Nonnegative and spectral matrix theory
Lecture notes

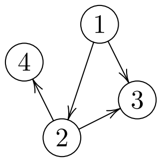
Dario Fasino, University of Udine (Italy)

Lecture notes for the first part of the course "Nonnegative and spectral matrix theory with applications to network analysis", held within the Rome-Moscow school on Matrix Methods and Applied Linear Algebra, August 2014. The author is partially supported by Istituto Nazionale di Alta Matematica. This document is intended for internal circulation only.

1 Basic definitions

A *(directed) graph* is a pair $G = (V, E)$ where V is a finite set of *nodes* (or *vertices*) and $V \subseteq E \times E$ is a set of *(oriented) edges*. Nodes can be visualized as points in the plane, and edges as arrows joining nodes. Hereafter, I generally assume $V = \{1, \dots, n\}$ and write $i \to j$ to indicate that $(i, j) \in E$. Edges of the form (i, i) are called *loops*.

The graph $G = (V, E)$ can be completely described by its *adjacency matrix*, which is the $n \times n$ matrix A such that $A_{ij} = 1$ if $j \to i$ and $A_{ij} = 0$ otherwise.[1] The notation $A = A_G$ indicates that A is the adjacency matrix of G.

$G:$ 

$$\rightsquigarrow \qquad A_G = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

A graph is *undirected* (or *non-oriented*) when $i \to j \Leftrightarrow j \to i$, that is, when its adjacency matrix is symmetric. In that case, edges are depicted as lines instead of arrows, and the notation $i \sim j$ replaces both $i \to j$ and $j \to i$.

Let $A = A_G$, and let $v \in \mathbb{R}^n$. If we consider v_i as a score placed on node i, then it is useful to look at the matrix-vector product $w = Av$ as propagating the scores along the edges of G. A more formal result is the following, whose simple proof proceeds by induction and is omitted for brevity:

Lemma 1.1. *Let $A = A_G$. For any $k \in \mathbb{N}$ and $i, j = 1, \dots, n$ the value of $(A^k)_{ij}$ is equal to the number of different walks of length k starting from j and ending in i.*

Hereafter, a *walk of length* $k \geq 1$ in G is any sequence of nodes $i_0, i_1, \dots, i_k$ such that $i_{j-1} \to i_j$ (or $i_{j-1} \sim i_j$ in the undirected case) for $j = 1, \dots, k$

For any given matrix $A \in \mathbb{R}^{n\times n}$, the *graph associated to* A is the graph $G_A = (V, E)$ such that $V = \{1, \dots, n\}$ and $j \to i \Longleftrightarrow A_{ij} \neq 0$. Thus, if the entries of A belong to the set $\{0, 1\}$ then A is the adjacency matrix of G_A. Other useful notations are the following:

- An all-zeros matrix is denoted by O. An all-ones vector is denoted by $\mathbf{1}$.
- Inequality operators like $\geq$ or $>$ are extended to matrices and vectors in the componentwise sense; for example, $A \geq O$ means that all elements of A are nonnegative.

1.1 Irreducible matrices

Definition 1.2. *The matrix $A \in \mathbb{R}^{n\times n}$ is* reducible *if there is a permutation matrix P such that the matrix $B = PAP^T$ is in (lower) block triangular form:*

$$B = PAP^T = \begin{pmatrix} B_{11} & B_{12} \\ O & B_{22} \end{pmatrix},$$

[1] Various authors define the adjacency matrix as $A_{ij} = 1 \Leftrightarrow i \to j$. I prefer the other definition for simplicity of subsequent notations.

where diagonal blocks B_{11}, B_{22} are square matrices. An irreducible *matrix is a matrix that is not reducible.*

Definition 1.3. *A graph is* strongly connected *if any two nodes are connected by a walk.*

The two preceding definitions are connected by the following important result:

Theorem 1.4. *A matrix $A \in \mathbb{R}^{n\times n}$ is irreducible if and only if G_A is strongly connected.*

PROOF. Suppose that A is reducible. Apart of a permutation (which corresponds to a renumbering of the nodes of G_A) we can assume that A is already in reduced block triangular form:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix}, \qquad A_{11} \in \mathbb{R}^{n_1\times n_1}, \qquad A_{22} \in \mathbb{R}^{n_2\times n_2}, \tag{1}$$

with $n_1+n_2 = n$. Hence, in G_A there are no edges connecting nodes $1, \dots, n_1$ to nodes $n_1+1, \dots, n$. As a consequence, there are no walks going from nodes $n_1 + 1, \dots, n$ to nodes $1, \dots, n_1$, and the graph is not strongly connected.

