

# Nonlinear Perron-Frobenius Theory: Part 1

Bas Lemmens  
University of Kent (UK)

# Perron-Frobenius Theorem

If  $A$  is a nonnegative irreducible  $n \times n$  matrix, then the following assertions hold:

- (i)  $r(A)$  is a simple eigenvalue of  $A$  and the corresponding normalised eigenvector  $v$  is unique and positive.
- (ii) Any nonnegative eigenvector of  $A$  is a multiple of  $v$ .
- (iii) If, in addition,  $A$  has exactly  $q$  eigenvalues  $\lambda$  with  $|\lambda| = r(A)$ , then these eigenvalues are given by

$$r(A)e^{2\pi ik/q} \quad \text{for } 0 \leq k < q.$$

# Kreĭn-Rutman Theorem

A convex subset  $K$  of  $V$  is called a **cone** if

$$\mu K \subseteq K \text{ for all } \mu \geq 0 \quad \text{and} \quad K \cap (-K) = \{0\}.$$

Theorem Let  $K$  be a closed cone in Banach space  $V$  such that  $(K - K) = V$ . If  $A: V \rightarrow V$  is a continuous compact linear map that leaves  $K$  invariant and  $r(A) > 0$ , then there exists  $v \in K \setminus \{0\}$  such that  $Av = r(A)v$ .

(In these lectures  $V$  will be mostly **finite** dimensional)

# Special cones

1.  $\mathbb{R}_+^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_i \geq 0 \text{ for } 1 \leq i \leq n\}$ , **standard positive cone**.
2.  $C(K)_+ = \{f: K \rightarrow \mathbb{R} : f(x) \geq 0 \text{ for all } x \in K\}$ , where  $K$  is a compact Hausdorff space.
3. **Polyhedral cones**, i.e., ones that are the intersection of finitely many closed half-spaces.
4.  $\Lambda_{n+1} = \{x \in \mathbb{R}^{n+1} : x_1^2 - x_2^2 - \dots - x_{n+1}^2 \geq 0 \text{ and } x_1 \geq 0\}$ , **Lorentz cone**.
5. The cone of **positive-semidefinite matrices**,  $\Pi_n(\mathbb{R})$ , in the vector space of  $n \times n$  symmetric matrices.
6. The cone of **positive-semidefinite self-adjoint operators**,  $B(H)_{\text{sa}}^+$ , in the real vector space of self-adjoint operators on a Hilbert space  $H$ .

# Irreducibility

A linear map  $A: V \rightarrow V$  that leaves a solid closed cone  $K$  invariant is said to be **irreducible** if the only faces which are left invariant by  $A$  are  $\{0\}$  and  $K$ .

**Theorem (Vandergraft)** If  $A: V \rightarrow V$  is an irreducible linear map that leaves a solid closed cone  $K$  invariant, then the following assertions hold:

- (i)  $r(A)$  is a simple eigenvalue of  $A$  and the corresponding normalised eigenvector  $v$  is unique and lies in  $K^\circ$ .
- (ii) Any nonnegative eigenvector of  $A$  is a multiple of  $v$ .

The 3<sup>rd</sup> part of the Perron-Frobenius Theorem cannot be extended to arbitrary cones!

# Peripheral spectrum

**Theorem** If  $A: V \rightarrow V$  is an irreducible linear map that leaves a **polyhedral** cone  $K$  with  $N$  facets invariant, then there exists an integer  $q \geq 1$  such that each  $\lambda \in \sigma(A)$  with  $|\lambda| = r(A)$  satisfies  $\lambda^q - r(A)^q = 0$ .

**Open Problem** What is an optimal upper bound on  $q$  in terms of the number of facets and extreme rays of  $K$ ?

**Corollary** If  $A: V \rightarrow V$  is a linear map that leaves a polyhedral cone  $K$  with  $N$  facets invariant, then there exists an integer  $p \geq 1$  such that

$$\lim_{k \rightarrow \infty} A^{kp} x$$

exists for each  $x \in K$  with  $(\|A^k x\|)_k$  bounded. Moreover,  $p$  is the order of a permutation on  $N$  letters.

