Perron-Frobenius Theorem

Mladen Victor WICKERHAUSER

Washington University in St. Louis, Missouri
victor@wustl.edu
http://www.math.wustl.edu/~victor

Dimensionality Reduction and Manifold Estimation
PMF — University of Zagreb
Winter, 2022

Positive and Nonnegative Square Matrices

These arise in graph theory, probability, and other contexts.

- Nonnegative $M = M(i,j) \geq 0$, for $i,j = 1,\ldots,n$.
- Positive if $M(i,j) > 0$, all i,j.
- Irreducible if M is nonnegative and $\exp(M) - I$ is positive.

Lemma

M is irreducible if and only if $(\forall i,j)(\exists k) M^k(i,j) > 0$.

Proof.

Exercise. □

Local Similarity

Given points $\mathbf{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subset \mathbf{R}^d$ (or, more generally, in some metric space).

Define a nonnegative, symmetric similarity function s on a subset of $\mathbf{V} \times \mathbf{V}$ of sufficiently similar pairs:

$$s(i,j) = s(j,i) = \begin{cases} s(\mathbf{v}_i, \mathbf{v}_j), & \|\mathbf{v}_i - \mathbf{v}_j\| < \epsilon, \\ 0, & \text{otherwise.} \end{cases}$$

Here $\epsilon > 0$ is a threshold (in the original metric) that defines "sufficiently similar."

Remark. Specifying k nearest neighbors by metric is an alternative criterion for sufficiently similar.

Global Similarity

Goal: Extend the similarity function to all of $\mathbf{V} \times \mathbf{V}$.

Method 1: Combine similarity over all paths of nonzero similarity.

- like the initial step in multidimensional scaling
- like finding shortest paths in weighted graphs
- but searching over many paths has high complexity

Method 2: Construct a diffusion process

- similarity is like an infinitestimal generator
- seek existence of long-time equilibrium solutions
- computation: find stationary distributions for Markov chains

Diffusion Maps

Choose Method 2 for generality and speed.

Extend the similarity function to all of $\mathbf{V} \times \mathbf{V}$ by

- exponentiating an infinitesimal generator, as in diffusion
- iterating a transition matrix, as for a Markov chain

In the discrete case, these are both applications for the Perron-Frobenius theorem.

Graphs

Let G be a graph with vertices $\mathcal{V} = \{1, \dots, n\}$, edges $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$.

- Adjacency matrix:

$$A(i,j) = \begin{cases} 1, & (i,j) \in \mathcal{E}, \\ 0, & \text{otherwise.} \end{cases}$$

  Generalization: weighted graphs $A(i,j) = w_{ij} \geq 0$ if $(i,j) \in \mathcal{E}$.

- Degree matrix:

$$D(i,j) = \begin{cases} \#\{k : (i,k) \in \mathcal{E}\}, & i = j, \\ 0, & \text{otherwise.} \end{cases}$$

  This is a diagonal $n \times n$ matrix.
  For a weighted graph, use $D(i,i) = \sum_{j=1}^{n} w_{ij}$.

Transition Matrices

Suppose a graph has adjacency matrix A and degree matrix D.

Transition matrix:

$$T = D^{-1}A$$

Lemma

Row sums of T are always 1.

Proof.

Fix i, compute

$$\sum_{j=1}^{n} T(i,j) = D(i,i)^{-1} \sum_{j=1}^{n} A(i,j) = \frac{\sum_{j=1}^{n} w_{ij}}{\sum_{j=1}^{n} w_{ij}} = 1.$$

□

Stochastic Matrices

Row stochastic M: Nonnegative with unit row sums:

$$(\forall i) \sum_{j=1}^{n} M(i,j) = 1.$$

Column stochastic: Nonnegative with unit column sums:

$$(\forall j) \sum_{i=1}^{n} M(i,j) = 1.$$

Doubly stochastic: both row and column stochastic.

