# Nonlinear applications of the Birkhoff theorem

Antoine Gautier

QUANTPI

SIAM LA21 Minitutorial
Applied Nonlinear Perron-Frobenius Theory
May 18, 2021

# Nonnegative vectors and component-wise operations

$$\begin{aligned} \mathbb{R}^n_+ &= \{x \in \mathbb{R}^n \colon x_1, \ldots, x_n \geq 0\} \\ \mathbb{R}^n_{++} &= \{x \in \mathbb{R}^n \colon x_1, \ldots, x_n > 0\} \end{aligned}$$

$$\begin{aligned} x \succeq y \quad &\Leftrightarrow \quad x_i \geq y_i \quad \forall i \\ x \succ y \quad &\Leftrightarrow \quad x_i > y_i \quad \forall i \\ x \succneq y \quad &\Leftrightarrow \quad x \succeq y, \quad x \neq y \end{aligned}$$

$$x^\theta = (x_1^\theta, \ldots, x_n^\theta) \qquad \forall \theta \in \mathbb{R}, \quad x \succ 0$$

# Positive matrices

$$
\begin{aligned}
M \in \mathbb{R}^{m \times n}_{++} \quad & \Leftrightarrow \quad M_{ij} > 0 \quad && \forall i, j \\
& \Leftrightarrow \quad Mx \succ 0 \quad && \forall x \gneqq 0 \\
& \Leftrightarrow \quad Mx \succ My \quad && \forall x \gneqq y
\end{aligned}
$$

# Linear Perron-Frobenius theorem

$M \in \mathbb{R}^{n\times n}_{++}$

$$Mx = \lambda x, \qquad x \succeq 0, \qquad \|x\| = 1 \qquad (\star)$$

**Theorem** [Perron, Frobenius (1908)]

Exists unique solution $(u, \lambda)$ to $(\star)$,

$$u \succ 0 \qquad \text{and} \qquad \lambda = \rho(M)$$

and $\forall x^{(0)} \succ 0, \quad x^{(k)} = Mx^{(k-1)}$

$$\lim_{k\to\infty} \frac{x^{(k)}}{\|x^{(k)}\|} = u$$

# Computation: $\|M\|_{2,2}$, $M \in \mathbb{R}^{m\times n}_{++}$

$$\|M\|_{2,2} \quad = \quad \max_{x\neq 0} \frac{\|Mx\|_2}{\|x\|_2} \quad \leq \quad \max_{x\neq 0} \frac{\|\,|Mx|\,\|_2}{\|\,|x|\,\|_2}$$

$\Rightarrow$ $\exists$ global maximizer $u \succeq 0$

$$\nabla \frac{\|Mu\|_2}{\|u\|_2} = 0 \qquad \Leftrightarrow \qquad M^\top M u = \Big(\frac{\|Mu\|_2}{\|u\|_2}\Big)^2 u$$

$\Rightarrow$ critical point $\equiv$ eigenvector

Apply linear Perron-Frobenius theorem

# Application: $\|M\|_{2,2}$

$M \in \mathbb{R}^{m\times n}_{++}$

$$\|M\|_{2,2} \quad = \quad \max_{x\neq 0} \frac{\|Mx\|_2}{\|x\|_2}$$

**Corollary**

Exists unique $u \succeq 0$ s.t.

$$\|Mu\|_2 = \|M\|_{2,2}, \qquad \|u\|_2 = 1$$

and

$$\lim_{k\to\infty} \frac{x^{(k)}}{\|x^{(k)}\|_2} = u, \qquad x^{(k)} = M^\top M x^{(k-1)},\ x^{(0)} \succ 0$$

# Hilbert projective metric

$$d_H(x, y) \quad = \quad \max_{i,j} \ln\left(\frac{x_i}{y_i}\frac{y_j}{x_j}\right) \qquad \forall x, y \succ 0$$

$$(\{x \succ 0 \mid \|x\| = 1\}, d_H) \text{ complete metric space}$$

$$f(x) = \lambda x \qquad \Leftrightarrow \qquad d_H(f(x), x) = 0$$

$$M \in \mathbb{R}^{n\times n}_{++} \qquad \Rightarrow \qquad d_H\big(Mx, My\big) \quad \leq \quad d_H(x, y) \qquad \forall x, y \succ 0$$

