^n c_j$

Find nec. and suf. conditions for scaling:
$A' = [a_{ij}e^{x_i+y_j}], \mathbf{x} \in \mathbb{R}^m, \mathbf{y} \in \mathbb{R}^n$ such that A' has given row, column

Most known case: $m = n, \mathbf{r} = \mathbf{c} = \mathbf{1}$:
rescaling to doubly stochastic matrix

A: completely irreducible: $A \ne A_1 \oplus A_2, A_1 \in \mathbb{R}_+^{n_1\times n_2}$:
Exists permutation matrices P, Q: PAQ has positive diagonal and is irreducible

Scaling of nonnegative matrices II

Thm $A \in \mathbb{R}_+^{n\times n}$ rescable to d.s. iff there exists a d.s. matrix with the same zero pattern as A.

Scaling of nonnegative matrices II

Thm $A \in \mathbb{R}_+^{n\times n}$ rescable to d.s. iff there exists a d.s. matrix with the same zero pattern as A.

Sinkhorn algorithm 1964-67: Rescaling rows and columns converges geometrically to d.s.

Menon-1968 $A \in \mathbb{R}_+^{m\times n}$ rescable to have $\mathbf{r}, \mathbf{c}$ row and column sums if there exists a matrix $B \in \mathbb{R}_+^{m\times n}$ with the same zero pattern as A and $\mathbf{r}, \mathbf{c}$ row and column sums.
The rescaled matrix is unique

Scaling of nonnegative matrices II

Thm $A \in \mathbb{R}_+^{n\times n}$ rescable to d.s. iff there exists a d.s. matrix with the same zero pattern as A.

Sinkhorn algorithm 1964-67: Rescaling rows and columns converges geometrically to d.s.

Menon-1968 $A \in \mathbb{R}_+^{m\times n}$ rescable to have $\mathbf{r}, \mathbf{c}$ row and column sums if there exists a matrix $B \in \mathbb{R}_+^{m\times n}$ with the same zero pattern as A and $\mathbf{r}, \mathbf{c}$ row and column sums.
The rescaled matrix is unique

Brualdi 1968: $A \in \mathbb{R}_+^{m\times n}$, A completely irreducible
$A \neq A_1 \oplus A_2, A_1 \in \mathbb{R}^{m_1\times n_1}$.
$PAQ = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}$ then the columns sums of $\mathbf{c}$ corresponding to the columns of A_{11} are strictly less then the row sums of $\mathbf{r}$ of the rows of A_{11}.

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 row, column and depth sums:
$\mathbf{r} = (r_1, \ldots, r_m)^\top, \mathbf{c} = (c_1, \ldots, c_n)^\top, \mathbf{d} = (d_1, \ldots, d_l)^\top > \mathbf{0}$:

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 row, column and depth sums:
$\mathbf{r} = (r_1, \ldots, r_m)^\top, \mathbf{c} = (c_1, \ldots, c_n)^\top, \mathbf{d} = (d_1, \ldots, d_l)^\top > \mathbf{0}$:
$\sum_{j,k} t_{i,j,k} = r_i > 0,\ \sum_{i,k} t_{i,j,k} = c_j > 0, \sum_{i,j} t_{i,j,k} = d_k > 0$
$\sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j = \sum_{k=1}^{l} d_k$

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

In general $\mathcal{T}$ is rescalable if I. and II. imply
$x_i + y_j + z_k = 0$ if $t_{i,j,k} > 0$

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

In general $\mathcal{T}$ is rescalable if I. and II. imply
$x_i + y_j + z_k = 0$ if $t_{i,j,k} > 0$

$\mathcal{T}$ is rescalable iff there exists $\mathcal{T}' \in \mathbb{R}_+^{m\times n\times l}$ with $\mathbf{r}, c, \mathbf{d}$ row,column,depth sums, and the same zero pattern as $\mathcal{T}$.

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

In general $\mathcal{T}$ is rescalable if I. and II. imply
$x_i + y_j + z_k = 0$ if $t_{i,j,k} > 0$

$\mathcal{T}$ is rescalable iff there exists $\mathcal{T}' \in \mathbb{R}_+^{m \times n \times l}$ with $\mathbf{r}, c, \mathbf{d}$ row,column,depth sums, and the same zero pattern as $\mathcal{T}$.

Are variants Brualdi theorem hold in the tensor case?

Rescaling versus Newton method

Rescaling versus Newton method

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

Rescaling versus Newton method

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

Newton method works, since the scaling problem is equivalent finding the unique minimum of strict convex function (Use Armijo rule or first rescaling rows and columns)

Rescaling versus Newton method

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

Newton method works, since the scaling problem is equivalent finding the unique minimum of strict convex function (Use Armijo rule or first rescaling rows and columns)

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

Rescaling versus Newton method

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

Newton method works, since the scaling problem is equivalent finding the unique minimum of strict convex function (Use Armijo rule or first rescaling rows and columns)

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

True for matrices too

References I

- R.A. Brualdi, Convex sets of nonnegative matrices, *Canad. J. Math* 20(1968), 144-157.
- R.A. Brualdi, S.V. Parter and H. Schneider, The diagonal equivalence of a nonnegative matrix to a stochastic matrix, *J. Math. Anal. Appl.* 16 (1966), 31–50.
- K.C. Chang, K. Pearson, and T. Zhang, Perron-Frobenius theorem for nonnegative tensors, *Commun. Math. Sci.* 6, (2008), 507-520.
- S. Friedland, Positive diagonal scaling of a nonnegative tensor to one with prescribed slice sums, *arXiv:0908.2368v1*.
- S. Friedland, S. Gauber and L. Han, Perron-Frobenius theorem for nonnegative multilinear forms, *arXiv:0905.1626*.
- S. Gaubert and J. Gunawardena, The Perron-Frobenius theorem for homogeneous, monotone functions, *Trans. Amer. Math. Soc.* 356 (2004), 4931-4950.

References II

- L.H. Lim, Singular values and eigenvalues of tensors: a variational approach, *CAMSAP* 05, 1 (2005), 129-132.
- M.V. Menon, Matrix links, an extremisation problem and the reduction of a nonnegative matrix to one with with prescribed row and column sums,*Canad. J. Math* 20 (1968), 225-232.
- M.V. Menon and H. Schneider, The spectrum of a nonlinear operator associated with a matrix, *Linear Algebra Appl.* 321–334.
- R. Nussbaum, Hilbert's projective metric and iterated nonlinear maps, *Memoirs Amer. Math. Soc.*, 1988, vol. 75.
- R. Sinkhorn and P. Knopp, Concenring nonnegative matrices and doubly stochastic matrices, *Pac. J. Math.* 21 (1967), 343-348.