# Order-preserving maps

A cone  $K \subseteq V$  induces a **partial ordering**  $\leq$  on  $V$  by

$$x \leq y \quad \text{if and only if} \quad y - x \in K.$$

For **linear** maps  $A: V \rightarrow V$ ,

$$A(K) \subseteq K \quad \iff \quad Ax \leq Ay \text{ whenever } x \leq y.$$

A map  $f: X \rightarrow V$ , with  $X \subseteq V$ , is called **order-preserving** if

$$x \leq y \text{ implies } f(x) \leq f(y) \text{ for } x, y \in X.$$

# Nonlinear Perron-Frobenius

1. Is there a sensible definition of the **spectral radius** for order-preserving maps  $f: K \rightarrow K$ , and does there exist a corresponding eigenvector?
2. When does an order-preserving map have an **eigenvector in the interior of the cone**, and when is it unique?
3. How do the orbits,

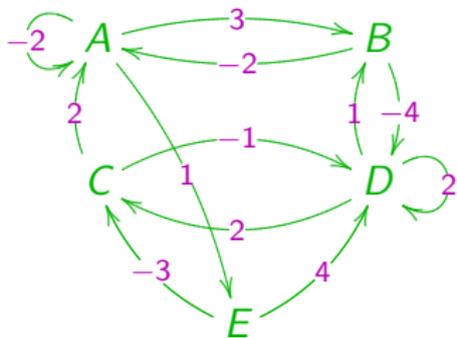
$$x, f(x), f^2(x) = f(f(x)), f^3(x) = f(f(f(x))), \dots,$$

of an order-preserving map  $f$  behave in the long-term?

4. When does an order-preserving map have the property that every bounded orbit converges to a periodic orbit?

Strikingly detailed answers exist for a variety of classes of order-preserving maps.

# Stochastic game



1. Two players MIN and MAX who move a token along one of the outgoing arcs.
2. To determine whose turn it is a coin is flipped. If head it is MAX's turn, and if tail it is MIN's turn.
3. MIN pays to MAX the weight of the arc along which the token is moved (zero-sum).

# The value of a game

If we fix an **initial state**  $i$  for the token and the **number of rounds**  $k$ , then the game has a **value**,  $V_i^k$ .

$V_i^k$  = least amount MAX can expect to win  
independent of actions of MIN  
= largest amount MIN can expect to lose  
independent of actions of MAX

$$V_i^k = \frac{1}{2} \left( \max_{i \rightarrow j} (w_{ij} + V_j^{k-1}) + \min_{i \rightarrow j} (w_{ij} + V_j^{k-1}) \right),$$

where  $V_j^0 = 0$  for all  $j$ .

# Shapley operator

The dynamic programming operator  $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}^n$  of a stochastic game is called the **Shapley operator**.

In this case

$$\Psi_i(x) = \frac{1}{2} \left( \max_{i \rightarrow j} (w_{ij} + x_j) + \min_{i \rightarrow j} (w_{ij} + x_j) \right).$$

Note that

$$x \leq y \text{ implies } \Psi(x) \leq \Psi(y)$$

and

$$\Psi(x + (\lambda, \dots, \lambda)) = \Psi(x) + (\lambda, \dots, \lambda) \text{ for all } x \in \mathbb{R}^n \text{ and } \lambda \in \mathbb{R}.$$

# Topical maps

A map  $h: \mathbb{R}^n \rightarrow \mathbb{R}^n$  is called **topical** if it is **order-preserving** with respect to  $\mathbb{R}_+^n$  and **additively homogeneous**, i.e.,

$$h(x + \lambda \mathbf{1}) = h(x) + \lambda \mathbf{1} \quad \text{for all } x \in \mathbb{R}^n \text{ and } \lambda \in \mathbb{R}.$$

Here  $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^n$ .

1. Max-plus maps,  $h: \mathbb{R}^n \rightarrow \mathbb{R}^n$ , where  $h_i(x) = \max_j a_{ij} + x_j$ .
2. Bellmann operators in Markov decision processes,  $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ , where

$$T_i(x) = \max_{\alpha} r_i(\alpha) + \sum_j p_{ij}(\alpha) x_j.$$

3. Shapley operators in stochastic games,  $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ , where

$$\Psi_i(x) = \max_{\alpha \in A_i} \min_{\beta \in B_j} r_i(\alpha, \beta) + \sum_j p_{ij}(\alpha, \beta) x_j.$$

# Log-Exp Transform

Given a topical map  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ , the coordinate-wise exp and ln functions yield commutative diagram:

$$\begin{array}{ccccc} & \mathbb{R}^n & \xrightarrow{f} & \mathbb{R}^n & \\ \text{exp} & \downarrow & & \downarrow & \text{exp} \\ & (\mathbb{R}_+^n)^\circ & \xrightarrow{h} & (\mathbb{R}_+^n)^\circ & \end{array}$$

The induced map  $h: (\mathbb{R}_+^n)^\circ \rightarrow (\mathbb{R}_+^n)^\circ$  is order-preserving with respect to  $\mathbb{R}_+^n$  and **homogeneous (of degree 1)**, so

$$h(\lambda x) = \lambda h(x) \quad \text{for all } x \in (\mathbb{R}_+^n)^\circ \text{ and } \lambda \geq 0.$$