# The Birkhoff-Hopf theorem

$M \in \mathbb{R}_{++}^{m \times n}$

**Theorem** [Birkhoff (1957)]

$$d_H\big(Mx, My\big) \quad \leq \quad \kappa(M)\, d_H(x, y) \qquad \forall x, y \succ 0$$

with

$$\kappa(M) = \tanh\left(\tfrac{1}{4}\ln(\Delta(M))\right) \qquad\qquad \Delta(M) = \max_{i,j,k,\ell} \frac{M_{ij} M_{k\ell}}{M_{i\ell} M_{ki}}$$

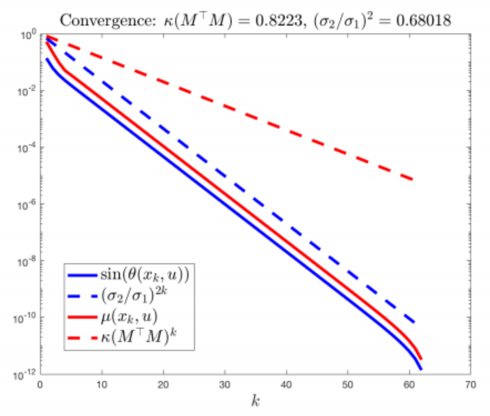
**Corollary**

$M \in \mathbb{R}_{++}^{n \times n}$

$$x^{(k)} = M^\top M x^{(k-1)}, \quad x^{(0)} \succeq 0, \quad \Rightarrow \quad \mu\left(\frac{x^{(k)}}{\|x^{(k)}\|_2}, u\right) \leq \kappa(M^\top M)^k\, c$$

# Convergence comparison

$A \in \mathbb{R}^{10\times 10}_{++}$




$$\sin(\angle(x^{(k)}, u)) \leq \left(\frac{\sigma_2}{\sigma_1}\right)^{2k} \qquad d_H(x^{(k)}, u) \leq \kappa(M^\top M)^k \, c$$

# Hilbert metric and Birkhoff ratio

$M \in \mathbb{R}_{++}^{n\times n}, \quad x, y \succ 0, \quad \theta \in \mathbb{R},$

$$d_H(x,y) \quad = \quad \max_{i,j} \ln\left(\frac{x_i}{y_i}\frac{y_j}{x_j}\right)$$

$$d_H(x^\theta, y^\theta) \quad = \quad |\theta|\, d_H(x,y)$$

**Lemma**

$$d_H\big(\,(Mx)^\theta\,,\,(My)^\theta\,\big) \quad \le \quad |\theta|\kappa(M)\, d_H(x,y)$$

# Application: Sinkhorn-Knopp theorem

$M \in \mathbb{R}^{n\times n}_{++}$

**Theorem** [Sinkhorn, Knopp (1967)]

Exists unique solution to

$$u \succ 0, \qquad (Mu)^{-1} = \lambda u, \qquad \|u\| = 1$$

and $\lim_{k\to\infty} \frac{x^{(k)}}{\|x^{(k)}\|} = u$, with

$$x^{(0)} \succ 0 \qquad x^{(k)} = (Mx^{(k-1)})^{-1}, \qquad \mu(x^{(k)}, x^{(0)}) \leq \kappa(M)^k\, C$$

$$D = \mathrm{diag}(\lambda^{1/2} u) \qquad \Rightarrow \qquad DMD \quad \text{stochastic matrix}$$

# Application: $\|M\|_{p,q}$

$M \in \mathbb{R}^{m\times n}_{++}$, $1 < p, q < \infty$

$$\|M\|_{p,q} \quad = \quad \max_{x\neq 0} \frac{\|Mx\|_q}{\|x\|_p}, \qquad\qquad \tau = \kappa(M)\kappa(M^\top)\frac{q-1}{p-1} < 1$$