Probability Vectors

Define a row pdf to be a probability function written as a row vector on the finite space $\Omega = \{1, \ldots, n\}$:

$$\mathbf{p} = \begin{pmatrix} p_1 & \ldots & p_n \end{pmatrix}; \qquad (\forall j) p_j \geq 0; \quad \sum_{j=1}^{n} p_j = 1.$$

Similarly, column pdf $\mathbf{q}$ is a column vector with nonnegative entries that sum to 1.

Lemma

For row stochastic M, if $\mathbf{p}$ is a row pdf, then $\mathbf{p}M$ is a row pdf. □

Also, column stochastic M maps column pdf $\mathbf{q}$ to column pdf $M\mathbf{q}$. Both proofs are left as exercises.

Finite Stationary Markov Chains

Stochastic process on the *finite* state space $\Omega = \{1, \ldots, n\}$.

Map initial pdf $\mathbf{p}_0$ to pdfs $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_k, \ldots$ by iterated application of stochastic M.

Stationary if the same M is used at each step.

Questions:

- does $\mathbf{p}_\infty \stackrel{\text{def}}{=} \lim_{k\to\infty} \mathbf{p}_k$ exist?
- can $\mathbf{p}_\infty$ be found by iteration? How fast will it converge?
- is $\mathbf{p}_\infty$ independent of $\mathbf{p}_0$?

If a limit $\mathbf{p}_\infty$ exists, it is called a *stationary distribution* for M.

Eigenvalue Problem

Stationary distributions $\mathbf{q} = \mathbf{p}_\infty$ (for the column stochastic case) solve the eigenvalue equation

$$\mathbf{q} = M\mathbf{q}$$

with column stochastic M having eigenvalue 1.

Since $\mathbf{q}$ is a (column) pdf, the solution is unique if and only if eigenvalue 1 has multiplicity 1. (Prove this as an exercise.)

Solution $\mathbf{q}$ is a limit of iterations of M if all other eigenvalues λ of M satisfy $|\lambda| < 1$.

Convergence $\|\mathbf{p}_\infty - \mathbf{p}_k\| = O(|\lambda|^{-k})$ as $k \to \infty$, where $|\lambda| < 1$ is largest-magnitude eigenvalue with $|\lambda| < 1$.

Spectral Radius and Matrix Norms

Spectral radius for $n \times n$ matrix M with eigenvalues $\{\lambda_i\} \subset \mathbf{C}$:

$$\rho(M) \stackrel{\text{def}}{=} \max\{|\lambda_1|, \ldots, |\lambda_n|\},$$

Matrix norm for $n \times n$ matrices M, N and scalars c, satisfies:

- $\|M\| \geq 0$, with $\|M\| = 0 \iff M = 0$; $\|cM\| = |c|\,\|M\|$.
- $\|M + N\| \leq \|M\| + \|N\|$ and $\|MN\| \leq \|M\|\,\|N\|$.

Theorem

Any two norms on a finite-dimensional vector space are equivalent: $\|\cdot\|_\alpha \sim \|\cdot\|_\beta$, *meaning* $(\exists K > 0)(\forall M)\ \|M\|_\alpha \leq K\|M\|_\beta$. □

Proof.

See `mfmm30-32.pdf` on class website. Note that $K = K(\alpha, \beta, n)$ depends on the norms and on the dimension. □

Example Matrix Norms

Fredholm Norm: $\|M\|_F \stackrel{\text{def}}{=} \left(\sum_{i,j} |M(i,j)|^2\right)^{1/2}$ (this is Euclidean norm on $\mathbf{C}^{n\times n}$, the matrix coefficients)

One Norm: $\|M\|_1 \stackrel{\text{def}}{=} \max_j \sum_i |M(i,j)|$

Infinity Norm: $\|M\|_\infty \stackrel{\text{def}}{=} \max_i \sum_j |M(i,j)|$

Operator Norm: $\|M\|_{\text{op}} \stackrel{\text{def}}{=} \sup_{\mathbf{x}\neq 0} \frac{\|M\mathbf{x}\|}{\|\mathbf{x}\|} = \sup_{\|\mathbf{x}\|=1} \|M\mathbf{x}\|.$

