From nonnegative matrices to nonnegative tensors

Shmuel Friedland
Univ. Illinois at Chicago

Workshop on Linear Algebra & Applications, 17 October, 2011
Hamilton Institute, National University of Ireland

Overview

1. Perron-Frobenius theorem for *irreducible* nonnegative matrices
2. Irreducibility and weak irreducibility for tensors
3. Perron-Frobenius for irreducible tensors
4. Kingman, Karlin-Ost and Friedland inequalities
5. Maxplus eigenvalues of tensors
6. Friedland-Karlin characterization of spectral radius
7. Diagonal scaling
8. Rank one approximation of tensors
9. Nonnegative multilinear forms

Nonnegative irreducible and primitive matrices

$A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$ induces digraph $DG(A) = DG = (V, E)$

$V = [n] := \{1, \ldots, n\},\ E \subseteq [n] \times [n],\ (i,j) \in E \iff a_{ij} > 0$

DG strongly connected, SC,
if for each pair $i \neq j$ there exists a dipath from i to j

Claim: DG SC iff for each $\emptyset \neq I \subset [n]$
$\exists j \in [n] \setminus I$ and $i \in I$ s.t. $(i,j) \in E$

A -primitive if $A^N > 0$ for some $N > 0 \iff A^N(\mathbb{R}_+^n \setminus \{\mathbf{0}\}) \subset \mathrm{int}\, \mathbb{R}_+^n$

A primitive $\iff$ A irreducible and g.c.d of all cycles in $DG(A)$ is one

Perron-Frobenius theorem

PF: $A \in \mathbb{R}_+^n$ irreducible. Then $0 < \rho(A) \in \mathrm{spec}\,(A)$ algebraically simple $\mathbf{x}, \mathbf{y} > \mathbf{0}$ $A\mathbf{x} = \rho(A)\mathbf{x}, A^\top \mathbf{y} = \rho(A)\mathbf{y}$.

$A \in \mathbb{R}_+^{n \times n}$ primitive iff in addition to above $|\lambda| < \rho(A)$ for $\lambda \in \mathrm{spec}\,(A) \setminus \{\rho(A)\}$

Collatz-Wielandt:

$$\rho(A) = \min_{\mathbf{x}>\mathbf{0}} \max_{i \in [n]} \frac{(A\mathbf{x})_i}{x_i} = \max_{\mathbf{x}>\mathbf{0}} \min_{i \in [n]} \frac{(A\mathbf{x})_i}{x_i}$$

Irreducibility and weak irreducibility of nonnegative tensors

$\mathcal{F} := [f_{i_1,\ldots,i_d}]_{i_1=\ldots=i_d}^{n} \in (\mathbb{C}^n)^{\otimes d}$ is called d-cube tensor, $(d \geq 3)$

$\mathcal{F} \geq 0$ if all entries are nonnegative

$\mathcal{F}$ irreducible: for each $\emptyset \neq I \subsetneq [n]$, there exists $i \in I, j_2, \ldots, j_d \in J := [n] \setminus I$ s.t. $f_{i,j_2,\ldots,j_d} > 0$.

$D(\mathcal{F})$ digraph $([n], A)$: $(i,j) \in A$ if there exists $j_2, \ldots j_d \in [n]$ s.t. $f_{i,j_2,\ldots,j_d} > 0$ and $j \in \{j_2, \ldots, j_d\}$.

$\mathcal{F}$ weakly irreducible if $D(\mathcal{F})$ is strongly connected.

Claim: irreducible implies weak irreducible

For $d = 2$ irreducible and weak irreducible are equivalent

Example of weak irreducible and not irreducible $n = 2, d = 3$,
$f_{1,1,2}, f_{1,2,1}, f_{2,1,2}, f_{2,2,1} > 0$
and all other entries of $\mathcal{F}$ are zero

Perron-Frobenius theorem for nonnegative tensors I

$\mathcal{F} = [f_{i_1,\ldots,i_d}] \in (\mathbb{C}^n)^{\otimes d}$ maps $\mathbb{C}^n$ to itself

$(\mathcal{F}\mathbf{x})_i = f_{i,\bullet}\mathbf{x} := \sum_{i_2,\ldots,i_d\in[n]} f_{i,i_2,\ldots,i_d} x_{i_2}\ldots x_{i_d},\ \ i \in [n]$

Note we can assume $f_{i,i_2,\ldots,i_d}$ is symmetric in $i_2,\ldots,i_d$.