# Homogeneous maps

1. Means and their generalisations, e.g., the Gauss arithmetic-geometric mean:

$$f(x) = ((x_1 + x_2)/2, \sqrt{x_1 x_2}).$$

Mathematical biology/demography

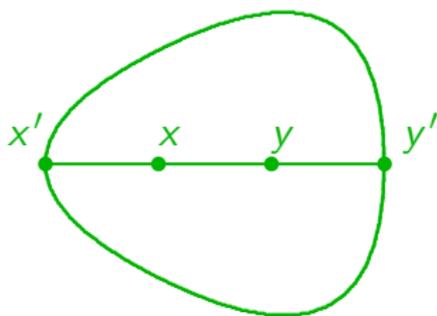
2. Nonlinear matrix equations, e.g.,

$$f(X) = B^* X B + (\text{tr}(A X) X)^{1/2}$$

on  $\Pi_n(\mathbb{R})$ , where  $A \in \Pi_n(\mathbb{R})$ .

3. *DAD*-operators in matrix scaling problems.
4. Reproduction-decimation operators in the theory of fractal diffusions, nonnegative tensors, etc

# Hilbert geometries



$\Omega$  an **open bounded convex set** in a finite dim. vector space  $V$ .

On  $\Omega$  **Hilbert's metric** is defined by

$$\delta(x, y) = \log \left( \frac{|x'y|}{|x'x|} \cdot \frac{|y'x|}{|y'y|} \right) \quad \text{for } x \neq y \in \Omega,$$

and  $\delta(x, x) = 0$  for all  $x \in \Omega$ .

# Geometric properties

Ellipsoids	strictly convex sets	partly flat partly curved	polytopes	simplices
$(X, \frac{1}{2}\delta)$ $\simeq$ $\mathbb{H}^n$	Gromov hyperbolic spaces	$\{A \in \Pi_n(\mathbb{R})^\circ : \text{tr}(A) = 1\}$	$(X, \delta)$ $\xrightarrow{\text{isom}}$ $(\mathbb{R}^m, \ \cdot\ _\infty)$	$(\Delta_n, \delta)$ $\simeq$ $(\mathbb{R}^n, \ \cdot\ )$
uniquely geodesic		non-uniquely geodesic		

$$\|x\|_\infty = \max_i |x_i| \quad (\text{sup-norm}).$$

## Parts

For  $x, y \in C$  we say  $x \sim_C y$  if there exist  $0 < \alpha \leq \beta$  such that

$$\alpha y \leq x \leq \beta y.$$

Easy to check that  $\sim_C$  is an equivalence relations on  $C$ .

The equivalence classes are called the parts of  $C$ .

The interior of  $C$  is a part, if it is non-empty.

For  $x \sim_C y$  define

$$M(x/y) = \inf\{\beta > 0 : x \leq \beta y\}$$

and

$$m(x/y) = \sup\{\alpha > 0 : \alpha y \leq x\}.$$

# Hilbert's metric on cones

**Hilbert's (projective) metric** on  $C$  is given by,

$$d_H(x, y) = \log \left( \frac{M(x/y)}{m(x/y)} \right) \quad \text{for all } x \sim_C y.$$

(Pioneered by Garrett Birkhoff and Hans Samelson)

Note that

$$d_H(\lambda x, \mu y) = d_H(x, y) \quad \text{for all } \lambda, \mu > 0 \text{ and } x \sim_C y.$$

Hilbert's metric is a genuine metric on the set of rays in each part of  $C$ , if  $C$  is a closed cone in a normed space.

## On $C(K)_+$

Suppose that  $f(t), g(t) > 0$  for all  $t \in K$ . Then

$$f \leq \beta g \quad \text{if and only if} \quad \frac{f(t)}{g(t)} \leq \beta \quad \text{for all } t .$$

So,

$$M(f/g) = \sup_t \frac{f(t)}{g(t)}$$

and

$$d_H(f, g) = \log \sup_t \frac{f(t)}{g(t)} - \log \inf_t \frac{f(t)}{g(t)} .$$

## On $B(H)_{\text{sa}}^+$

Suppose  $A, B$  in  $B(H)_{\text{sa}}^+$  invertible. Then

$$\begin{aligned} A \leq \beta B &\Leftrightarrow B^{-1/2}AB^{-1/2} \leq \beta I \\ &\Leftrightarrow \sigma(B^{-1/2}AB^{-1/2}) \subseteq [0, \beta] \\ &\Leftrightarrow \sigma(B^{-1}A) \subseteq [0, \beta] \end{aligned}$$