Conversely, if G_A is not strongly connected then there are two distinct nodes, say i and j, such that there is no walk from i to j. Let $\mathcal{J}$ be the set of all nodes that cannot be reached by a walk starting from i, and let $\mathcal{I}$ be its complementary set. Without loss of generality, we can suppose that $\mathcal{J} = \{1, \dots, n_1\}$ (note that it's not empty, since $j \in \mathcal{J}$) and $\mathcal{I} = \{n_1 + 1, \dots, n\}$. Clearly, A has the form (1), hence it is reducible. ■

2 A glimpse to Perron–Frobenius theory

Quoting from [2, p. 662]: *"The Perron–Frobenius theory is elegant. It is a testament to the fact that beautiful mathematics eventually tends to be useful, and useful mathematics eventually tends to be beautiful."*

Theorem 2.1 (Perron–Frobenius)**.** *Let $A \in \mathbb{R}^{n\times n}$ be an irreducible, nonnegative matrix. Then,*

1. *A has a positive eigenvalue equal to $\rho(A)$.*
2. *To $\rho(A)$ corresponds a positive eigenvector x.*
3. *$\rho(A)$ is a simple eigenvalue of A, that is, corresponds to a single Jordan block of order 1.*
4. *$\rho(A)$ increases (or decreases) when any entry of A increases (or decreases, respectively). That is, if A and B are two nonnegative, irreducible matrices with $O \le A \le B$ and $A \ne B$ then $\rho(A) < \rho(B)$.*

It is usual to call $\rho(A)$ the *Perron eigenvalue* of A. Any associated positive eigenvector is a *Perron eigenvector*. Before proceeding to the proof of Theorem 2.1, it is useful to collect hereafter some facts from matrix theory.

Lemma 2.2. *If $A \in \mathbb{R}^{n\times n}$ is an irreducible, nonnegative matrix then $(I + A)^{n-1} > O$.*

Hint: $(I + A)^{n-1}$ is a linear combination of $I, A, \dots, A^{n-1}$ with positive coefficients. Use Theorem 1.4 and the "score propagation" argument to show that any column of $(I + A)^{n-1}$ is positive.

Lemma 2.3. *For any square matrix M, if $\rho(M) < 1$ then the matrix series $\sum_{k=0}^{\infty} M^k$ converges, and $\sum_{k=0}^{\infty} M^k = (I - M)^{-1}$.*

Lemma 2.4. *Let $Ax = \lambda x$. Then, λ is multiple if and only if there exists y such that $A^T y = \lambda y$ and $x^T y = 0$.*

Hint: The proof of the last two facts relies on the Jordan normal form of A. Actually, it is sufficient to consider just two simple cases: (1) A is diagonalizable; (2) A consists of a single nontrivial Jordan block. In fact, complete proofs can be established on the basis of the discussion of these two cases.

2.1 Proof of Theorem 2.1

The following proof is largely based on the unpublished technical report by G. W. Stewart [3]. Various details have been added or adjusted after discussing with students of the Rome-Moscow school.

First of all, observe that if $A \geq O$ is irreducible then $\rho(A) > 0$. In fact, if $\rho(A) = 0$ then A is *nilpotent*, that is, there exists a positive integer m such that $A^m = O$. In particular, $A^m e_1 = 0$. According to the "score propagation" interpretation of matrix-vector products $Ae_1, A^2e_1, \ldots$, we must conclude that all walks in G_A starting from node 1 sooner or later arrive to nodes without outgoing edges. This fact contradicts the fact that, owing to Theorem 1.4, G_A is strongly connected.

By replacing A with $A/\rho(A)$, we may assume that $\rho(A) = 1$. For any $0 < \tau < 1$, the matrix $I - \tau A$ is nonsingular (hint: its eigenvalues lie in the circle $\{z \in \mathbb{C} : |1 - z| \leq \tau\}$). Hence, let

$$x_\tau = (I - \tau A)^{-1}\mathbf{1} = \mathbf{1} + \tau A\mathbf{1} + \tau^2 A^2 \mathbf{1} + \ldots$$

The series converges (owing to Lemma 2.3) and all terms are ≥ 0. Hence $x_\tau \geq 0$. Its 1-norm is

$$\|x_\tau\|_1 = \mathbf{1}^T x_\tau = \sum_{k=0}^{\infty} \tau^k \mathbf{1}^T A^k \mathbf{1}.$$

We want to prove that $\|x_\tau\|_1 \to \infty$ as $\tau \to 1$. In fact, note that

$$\mathbf{1}^T A^k \mathbf{1} \geq \max_i (A^k \mathbf{1})_i = \|A^k\|_\infty \geq \rho(A^k) = \rho(A)^k = 1,$$

hence $\|x_\tau\|_1 \geq \sum_{k=0}^{\infty} \tau^k = 1/(1-\tau)$.