Lemma

$\|M\|_{\text{op}} = \rho(M^*M)^{1/2}$ *is the largest singular value of M.*

Proof.

$$\|M\|_{\text{op}}^2 = \sup_{\|\mathbf{x}\|=1} \|M\mathbf{x}\|^2 = \sup_{\|\mathbf{x}\|=1} \langle M^*M\mathbf{x}, \mathbf{x}\rangle = \rho(M^*M). \qquad \square$$

Induced Operator Norms

Let $\|\cdot\|_X$ be any norm on $\mathbf{C}^n$.

For $n \times n$ matrix M, define its induced operator norm by

$$\|M\|_{\mathrm{X,op}} \stackrel{\mathrm{def}}{=} \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|M\mathbf{x}\|_X}{\|\mathbf{x}\|_X}.$$

The resulting function $\|\cdot\|_{\mathrm{X,op}}$ is a matrix norm.

Lemma

Let $\|\cdot\|$ be any matrix norm. Then $\|I\| \geq 1$.

Proof.

$I \neq 0$, so $\|I\| > 0$, and $\|I\|^2 \geq \|I^2\| = \|I\|$, so $\|I\| \geq 1$. □

Continuity of Matrix Norms

Fix n and let $\|\cdot\|$ be any matrix norm on $n \times n$ matrices.

Lemma

$M \mapsto \|M\|$ is a continuous function on the coefficients of M.

Proof.

Since $\|M\| \leq \|M - N\| + \|N\|$ and $\|N\| \leq \|N - M\| + \|M\|$, it follows that

$$\Big| \|M\| - \|N\| \Big| \leq \|M - N\|.$$

Since $\|\cdot\| \sim \|\cdot\|_F$, there is some $0 < K < \infty$ such that $\|M - N\| \leq K\|M - N\|_F$. Conclude that

$$\Big| \|M\| - \|N\| \Big| \leq \|M - N\| \leq K\|M - N\|_F$$

so that $\|\cdot\|$ is (Lipschitz) continuous with respect to Euclidean norm on $\mathbf{C}^{n\times n}$, the vector space of matrix coefficients. □

Matrix Norm and Boundedness

Fix n and let $\| \cdot \|$ be any matrix norm on $n \times n$ matrices.

Lemma

There is some constant $K > 0$ such that, for all $n \times n$ matrices M and all vectors $\mathbf{x}$, $\|M\mathbf{x}\| \leq K\|M\| \, \|\mathbf{x}\|$, where $\|\mathbf{x}\|$ is the Euclidean norm of $\mathbf{x} \in \mathbf{C}^n$.

Proof.

Define the matrix $X(i,j) = \mathbf{x}_i$ (each column is a copy of $\mathbf{x}$).
Then $\|X\|_F = \sqrt{n}\,\|\mathbf{x}\|$, and $\|MX\|_F = \sqrt{n}\,\|M\mathbf{x}\|$.
But there exists $K > 0$ such that $\|M\|_F \leq K\|M\|$, so

$$\|M\mathbf{x}\| = \frac{1}{\sqrt{n}}\|MX\|_F \leq \frac{1}{\sqrt{n}}\|M\|_F\|X\|_F \leq K\|M\| \, \|\mathbf{x}\|,$$

by the equivalence of matrix norms $\| \cdot \| \sim \| \cdot \|_F$. □

Norm versus Spectral Radius

Suppose that $\|\cdot\|$ is any matrix norm.

Lemma

If $\rho(M) > 1$, *then* $\lim_{k\to\infty} \|M^k\| = \infty$.

Proof.