$\mathcal{F}$ has eigenvector $\mathbf{x} = (x_1,\ldots,x_n)^\top \in \mathbb{C}^n$ with eigenvalue λ:

$(\mathcal{F}\mathbf{x})_i = \lambda x_i^{d-1}$ for all $i \in [n]$

Assume: $\mathcal{F} \geq 0, (\mathcal{F}\mathbb{R}_+^n \setminus \{\mathbf{0}\}) \subseteq \mathbb{R}_+^n \setminus \{\mathbf{0}\}$

$$\mathcal{F}_1 : \Pi_n \to \Pi_n, \quad \mathbf{x} \mapsto \frac{1}{\sum_{i=1}^n (\mathcal{F}\mathbf{x})_i^{\frac{1}{d-1}}} (\mathcal{F}\mathbf{x})^{\frac{1}{d-1}}$$

Brouwer fixed point: $\mathbf{x} \gneq \mathbf{0}$ eigenvector with $\lambda > 0$ eigenvalue

Problem When there is a unique positive eigenvector with maximal eigenvalue?

Perron-Frobenius theorem for nonnegative tensors II

Theorem Chang-Pearson-Zhang 2009 [2]
Assume $\mathcal{F} \in ((\mathbb{R}^n)^{\otimes d})_+$ is irreducible.
Then there exists a unique nonnegative eigenvector which is positive with the corresponding maximum eigenvalue λ.
Furthermore the Collatz-Wielandt characterization holds

$$\lambda = \min_{\mathbf{x}>0} \max_{i\in[n]} \frac{(\mathcal{F}\mathbf{x})_i}{x_i^{d-1}} = \max_{\mathbf{x}>0} \min_{i\in[n]} \frac{(\mathcal{F}\mathbf{x})_i}{x_i^{d-1}}$$

Theorem Friedland-Gaubert-Han 2011 [5]
Assume $\mathcal{F} \in ((\mathbb{R}^n)^{\otimes d})_+$ is weakly irreducible.
Then there exists a unique positive eigenvector with the corresponding maximum eigenvalue λ.
Furthermore the Collatz-Wielandt characterization holds

Generalization of Kingman inequality: Friedland-Gaubert

Kingman's inequality: $D \subset \mathbb{R}^m$ convex,
$A : D \to \mathbb{R}_+^{n\times n}$, $A(\mathbf{t}) = [a_{ij}(\mathbf{t})]$, each $\log a_{ij}(\mathbf{t}) \in [-\infty, \infty)$ is convex,
(entrywise logconvex)
then $\log \rho(A) : D \to [-\infty, \infty)$ convex, ($\rho(A(\cdot))$ logconvex)

Generalization: $\mathcal{F} : D \to ((\mathbb{R}^n)^{\otimes d})_+$ entrywise logconvex
then $\rho(\mathcal{T}(\cdot))$ is logconvex (L. Qi & collaborators)

Proof Outline:
$\mathcal{F}^{\circ s} = [f^s_{i_1,\ldots,i_d}],\ (0^0 = 0),\ \mathcal{F} \circ \mathcal{G} = [f_{i_1,\ldots,i_d} g_{i_1,\ldots,i_d}]$

GKI: $\rho(\mathcal{F}^{\circ\alpha} \circ \mathcal{G}^{\circ\beta}) \le (\rho(\mathcal{F}))^\alpha(\rho(\mathcal{G}))^\beta,\ \alpha, \beta \ge 0, \alpha + \beta = 1\ (*)$
Assume $\mathcal{F}, \mathcal{G} > 0,\ \mathcal{F}\mathbf{x} = \rho(\mathcal{F})\mathbf{x}^{\circ(d-1)},\ \mathcal{G}\mathbf{x} = \rho(\mathcal{G})\mathbf{y}^{\circ(d-1)}$

Hölder's inequality for $p = \alpha^{-1}, q = \beta^{-1}$ yields
$((\mathcal{F}^{\circ\alpha} \circ \mathcal{G}^{\circ\beta})(\mathbf{x}^{\circ\alpha} \circ \mathbf{y}^{\circ\beta}))_i \le (\mathcal{F}\mathbf{x})_i^\alpha(\mathcal{G}\mathbf{x})_i^\beta = (\rho(\mathcal{F}))^\alpha(\rho(\mathcal{G}))^\beta(x_i^\alpha y_i^\beta)^{d-1}$
Collatz-Wielandt implies $(*)$