**Theorem** [G., Hein, Tudisco (2021)]

Exists unique $u \succeq 0$ s.t.

$$\|Mu\|_q = \|M\|_{p,q}, \qquad \|u\|_p = 1$$

and $\lim_{k\to\infty} \frac{x^{(k)}}{\|x^{(k)}\|_p} = u$ with

$$x^{(0)} \succ 0, \qquad x^{(k)} = \big(M^\top (Mx^{(k-1)})^{p-1}\big)^{\frac{1}{q-1}} \qquad \mu(x^{(k)}, x^{(0)}) \leq \tau^k\, C$$

# Examples

$1 < p, q < \infty$, $\epsilon > 0$,

$$\|M\|_{p,q} = \max_{x \neq 0} \frac{\|Mx\|_q}{\|x\|_p}, \qquad \tau = \kappa(M)\kappa(M^\top)^{\frac{q-1}{p-1}}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 1 & 1 \end{pmatrix} \quad \Rightarrow \quad \tau = \frac{9}{400} \frac{q-1}{p-1}$$

$$B = \begin{pmatrix} \epsilon & 1 \\ 1 & \epsilon \end{pmatrix} \quad \Rightarrow \quad \tau = \left(\frac{1-\epsilon}{1+\epsilon}\right)^2 \frac{q-1}{p-1}$$

# Generalization: Composition of $p$-norms

$A, B \in \mathbb{R}^{n\times n}_{++}$, $1 < s \leq \theta \leq p < \infty$, $2 \leq q, r < \infty$

$$\|A\| = \max_{x+y\neq 0} \frac{\|Ax + Ay\|_p}{\|(\|x\|_r, \|y\|_s)\|_2} \quad \Rightarrow \quad \tau = \kappa(A)^2(p-1)$$

$$\|[A,B]\| = \max_{(x,y)\neq 0} \frac{\|Ax\|_p^\theta + \|By\|_q^\theta}{\|x\|_r^\theta + \|y\|_s^\theta} \quad \Rightarrow \quad \tau = \kappa([A,B])^2 \frac{p+q-\theta-1}{s-1}$$

$$\|B\| = \max_{\|x\|_r=1} \|A\,|Bx|^p\|_q \quad \Rightarrow \quad \tau = \kappa(B)^2 \frac{pq-1}{r-1}$$

# Multilinear Birkhoff theorem

# Symmetric positive tensors

$T \in \mathbb{R}_{++}^{n \times n \times n}$

$$T(x, y, z) = \sum_{i,j,k=1}^{n} T_{ijk} x_i y_j z_k$$

$T$ super-symmetric $\Leftrightarrow$ $T_{ijk} = T_{ikj} = T_{jik} = T_{jki} = T_{kij} = T_{kji} \quad \forall i, j, k$

# Multilinear Birkhoff-Hopf

$T \in \mathbb{R}_{++}^{n\times n\times n}$ super-symmetric,

$$\nabla_x T(\cdot, y, z)_i = \sum_{j,k=1}^n T_{ijk} y_j z_k, \qquad \nabla_{x,y}^2 T(\cdot,\cdot,z)_{i,j} = \sum_{k=1}^n T_{ijk} z_k$$

**Theorem** [G., Tudisco]

$$d_H\big(\nabla_x T(\cdot, y, z), \nabla_x T(\cdot, y', z')\big) \leq \kappa(T)\Big(d_H(y, y') + d_H(z, z')\Big)$$

with

$$\kappa(T) = \tanh\left(\tfrac{1}{4}\ln(\Delta(T))\right), \quad \Delta(T) = \sup_{z\succ 0}\max_{i,j,k,l} \frac{\nabla_{x,y}^2 T(\cdot,\cdot,z)_{i,j}\,\nabla_{x,y}^2 T(\cdot,\cdot,z)_{k,\ell}}{\nabla_{x,y}^2 T(\cdot,\cdot,z)_{i,\ell}\,\nabla_{x,y}^2 T(\cdot,\cdot,z)_{k,j}}$$