Since $\rho(M) > 1$, M has an eigenvalue λ with $|\lambda| > 1$. Let $\mathbf{v} \neq \mathbf{0}$ be an eigenvector for λ. Then as $k \to \infty$,

$$\|M^k\|_{\mathrm{op}} = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|M^k\mathbf{x}\|}{\|\mathbf{x}\|} \geq \frac{\|M^k\mathbf{v}\|}{\|\mathbf{v}\|} = |\lambda|^k \to \infty.$$

But $\|\cdot\|_{\mathrm{op}} \sim \|\cdot\|$, so $\|M^k\| \geq \frac{1}{K}\|M^k\|_{\mathrm{op}}$ for some $0 < K < \infty$, so $\|M^k\| \to \infty$ as $k \to \infty$. □

Special Case: Nilpotent Matrices

If M is nilpotent, namely $M^k = 0$ for some k, then $\rho(M) = 0$, because any eigenvalue λ with eigenvector $\mathbf{v} \neq \mathbf{0}$ satisfies

$$\mathbf{0} = 0\mathbf{v} = M^k\mathbf{v} = \lambda^k\mathbf{v}, \implies \lambda^k = 0, \implies \lambda = 0.$$

Conversely, if $\rho(M) = 0$, then M is nilpotent. This follows from the Cayley-Hamilton theorem below.

If M is diagonalizable, then $\rho(M)$ is its largest singular value, but this is false for more general M. Example: nonzero nilpotent

$$N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad N^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad N^T N = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

with eigenvalues $0, 0$ so $\rho(N) = 0$, but with singular values $0, 1$.

Jordan Canonical Form

Theorem

A square matrix M with eigenvalues $\{\lambda_i\}$ has a Jordan canonical form: $M = SJS^{-1}$ with invertible S and block diagonal

$$J = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_m \end{pmatrix}, \quad \text{for } J_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix} = \lambda_i I + N_i$$

- Block J_i corresponds to eigenvalue λ_i.
- The order n_i of J_i (and of N_i) is at most the multiplicity of λ_i.
- N_i is nilpotent, with $N_i^k = 0$ for all $k \geq n_i$. □

Corollary: $M^k = (SJS^{-1})^k = SJ^kS^{-1}$. □

Cayley-Hamilton Theorem

Theorem

If χ is the characteristic polynomial of matrix M, then $\chi(M) = 0$.

Proof.

Let $M = SJS^{-1}$ be the Jordan canonical form of M. Then

$$\chi(M) = S\chi(J)S^{-1} = S \begin{pmatrix} \chi(J_1) & & 0 \\ & \ddots & \\ 0 & & \chi(J_m) \end{pmatrix},$$

where $J_i = \lambda_i I + N_i$ is a Jordan block. Let n_i be its order, so nilpotent $N_i^{n_i} = 0$. Now write $\chi(z) = \prod_j (z - \lambda_j)^{n_j}$ to see

$$\chi(J_i) = (\lambda_i I + N_i - \lambda_i I)^{n_i} \prod_{j \neq i} (J_i - \lambda_j I)^{n_j} = N_i^{n_i} \prod_{i \neq j} (J_i - \lambda_j I)^{n_j} = 0.$$

Conclude that $\chi(J) = 0$, so therefore $\chi(M) = 0$. □

Powers of Jordan Blocks

Lemma

Let $J = \lambda I + N$ be an $m \times m$ Jordan block for eigenvalue λ. Then $\lim_{k\to\infty} J^k = 0$ if and only if $|\lambda| < 1$.

Proof.

Obviously true for $m = 1$, so suppose $m > 1$ with nilpotent $N \neq 0$.

Since $N^m = 0$, expand $J^k = (\lambda I + N)^k$, for $k \geq m - 1$, as

$$J^k = \lambda^k I + \binom{k}{1}\lambda^{k-1}N + \cdots + \binom{k}{m-1}\lambda^{k+1-m}N^{m-1}.$$

If $|\lambda| < 1$, then $J^k = O(k^{m-1}|\lambda|^{k+1-m}) \to 0$ as $k \to \infty$.