Karlin-Ost and Friedland inequalities-FG

$\rho(\mathcal{F}^{\circ s})^{\frac{1}{s}}$ non-increasing on $(0,\infty)$ $(*)$

Assume $\mathcal{F}>0,\ s>1$ use $\|\mathbf{y}\|_s$ non-increasing

$(\mathcal{F}^{\circ s}\mathbf{x}^{\circ s})_i^{\frac{1}{s}} \le (\mathcal{F}\mathbf{x})_i = \rho(\mathcal{F})x_i^{d-1}$

use Collatz-Wielandt

$\rho_{\rm trop}(\mathcal{F}) = \lim_{s\to\infty}\rho(\mathcal{F}^{\circ s})^{\frac{1}{s}}$ - the tropical eigenvalue of $\mathcal{F}$.

if $\mathcal{F}$ weakly irreducible then $\mathcal{F}$ has positive tropical eigenvector

$\max_{i_2,\ldots,i_d} f_{i,i_2,\ldots,i_d}x_{i_2}\ldots x_{i_d} = \rho_{\rm trop}(\mathcal{F})x_i^{d-1}, \quad i\in[n], \mathbf{x}>\mathbf{0}$

Cor:

$\rho(\mathcal{F}\circ\mathcal{G}) \le \rho(\mathcal{F}^{\frac{1}{2}}\circ\mathcal{G}^{\frac{1}{2}})^2 \le \rho(\mathcal{F})\rho(\mathcal{G})$

$\rho(\mathcal{F}\circ\mathcal{G}) \le \rho(\mathcal{F}^{\circ p})^{\frac{1}{p}}\rho(\mathcal{G}^{q})^{\frac{1}{q}},\ \frac{1}{p}+\frac{1}{q}=1$

$p=1, q=\infty \quad \Rightarrow \rho(\mathcal{F}\circ\mathcal{G}) \le \rho(\mathcal{F})\rho_{\rm trop}(\mathcal{G})$

$\mathrm{pat}(\mathcal{G})$ *pattern of* $\mathcal{G}$, tensor with $0/1$ entries obtained by replacing every non-zero entry of $\mathcal{G}$ by 1.

$\mathcal{F}=\mathrm{pat}(\mathcal{G}) \Rightarrow \rho(\mathcal{G}) \leqslant \rho(\mathrm{pat}(\mathcal{G}))\rho_{\rm trop}(\mathcal{G})$

Characterization of $\rho_{\text{trop}}(\mathcal{F})$ – I

Friedland 1986: $\rho_{\text{trop}}(A)$
is the maximum geometric average of cycle products of $A \in \mathbb{R}_+^{n\times n}$.

$D(\mathcal{F}) := ([n], Arc)$, $(i,j) \in Arc$ iff $\sum_{j_2,\ldots,j_d} f_{i,j_2,\ldots,j_d} x_{j_2} \ldots x_{j_d}$ contains x_j.
$d-1$ cycle on $[m]$ vertices is $d-1$ outregular strongly connected subdigraph $D = ([m], Arc)$ of $D(\mathcal{F})$,
i.e. the digraph adjacency matrix $A(D) = [a_{ij}]) \in \mathbb{Z}_+^{m\times m}$ of subgraph is irreducible with each row sum $d-1$.

$A(D)\mathbf{1} = (d-1)\mathbf{1}, \mathbf{v}^\top A(D) = (d-1)\mathbf{v}^\top, \mathbf{v} = (v_1,\ldots,v_m)^\top > \mathbf{0}$
probability vector
Assume for simplicity $d-1$ cycle on $[m]$
weighted-geometric average: $\prod_{i=1}^m (f_{i,j_2(i),\ldots,j_d(i)})^{v_i}$

Friedland-Gaubert: $\rho_{\text{trop}}(\mathcal{F})$ is the maximum weighted-geometric average of $d-1$ cycle products of $\mathcal{F} \in ((\mathbb{R}^n)^{\otimes d})_+$

Cor. $\rho_{\text{trop}}(\mathcal{F})$ is logconvex in entries of $\mathcal{T}$.