# Application: Tensor singular values

**Lemma**

$$1 < T_{ijk} < 3, \qquad \forall i,j,k \qquad \Rightarrow \qquad \kappa(T) < \frac{1}{2}$$

$$\begin{cases} \nabla_x T(\cdot, y, z) &= \lambda_1 x^{\alpha} \\ \nabla_y T(x, \cdot, z) &= \lambda_2 y^{\alpha} \\ \nabla_z T(x, y, \cdot) &= \lambda_3 z^{\alpha} \end{cases} \qquad x, y, z \succ 0$$

$$|\alpha| \leq 1 \qquad \Rightarrow \qquad \text{strict contraction}$$

$$\mu((x,y,z),(x',y',z')) = d_H(x,x') + d_H(y,y') + d_H(z,z')$$

# Application: $\|T\|_{2,2,2}$ (NP-hard)

$T \in \mathbb{R}^{n\times n\times n}, \ T(x,y,z) = \sum_{i,j,k=1}^{n} T_{ijk} x_i y_j z_k$

$$\|T\|_{2,2,2} = \max_{x,y,z\neq 0} \frac{T(x,y,z)}{\|x\|_2\|y\|_2\|z\|_2} \qquad (\star)$$

$$\begin{cases} \nabla_x T(\cdot,y,z) &= \lambda_1 x \\ \nabla_y T(x,\cdot,z) &= \lambda_2 y \\ \nabla_z T(x,y,\cdot) &= \lambda_3 z \end{cases} \qquad x,y,z \succ 0$$

**Corollary** [G., Tudisco (2019)]

$1 < T_{ijk} < 3, \quad \forall i,j,k \quad \Rightarrow \quad (\star)$ computable with linear convergence rate

# Application: Discrete generalized Schrödinger equation

$T \in \mathbb{R}^{n\times n\times n}$, $T(x, y, z) = \sum_{i,j,k=1}^{n} T_{ijk} x_i y_j z_k$

$$\begin{cases} \nabla_x T(\cdot, y, z) &= \lambda_1 x^{-1} \\ \nabla_y T(x, \cdot, z) &= \lambda_2 y^{-1} \\ \nabla_z T(x, y, \cdot) &= \lambda_3 z^{-1} \end{cases} \qquad x, y, z \succ 0 \qquad (\star)$$

**Corollary** [G., Tudisco (2019)]

$1 < T_{ijk} < 3, \quad \forall i, j, k \qquad \Rightarrow \qquad (\star)$ computable with linear convergence rate

# Tightness

$$M \in \mathbb{R}_{++}^{2\times 2}, \quad T \in \mathbb{R}_{++}^{2\times 2\times 2}, \quad 0 < \epsilon < 1, \quad \alpha \geq 0$$

$$M_{ij} = \begin{cases} 1 & \text{if } i = j \\ \epsilon & \text{otherwise} \end{cases} \qquad T_{ijk} = \begin{cases} 1 & \text{if } i = j = k \\ \epsilon & \text{otherwise} \end{cases} \qquad \forall i, j, k = 1, 2$$

**Lemma**

$$(Mu)^\alpha = \lambda u \qquad u \succ 0$$

$$\kappa(M)|\alpha| \leq 1 \quad \Leftrightarrow \quad \text{unique solution} \quad \Leftrightarrow \quad \text{convergence of power method}$$

$$\big(\nabla_x T(\cdot, v, v)\big)^\alpha = \lambda v \qquad v \succ 0$$

$$2\kappa(T)|\alpha| \leq 1 \quad \Leftrightarrow \quad \text{unique solution} \quad \Leftrightarrow \quad \text{convergence of power method}$$

Thank you

https://ftudisco.github.io/siam-nonlinear-pf-tutorial/