If $|\lambda| \geq 1$, then $J^k N^{m-1} = \lambda^k N^{m-1}$ does not converge to 0 as $k \to \infty$, and since N^{m-1} is constant, neither does J^k . □

Powers of Square Matrices

Corollary

Let M be a square matrix with spectral radius $\rho(M)$. Then $\lim_{k\to\infty} M^k = 0$ *if and only if* $\rho(M) < 1$.

Proof.

Let $M = SJS^{-1}$ be the Jordan canonical decomposition of M. Then $M^k = SJ^kS^{-1}$ for all $k = 1, 2, \ldots,$ and since S is nonsingular, $\lim_{k\to\infty} M^k = 0$ if and only if $\lim_{k\to\infty} J^k = 0$.

If $\rho(M) < 1$, then $\lim_{k\to\infty} J^k = 0$, so $\lim_{k\to\infty} M^k = 0$.

But if $\rho(M) \geq 1$, then there exists some eigenvalue λ of M with $|\lambda| \geq 1$, so $\lim_{k\to\infty} J^k \neq 0$, so $\lim_{k\to\infty} M^k \neq 0$. □

Note: Every matrix norm is a continuous function of the matrix coefficients, so $\lim_{k\to\infty} \|M^k\| = 0$ if and only if $\rho(M) < 1$.

Zero Spectral Radius Implies Nilpotent

Corollary

Let M be an $n \times n$ matrix with spectral radius $\rho(M) = 0$. Then there exists $1 \leq k \leq n$ such that $M^k = 0$.

Proof.

Let $M = SJS^{-1}$ be the Jordan canonical decomposition of M. Since $\rho(M) = 0$, all eigenvalues of M must be zero, so every Jordan block $J_i = N_i$ is nilpotent with order $n_i \leq n$ equal to the order of block J_i.
Let $k = \max_i n_i$. Then $1 \leq k \leq n$, and $(\forall i)\, J_i^k = 0$.
Thus $J^k = 0$, so $M^k = SJ^kS^{-1} = 0$.

Alternate proof: Every eigenvalue is zero, so $\chi(z) = z^n$, so by the Cayley-Hamilton theorem, $\chi(M) = M^n = 0$. □

Gel'fand's Formula

Lemma

For any $n \times n$ matrix M and norm $\|\cdot\|$, $\rho(M) = \lim_{k\to\infty} \|M^k\|^{1/k}$.

Proof.

If $\rho(M) = 0$, then $M^n = 0$ by the Cayley-Hamilton Theorem. Hence $M^k = 0$ for all $k \geq n$, so $\lim_{k\to\infty} \|M^k\|^{1/k} = 0 = \rho(M)$.

Else $\rho(M) > 0$, so let $0 < \epsilon < \rho(M)$ be given and put

$$M_- \stackrel{\text{def}}{=} \frac{1}{\rho(M) - \epsilon} M, \qquad M_+ \stackrel{\text{def}}{=} \frac{1}{\rho(M) + \epsilon} M.$$

Then $0 < \rho(M_+) < 1 < \rho(M_-)$, so $\|M_+^k\| \to 0$ while $\|M_-^k\| \to \infty$ as $k \to \infty$. Hence for all sufficiently large k,

$$\frac{\|M^k\|}{(\rho(M)+\epsilon)^k} = \|M_+^k\| < 1 < \|M_-^k\| = \frac{\|M^k\|}{(\rho(M)-\epsilon)^k},$$

so $\rho(M) - \epsilon < \|M^k\|^{1/k} < \rho(M) + \epsilon$. □

Fixed Point Existence

Theorem (Brouwer)

If $f : X \to X$ is a continuous endomorphism on compact convex $X \subset \mathbf{C}^n$, then f has a fixed point: $(\exists x \in X)\, f(x) = x$. $\square$

Application: for invertible $n \times n$ matrix M with $\|M\|_{\mathrm{op}} \leq 1$, the map

$$\mathbf{x} \mapsto M\mathbf{x}$$

is defined and continuous from the closed unit ball in $\mathbf{C}^n$ into itself, and thus has a fixed point.