Characterization of $\rho_{\mathrm{trop}}(\mathcal{F})$ II

More general results Akian-Gaubert [1]
$\mathcal{Z} = (z_{i_1,\ldots,i_d}) \in ((\mathbb{R}^n)^{\otimes d})_+$ *occupation measure*:
$\sum_{i_1,\ldots,i_d} z_{i_1,\ldots,i_d} = 1$ and for all $k \in [n]$

$\sum_{i,\{j_2,\ldots,j_d\}\ni k} z_{i,j_2,\ldots,i_d} = (d-1)\sum_{m_2,\ldots,m_d} z_{k,m_2,\ldots,m_d}$

first sum is over $i \in [n]$ and all $j_2,\ldots,j_d \in [n]$ s. t. $k \in \{j_2,\ldots,j_d\}$
Def: $\mathbf{Z}_{n,d}$ all occupation measures

Thm: $\log \rho_{\mathrm{trop}}(\mathcal{F}) = \max_{\mathcal{Z}\in\mathbf{Z}_{n,d}} \sum_{j_1,\ldots,j_d\in[n]} z_{j_1,\ldots,j_d} \log f_{i_1,\ldots,i_d}$

Proof: The extreme points of occupational measures correspond to geometric average

Diagonal similarity of nonnegative tensors

$\mathcal{F} = [f_{i_1,\ldots,i_d}] \in ((\mathbb{R}^n)^{\otimes d})_+$ is diagonally similar to
$\mathcal{G} = [g_{i_1,\ldots,i_d}] \in ((\mathbb{R}^n)^{\otimes d})_+$ if

$g_{i_1,\ldots,i_d} = e^{-(d-1)t_{i_1} + \sum_{j=2}^{d} t_{i_j}} f_{i_1,\ldots,i_d}$ for some $\mathbf{t} = (t_1, \ldots, t_n)^\top \in \mathbb{R}^n$

Diagonally similar tensors have the same eigenvalues and spectral radius

generalization of Engel-Schneider [3], (Collatz-Wielandt)

$\rho_{\mathrm{trop}}(\mathcal{F}) = \inf_{(t_1,\ldots,t_n)^\top \in \mathbb{R}^n} \max_{i_1,\ldots,i_d} e^{-(d-1)t_{i_1} + \sum_{j=2}^{d} t_{i_j}} f_{i_1,\ldots,i_d}$

Generalized Friedland-Karlin inequality I

Friedland-Karlin 1975: $A \in \mathbb{R}_+^{n\times n}$ irreducible, $A\mathbf{u} = \rho(A)\mathbf{u}, A^\top \mathbf{v} = \rho(A)\mathbf{v}$,
$\mathbf{u} \circ \mathbf{v} = (u_1 v_1, \ldots, u_n v_n) > \mathbf{0}$ probability vector:
$\log \rho(\mathrm{diag}(e^{\mathbf{t}})A) \ge \log \rho(A) + \sum_{i=1}^n u_i v_i t_i$
(graph of convex function above its supporting hyperplane)

$(e^{\mathbf{t}}\mathcal{F})_{i_1,\ldots,i_d} = e^{t_{i_1}} f_{i_1,\ldots,i_d}$

GFKI: $\mathcal{F}$ is weakly irreducible.
$A := D(\mathbf{u})^{-(d-2)}\partial(\mathcal{F}\mathbf{x})(\mathbf{u}),\ A\mathbf{u} = \rho(A)\mathbf{u}, A^\top \mathbf{v} = \rho(A)\mathbf{v}$ and $\mathbf{u} \circ \mathbf{v} > \mathbf{0}$
probability vector

$\log \rho(\mathrm{diag}(e^{\mathbf{t}})\mathcal{F}) \ge \log \rho(\mathcal{F}) + \sum_{i=1}^n u_i v_i t_i$

$\mathcal{F}$ super-symmetric: $\mathcal{F}\mathbf{x} = \nabla\phi(\mathbf{x})$, ϕ homog. pol. degree d

$\log \rho(\mathrm{diag}(e^{\mathbf{t}})\mathcal{F}) \ge \log \rho(\mathcal{F}) + \sum_{i=1}^n u_i^d t_i, \quad \sum_{i=1}^n u_i^d = 1$

Generalized Friedland-Karlin inequality II

$\min_{\mathbf{x}>0} \sum_{i=1}^n u_i v_i \log \frac{(\mathcal{F}\mathbf{x})_i}{x_i^{d-1}} = \log \rho(\mathcal{F})$ $(*)$
equality iff $\mathbf{x}$ the positive eigenvector of $\mathcal{F}$.