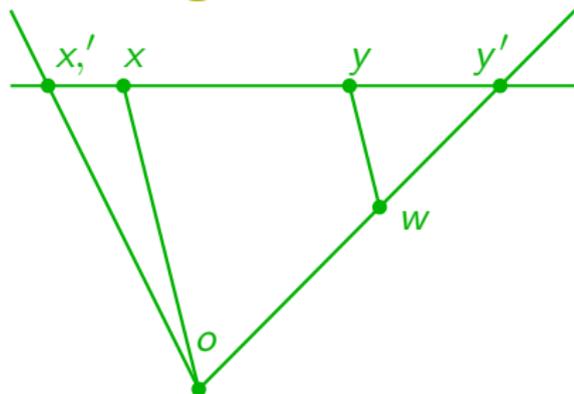
So,

$$M(A/B) = \sup\{\lambda: \lambda \in \sigma(B^{-1}A)\}$$

and

$$d_H(A, B) = \log \sup\{\lambda \in \sigma(B^{-1}A)\} - \log \inf\{\lambda \in \sigma(B^{-1}A)\}$$

## Relation with Hilbert geometries



**Theorem** If  $C \subseteq V$  is a closed cone with non-empty interior and  $H$  is an affine hyperplane such that  $C^\circ \cap H$  is nonempty and bounded, then  $\delta = d_H$  on  $C^\circ \cap H$ .

“PROOF” Note  $x \leq M(x/y)y$ , so  $w := y - M(x/y)^{-1}x \in \partial C$ . As  $\Delta oxy'$  is similar to  $\Delta wy y'$ , we get

$$\frac{|y'x|}{|y'y|} = \frac{|ox|}{|wy|} = M(x/y).$$

# Nonexpansiveness

Suppose  $C$  and  $K$  are cones in  $V$  and  $W$ , resp. If  $h: C \rightarrow K$  is an order-preserving homogenous (degree 1) map, then for  $x, y \in C$  with

$$\alpha y \leq x \leq \beta y,$$

we have that

$$\alpha h(y) \leq h(x) \leq \beta h(y).$$

Thus,

$$M(h(x)/h(y)) \leq M(x/y) \quad \text{and} \quad m(x/y) \leq m(h(x)/h(y))$$

and hence  $h$  is nonexpansive under  $d_H$ , i.e.,

$$d_H(h(x), h(y)) \leq d_H(x, y) \quad \text{for all } x \sim_C y$$

# Birkhoff's contraction ratio

Given a linear map  $L: V \rightarrow W$  with  $L(C) \subseteq K$  the **Birkhoff contraction ratio** is defined by

$$\kappa(L) = \inf\{c \geq 0: d_H(Lx, Ly) \leq cd_H(x, y) \text{ for all } x \sim_C y\}.$$

When is  $\kappa(L) < 1$  and how can we compute it?

The **projective diameter** of  $L$  is given by

$$\Delta(L) = \sup\{d_H(Lx, Ly): x, y \in C \text{ with } Lx \sim_K Ly\}.$$

# Birkhoff's theorem

**Theorem** Let  $C$  be a cone in a vector space  $V$  and  $K$  be a cone in a vector space  $W$ . If  $L: V \rightarrow W$  is a linear map with  $L(C) \subseteq K$ , then

$$\kappa(L) = \tanh\left(\frac{1}{4}\Delta(L)\right),$$

where  $\tanh(\infty) = 1$ .

So, if  $\Delta(L) < \infty$ , then  $L$  is a Lipschitz contraction on each part of  $C$ , with contraction constant  $\tanh(\Delta(L)/4) < 1$ .

Brauer, Bushell, Eveson, Hopf, Krasnoselskii, Nussbaum, Ostrowski, Thompson, . . . .

# Banach fixed point theorem

If  $(V, C, u)$  is a complete order-unit space and  $M = \{x \in C^\circ : \|x\|_u = 1\}$ , then the topology of the metric space  $(M, d_H)$  coincides with the norm topology.

So, if  $L: V \rightarrow V$  is a linear map with  $L(C^\circ) \subseteq C^\circ$  and  $\Delta(L) < \infty$ , then  $L$  has a unique eigenvector in  $C^\circ$ .

# Positive matrices

If  $L$  is given by a positive  $m \times n$  matrices  $A = (a_{ij})$ , so  $a_{ij} > 0$  for all  $i$  and  $j$ , there exists the following explicit formula:

$$\Delta(A) = \max_{i,j} d_H(Ae_i, Ae_j) = \log \left( \max_{i,j,p,q} \frac{a_{pi}a_{qj}}{a_{pj}a_{qi}} \right) < \infty,$$

where  $e_1, \dots, e_n$  denote the standard basis vectors in  $\mathbb{R}^n$ .