Problem: avoid the trivial fixed point $M\mathbf{0} = \mathbf{0}$.

Power Method

Lemma

If M has a maximal eigenvalue $\lambda = r$, with $|\lambda| < r$ for all its other eigenvalues, then the iteration $\mathbf{x}_{k+1} = \frac{1}{r} M\mathbf{x}_k$ starting from almost any $\mathbf{x}_0$ (that is, any $\mathbf{x}_0$ with a nonzero projection into the r-eigenspace) will converge to an r-eigenvector.

Proof.

Write $\mathbf{x}_0 = \mathbf{v} \oplus \mathbf{u}$ with $\mathbf{v} \neq \mathbf{0}$ in r-eigenspace X_r and $\mathbf{u} \in X_r^{\perp}$.
Then $(\frac{1}{r}M)^k\mathbf{v} = \mathbf{v}$ while $(\frac{1}{r}M)^k\mathbf{u} \to \mathbf{0}$ as $k \to \infty$. □

Remark. The same holds for iteration with renormalization:

$$\mathbf{x}_{k+1} = \frac{1}{\|M\mathbf{x}_k\|} M\mathbf{x}_k, \qquad k = 0, 1, 2, \ldots$$

For almost every $\mathbf{x}_0$, $\lim_{k\to\infty} \mathbf{x}_k$ is a unit r-eigenvector.

Perron-Frobenius I

Theorem

For any positive $n \times n$ matrix M with spectral radius $r = \rho(M)$:

1. $0 < \min_i \sum_j M(i,j) \le r \le \max_i \sum_j M(i,j)$,
2. *r is an eigenvalue for M,*
3. *every other eigenvalue λ of M satisfies $|\lambda| < r$,*
4. *there exists a positive r-eigenvector $\mathbf{v}$ of M, namely* $\mathbf{v} = (v_1, \ldots, v_n)$ *with* $(\forall i) v_i > 0$,
5. *eigenvalue r has multiplicity 1, and*
6. *every other eigenvector with all positive coordinates is a positive scalar multiple of $\mathbf{v}$.*

PFI.1: Lower Bound for Spectral Radius

If M is positive, then M^k is positive for all $k > 0$, so

$$\mu_k \stackrel{\text{def}}{=} \min_i \sum_j M^k(i,j) \; > \; 0, \qquad k = 1, 2, \ldots$$

Let $\mathbf{1} = (1, \ldots, 1)$ and compute

$$\|M^k\|_{\text{op}} \geq \frac{\|M^k \mathbf{1}\|}{\|\mathbf{1}\|} = \frac{1}{\sqrt{n}} \sqrt{\sum_i \left[\sum_j M^k(i,j)\right]^2} \geq \mu_k$$

But $\mu_{k+1} \geq \mu_1 \mu_k$ (Exercise!), so $\mu_k \geq \mu_1^k$. Now apply Gel'fand:

$$\rho(M) = \lim_{k\to\infty} \|M^k\|_{\text{op}}^{1/k} \geq \lim_{k\to\infty} (\mu_k)^{1/k} \geq \lim_{k\to\infty} (\mu_1^k)^{1/k} \geq \mu_1,$$

which means that $\boxed{\rho(M) \geq \min_i \sum_j M(i,j) > 0}$.