Gen. Donsker-Varadhan: $\rho(\mathcal{F}) = \max_{\mathbf{p}\in\Pi_n} \inf_{\mathbf{x}>\mathbf{0}} \sum_{i=1}^n p_i \frac{(\mathcal{F}\mathbf{x})_i}{x_i^{d-1}}$ $(**)$
Prf: For $\mathbf{x} = \mathbf{u}$ RHS $(**) \le \rho(\mathcal{T})$.
For $\mathbf{p} = \mathbf{u} \circ \mathbf{v}$ $(*) \Rightarrow$ RHS $(**) = \rho(\mathcal{F})$.

Gen. Cohen: $\rho(\mathcal{F})$ convex in $(f_{1,\ldots,1}, \ldots, f_{n,\ldots,n})$:
$\rho(\mathcal{F} + \mathcal{D}) = \max_{\mathbf{p}\in\Pi_n}(\sum_{i=1}^n p_i d_{i,\ldots,i} + \inf_{\mathbf{x}>\mathbf{0}} \sum_{i=1}^n p_i \frac{(\mathcal{F}\mathbf{x})_i}{x_i^{d-1}})$

GFK: $\mathcal{F}$ weakly irreducible, positive diagonal, $\mathbf{u}, \mathbf{v} > 0, \mathbf{u} \circ \mathbf{v} \in \Pi_n$,
$\exists \mathbf{t}, \mathbf{s} \in \mathbb{R}^n$ s.t. $e^{t_{i_1}} f_{i_1,\ldots,i_d} e^{s_{i_2}+\ldots+s_{i_d}}$ with eigenvector $\mathbf{u}$
and $\mathbf{v}$ left eigenvector of $D(\mathbf{u})^{-(d-2)} \partial \mathcal{F}\mathbf{x}(\mathbf{u})$

PRF: Strict convex function $g(\mathbf{z}) = \sum_{i=1}^n u_i v_i (\log \mathcal{F} e^{\mathbf{z}} - (d-1) z_i)$
achieves unique minimum for some $\mathbf{z} = \log \mathbf{x}$, as $g(\partial(\mathbb{R}_+^n \setminus \{\mathbf{0}\}) = \infty$

$\mathcal{F}$ super-symmetric and $\mathbf{v} = \mathbf{u}^{d-1}$ then $\mathbf{t} = \mathbf{s}$

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: $\mathcal{T} \ne \mathcal{T}_1 \oplus \mathcal{T}_2$.
For matrices indecomposability is equivalent to strict convexity

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

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
How many distinct singular values are for a generic tensor?

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

Nonnegative multilinear forms

Associate with $\mathcal{T} = [t_{i_1,\ldots,i_d}] \in \mathbb{R}_+^{m_1\times\ldots\times m_d}$
a multilinear form $f(\mathbf{x}_1,\ldots\mathbf{x}_d) : \mathbb{R}^{m_1\times\ldots\times m_d} \to \mathbb{R}$

$f(\mathbf{x}_1,\ldots,\mathbf{x}_d) = \sum_{i_j\in[m_j],j\in[d]} t_{i_1,\ldots,i_d} x_{i_1,1}\ldots x_{i_d,d}$,
$\mathbf{x}_i = (x_{1,i},\ldots,x_{m_i,i} \in \mathbb{R}^{m_i}$

For $\mathbf{u}\in\mathbb{R}^m, p\in(0,\infty]$ let $\|\mathbf{u}\|_p := (\sum_{i=1}^m |u_i|^p)^{\frac{1}{p}}$ and
$\mathrm{S}_{p,+}^{m-1} := \{\mathbf{0}\le\mathbf{u}\in\mathbb{R}^m,\ \|\mathbf{u}\|_p = 1\}$

For $p_1,\ldots,p_d\in(1,\infty)$ critical point $(\xi_1,\ldots,\xi_d)\in\mathrm{S}_{p_1,+}^{m_1-1}\times\ldots\times\mathrm{S}_{p_d,+}^{m_d-1}$
of $f|\mathrm{S}_{p_1,+}^{m_1-1}\times\ldots\times\mathrm{S}_{p_d,+}^{m_d-1}$ satisfies Lim [4]:
$\sum t_{i_1,\ldots,i_d} x_{i_1,1}\ldots x_{i_{j-1},j-1} x_{i_{j+1},j+1}\ldots x_{i_d,d} = \lambda x_{i_j,j}^{p_j-1}$,
$i_j\in[m_j], \mathbf{x}_j\in\mathrm{S}_{m_j,+}^{p_j-1}, j\in[d]$

Perron-Frobenius theorem for nonnegative multilinear forms

Theorem- Friedland-Gauber-Han [5]
$f : \mathbb{R}^{m_1} \times \ldots \times \mathbb{R}^{m_d} \to \mathbb{R}$, a nonnegative multilinear form,
$\mathcal{T}$ weakly irreducible and $p_j \geq d$ for $j \in [d]$.
Then f has unique positive critical point on $\mathrm{S}_+^{m_1-1} \times \ldots \times \mathrm{S}_+^{m_d-1}$.
If $\mathcal{F}$ is irreducible then f has a unique nonnegative critical point which is necessarily positive

Eigenvectors of homogeneous monotone maps on $\mathbb{R}^n_+$

Hilbert metric on $\mathbb{PR}^n_{>0}$: for $\mathbf{x} = (x_1, \ldots, x_n)^\top, \mathbf{y} = (y_1, \ldots, y_n)^\top > \mathbf{0}$.
Then $\mathrm{dist}(\mathbf{x}, \mathbf{y}) = \max_{i\in[n]} \log \frac{y_i}{x_i} - \min_{i\in[n]} \log \frac{y_i}{x_i}$.

$\mathbf{F} = (F_1, \ldots, F_n)^\top : \mathbb{R}^n_{>0} \to \mathbb{R}^n_{>0}$ homogeneous:
$\mathbf{F}(t\mathbf{x}) = t\mathbf{F}(\mathbf{x})$ for $t > 0, \mathbf{x} > \mathbf{0}$, and monotone $\mathbf{F}(\mathbf{y}) \geq \mathbf{F}(\mathbf{x})$ if $\mathbf{y} \geq \mathbf{x} > \mathbf{0}$.
$\mathbf{F}$ induces $\hat{\mathbf{F}} : \mathbb{PR}^n_{>0} \to \mathbb{PR}^n_{>0}$

$\mathbf{F}$ nonexpansive with respect to Hilbert metric
$\mathrm{dist}(\mathbf{F}(\mathbf{x}), \mathbf{F}(\mathbf{y})) \leq \mathrm{dist}(\mathbf{x}, \mathbf{y})$.

$\alpha_{max}\mathbf{x} \leq \mathbf{y} \leq \beta_{\min}\mathbf{x} \Rightarrow$
$\alpha_{\max}\mathbf{F}(\mathbf{x}) = \mathbf{F}(\alpha_{\max}\mathbf{x}) \leq \mathbf{F}(\mathbf{y}) \leq \mathbf{F}(\beta_{\min}\mathbf{x}) = \beta_{\min}\mathbf{F}(\mathbf{x})$
$\Rightarrow \mathrm{dist}(\mathbf{F}(\mathbf{x}), \mathbf{F}(\mathbf{y})) \leq \log \frac{\beta_{\min}}{\alpha_{\max}} = \mathrm{dist}(\mathbf{x}, \mathbf{y})$

$\mathbf{x} > \mathbf{0}$ eigenvector of $\mathbf{F}$ if $\mathbf{F}(\mathbf{x}) = \lambda\mathbf{F}(\mathbf{x})$.
So $\mathbf{x} \in \mathbb{PR}^n_+$ fixed point of $\mathbf{F}|\mathbb{PR}^n_+$.

Existence of positive eigenvectors of $\mathbf{F}$

1. If $\mathbf{F}$ contraction: $\mathrm{dist}(\mathbf{F}(\mathbf{x}), \mathbf{F}(\mathbf{y})) \le K\mathrm{dist}(\mathbf{x}, \mathbf{y})$ for $K < 1$, then $\mathbf{F}$ has unique fixed point in $\mathbb{PR}_+^n$ and power iterations converge to the fixed point

2. Use Brouwer fixed and irreducibility to deduce existence of positive eigenvector

3. Gaubert-Gunawardena 2004:
for $u \in (0, \infty), J \subseteq [n]$ let $\mathbf{u}_J = (u_1, \ldots, u_n)^\top > \mathbf{0}$: $u_i = u$ if $i \in J$ and $u_i = 1$ if $i \notin J$. $F_i(\mathbf{u}_J)$ nondecreasing in u.
di-graph $\mathcal{G}(\mathbf{F}) \subset [n] \times [n]$: $(i, j) \in \mathcal{G}(\mathbf{F})$ iff $\lim_{u \to \infty} F_i(\mathbf{u}_{\{j\}}) = \infty$.