PFI.1: Upper Bound for Spectral Radius

For positive M, put

$$\gamma_k \stackrel{\text{def}}{=} \max_i \sum_j M^k(i,j) = \|M^k\|_\infty.$$

By the submultiplicativity of the matrix norm $\|\cdot\|_\infty$,

$$\gamma_{k+1} = \|M^{k+1}\|_\infty \le \|M\|_\infty \|M^k\|_\infty = \gamma_1 \gamma_k,$$

so $\gamma_k \le \gamma_1^k$ for all k. Apply the Gel'fand formula with this norm,

$$\rho(M) = \lim_{k\to\infty} \|M^k\|_\infty^{1/k} \le \lim_{k\to\infty} (\gamma_k)^{1/k} \le \lim_{k\to\infty} (\gamma_1^k)^{1/k} = \gamma_1.$$

Conclude that $\boxed{\rho(M) \le \max_i \sum_j M(i,j)}$.

Gershgorin's Theorem

The bounds on $\rho(M)$ are a special case of:

Theorem

Suppose M is an $n \times n$ matrix over $\mathbf{C}$. For $i = 1, \ldots, n$, define the Gershgorin disc $G_i \subset \mathbf{C}$ by

$$G_i \stackrel{\text{def}}{=} \Big\{ z \in \mathbf{C} : |z - M(i,i)| \leq \sum_{j \neq i} |M(i,j)| \Big\}.$$

Then every eigenvalue of M lies in $\bigcup_i G_i$.

Proof.

This relatively simple proof is left as an exercise. □

Thus, the largest eigenvalue $z = \rho(M)$ of positive M must satisfy $z \leq M(i,i) + \sum_{j \neq i} M(i,j)$ for some i, so $z \leq \max_i \sum_j M(i,j)$.

Proof of PFI.2 and PFI.3

Assume that $\rho(M) = 1$, else use $M/\rho(M)$. Thus for eigenvalues λ:

$$\boxed{(\forall\lambda)\,|\lambda| \leq 1} \qquad (\exists\lambda)\,|\lambda| = 1.$$

Suppose $|\lambda| = 1$ but $\lambda \neq 1$. Then $(\exists m \in \mathbf{Z}^+)\mathrm{Re}\,\lambda^m < 0$.

Let $\epsilon = \frac{1}{2}\min_j M^m(j,j) > 0$. Then $T \stackrel{\text{def}}{=} M^m - \epsilon I$ is a positive matrix, with an eigenvalue $\lambda^m - \epsilon$, so $\rho(T) \geq |\lambda^m - \epsilon| > 1$. Now

$$(\forall i,j)\,0 < T(i,j) \leq M^m(i,j) \implies (\forall i,j,k)\,0 < T^k(i,j) \leq M^{mk}(i,j).$$

Thus $(\forall k)\,\|T^k\|_F \leq \|M^{mk}\|_F$. Apply Gel'fand with $\|\cdot\|_F$ to get

$$\rho(T) \leq \rho(M^m) = \rho(M)^m = 1.$$

Contradiction! so $\boxed{\lambda = 1}$ is the unique eigenvalue with $|\lambda| = 1$.

PFI.4: Positive Eigenvector

Some positive $\mathbf{x}_0$ near $\mathbf{1}$ will have a nonzero component in the r-eigenspace. The power method converges from that $\mathbf{x}_0$:

$$\mathbf{x}_{k+1} = \frac{1}{\|M\mathbf{x}_k\|} M\mathbf{x}_k, \qquad k = 0, 1, 2, \ldots$$

For all k, $\mathbf{x}_k$ has all positive coordinates, so the r-eigenvector $\mathbf{v} = \lim_{k\to\infty} \mathbf{x}_k$ has nonnegative coordinates $v_1, \ldots, v_n$.

But if $v_i = 0$ for some i, then $M\mathbf{v} = r\mathbf{v}$ implies

$$0 = v_i = \frac{1}{r} \sum_{j=1}^{n} M(i,j) v_j, \implies (\forall j) v_j = 0,$$

since $(\forall j) M(i,j) > 0$. This is a contradiction since $\|\mathbf{v}\| = r > 0$ by construction. Conclude that $\mathbf{v}$ is a positive eigenvector.

PFI.5: Multiplicity One

Let **v** be a positive *r*-eigenvector.