Thm: $\mathbf{F} : \mathbb{R}_{>0}^n \to \mathbb{R}_{>0}^n$ homogeneous and monotone. If $\mathcal{G}(\mathbf{F})$ strongly connected then $\mathbf{F}$ has positive eigenvector

Thm 2.5, Nussbaum 88: $\mathbf{F} : \mathbb{R}^n_{>0} \to \mathbb{R}^n_{>0}$ homogeneous and monotone. Assume; $\mathbf{u} > \mathbf{0}$ eigenvector $\mathbf{F}$ with the eigenvalue $\lambda > 0$, $\mathbf{F}$ is C^1 in some open neighborhood of $\mathbf{u}$, $A = \mathrm{D}\mathbf{F}(\mathbf{u}) \in \mathbb{R}^{n\times n}_+$ $\rho(A)(=\lambda)$ a simple root of $\det(xI - A)$. Then $\mathbf{u}$ is a unique eigenvector of $\mathbf{F}$ in $\mathbb{R}^n_{>0}$.

Cor 2.5, Nus88: In the above theorem assume $A = \mathrm{D}\mathbf{F}(\mathbf{u})$ is primitive.
Let $\psi \gneq \mathbf{0}, \psi^\top \mathbf{u} = 1$.
Define $\mathbf{G} : \mathbb{R}^n_{>0} \to \mathbb{R}^n_{>0}$ $\mathbf{G}(\mathbf{x}) = \frac{1}{\psi^\top \mathbf{F}(\mathbf{x})}\mathbf{F}(\mathbf{x})$
Then $\lim_{m\to\infty} \mathbf{G}^{\circ m}(\mathbf{x}) = \mathbf{u}$ for each $\mathbf{x} \in \mathbb{R}^n_{>0}$.

Outline of the uniqueness of pos. crit. point of f

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

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

References I

- M. Akian and S. Gaubert, Spectral theorem for convex monotone homogeneous maps, and ergodic control, *Nonlinear Analysis* 52 (2003), 637-679.
- K.C. Chang, K. Pearson, and T. Zhang, Perron-Frobenius theorem for nonnegative tensors, *Commun. Math. Sci.* 6, (2008), 507-520.
- G.M. Engel and H. Schneider, Diagonal similarity and equivalence for matrices over groups with 0, *Czechoslovak Mathematical Journal* 25 (1975), 389–403.
- S. Friedland, Limit eigenvalues of nonnegative matrices, *Linear Algebra Appl.* 74 (1986), 173-178.
- S. Friedland, S. Gaubert and L. Han, Perron-Frobenius theorem for nonnegative multilinear forms, *arXiv:0905.1626*, to appear in LAA.

References II

- S. Friedland and S. Karlin, Some inequalities for the spectral radius of nonnegative matrices and applications, *Duke Math. J.* 42 (1975), 459-490.
- S. Gaubert and J. Gunawardena, The Perron-Frobenius theorem for homogeneous, monotone functions, *Trans. Amer. Math. Soc.* 356 (2004), 4931-4950.
- S. Karlin and F. Ost, Some monotonicity properties of Schur powers of matrice and related inequalities, *Linear Algebra Appl.* 68 (1985), 47–65.
- L.H. Lim, Singular values and eigenvalues of tensors: a variational approach, *CAMSAP* 05, 1 (2005), 129-132.
- M. Ng, L. Qi and G. Zhou, Finding the largest eigenvalue of a nonnegative tensor, *SIAM J. Matrix Anal. Appl.* 31 (2009), 1090–1099.

References III

- R. Nussbaum, Hilbert's projective metric and iterated nonlinear maps, *Memoirs Amer. Math. Soc.*, 1988, vol. 75.
- L. Qi, Eigenvalues of a real supersymmetric tensor, *J. Symbolic Comput.* 40 (2005), 1302-1324.