Suppose that **u** is another *r*-eigenvector. Without loss, some component of **u** is positive, else use $-\mathbf{u}$.

For $\alpha > 0$, let $\mathbf{w} \stackrel{\text{def}}{=} \mathbf{v} - \alpha\mathbf{u}$. Any $\mathbf{w} \neq \mathbf{0}$ is an *r*-eigenvector.

There is a maximal positive α for which **w** is nonnegative. By maximality, some component of **w** must be 0.

However, any nonnegative *r*-eigenvector must in fact be positive by PFI.4. Hence $\mathbf{w} = \mathbf{0}$, so $\mathbf{u} = \frac{1}{\alpha}\mathbf{v}$.

Conclude that there cannot be another linearly independent *r*-eigenvector.

PFI.6: Positive Eigenvectors

Let $\mathbf{v} = (v_1, \ldots, v_n)$ be the r-eigenvector with all positive coordinates from PFI.5, so $(\forall i) v_i > 0$.

Let $\mathbf{x} = (x_1, \ldots, x_n)$ be another positive eigenvector, so $(\forall i) x_i > 0$. Then $\mathbf{x}$ is an r-eigenvector, since $\langle \mathbf{x}, \mathbf{v} \rangle > 0$ implies that $\mathbf{x}$ cannot be in the (orthogonal) eigenspace of any other eigenvalue of M.

Since the r-eigenspace is one-dimensional, $\mathbf{x} = \mathbf{c}\mathbf{v}$. Thus $(\forall i)\, x_i = c v_i$. This is possible if and only if $c > 0$.

Conclude that $\mathbf{x} = c\mathbf{v}$ is another positive r-eigenvector if and only if $c > 0$.

This completes the proof of Perron-frobenius I. □

Perron-Frobenius II

Theorem

If M is a nonnegative irreducible $n \times n$ matrix with $\rho(M) = r > 0$, then all results for PFI hold with these changes:

PFII.4: *there exists an eigenvector $\mathbf{v} = (v_1, \ldots, v_n)$ of M, with eigenvalue r, such that $\boxed{(\forall i) v_i \geq 0}$,*

PFII.6: *every other eigenvector with* $\boxed{\text{nonnegative}}$ *coordinates is a positive scalar multiple of $\mathbf{v}$,*

Proof.

Idea: since $N = \exp(M) - I$ is positive, apply PFI to N. But $M\mathbf{v} = \lambda\mathbf{v}$ implies $N\mathbf{v} = [\exp(\lambda) - 1]\mathbf{v}$. □

Markov Matrices

Row stochastic nonnegative M:

$$(\forall i,j)\, M(i,j) \geq 0; \qquad (\forall i) \sum_j M(i,j) = 1.$$

Say that such an M is *ergodic* if

- M is irreducible: $\exp(M) - I$ is positive, and
- M is aperiodic: $(\forall i)\, \mathrm{period}(i) = 1$, where

$$\mathrm{period}(i) \stackrel{\mathrm{def}}{=} \gcd\{k \geq 1 : M^k(i,i) \neq 0\}.$$

Lemma

If M is ergodic, then $\lim_{k\to\infty} M^k$ exists and has constant rows $\mathbf{v}$ satisfying $\mathbf{v}M = \mathbf{v}$.

Adjacency Matrices

Lemma

The adjacency matrix for a connected graph is irreducible.

Proof.

Form the transition matrix $T = D^{-1}A$, where A is the adjacency matrix and D is the degree matrix. This is row stochastic.
Since the graph is connected, every pair of vertices i, j are connected by a path whose probability is $T^k(i,j) > 0$, where k is the path length. Therefore,

$$(\forall i,j)(\exists k) T^k(i,j) > 0, \implies (\forall i,j)(\exists k) A^k(i,j) > 0.$$

This implies that $\exp(A) - I = \sum_{k\geq 1} \frac{1}{k!} A^k > 0$. $\square$

Diffusion Matrices

Normalize a similarity matrix to be row stochastic.