The normalized vectors $x_\tau/\|x_\tau\|_1$ lie on a compact (= closed and bounded) set. Hence, we can choose a sequence $\tau_i \to 1$ such that $x_{\tau_i}/\|x_{\tau_i}\|_1 \to x \geq 0$. Owing to continuity,

$$0 = \lim_{i\to\infty} \frac{1}{\|x_{\tau_i}\|_1}\mathbf{1} = \lim_{i\to\infty} \frac{1}{\|x_{\tau_i}\|_1}(I - \tau_i A)x_{\tau_i} = (I - A)x.$$

Thus $Ax = x$, that is, x is an eigenvector of A corresponding to $1 = \rho(A)$. To show that $x > 0$ it is sufficient to observe that, since $(I + A)x = 2x$ and $(I + A)^{n-1} > O$ (Lemma 2.2), we have $0 < (I + A)^{n-1}x = 2^{n-1}x$.

If $\rho(A)$ is not a simple eigenvalue then by Lemma 2.4 there exists a vector y such that $y^T A = \rho(A)y^T$ and $y^T x = 0$. Since x is positive, y^T must have both positive and negative entries (otherwise $y^T x > 0$). Let $\mathcal{P} = \{i : y_i \geq 0\}$ and $\mathcal{N} = \{i : y_i < 0\}$. Both sets are not empty and, without loss in generality, we can assume that $\mathcal{P} = \{1, \ldots, n_1\}$ and $\mathcal{N} = \{n_1 + 1, \ldots, n\}$.

Let z be the vector defined as

$$z_j = \begin{cases} y_j & j \in \mathcal{P} \\ 0 & j \in \mathcal{N}. \end{cases}$$

It is not difficult to see that $\rho(A)z^T \leq z^T A$. Moreover, because of irreducibility, there exists $i \in \mathcal{N}$ and $j \in \mathcal{P}$ such that $A_{ij} > 0$; otherwise, A is as in (1). Consequently,

$$\begin{aligned} \rho(A)z_j = \rho(A)y_j &= \sum_{i\in\mathcal{P}} A_{ij}y_i + \sum_{i\in\mathcal{N}} A_{ij}y_i \\ &< \sum_{i\in\mathcal{P}} A_{ij}y_i = \sum_{i=1}^{n} A_{ij}z_i = (z^T A)_j. \end{aligned}$$

Thus, the inequality $\rho(A)z^T \leq z^T A$ is strict in at least one entry in $\mathcal{P}$. Oving to the positivity of x we deduce

$$\rho(A)z^T x = z^T Ax < \rho(A)z^T x,$$

a contradiction.

Before completing the proof of Theorem 2.1, let's prove a result of independent interest:

Lemma 2.5. *Let $A \geq O$ be irreducible. Suppose that for some vector $w \geq 0$ and scalars $\alpha, \beta \geq 0$ we have $\alpha w \leq Aw \leq \beta w$, with strict inequalities in at least one entry. Then $\alpha < \rho(A) < \beta$.*

Proof. Since A^T is nonnegative and irreducible (why?), then there exists a vector $y > 0$ such that $y^T A = \rho(A) y^T$. We have $y^T w > 0$ and moreover,

$$\alpha y^T w < y^T A w < \beta y^T w,$$

and the claim follows from the identity $y^T A w = \rho(A) y^T w$. ■

Finally, let z be a Perroneigenvector of B. We have $Az \lneqq Bz = \rho(B)z$, and the last part of Theorem 2.1 follows from the rightmost inequality of Lemma 2.5. ■

Exercise 2.6. Let $A \geq O$ be irreducible. Let $\alpha = \min_i \sum_j A_{ij}$ and $\beta = \max_i \sum_j A_{ij}$. Prove that (1) if $\alpha = \beta$ then $\rho(A) = \alpha$; (2) if $\alpha < \beta$ then $\alpha < \rho(A) < \beta$.

2.2 Primitive matrices

One more useful result belonging to the Perron–Frobenius theory is the following:

Theorem 2.7 (O. Perron). *If $A > O$ then there is only one eigenvalue with modulus $\rho(A)$.*

Proof. For some sufficiently small $\varepsilon > 0$, the matrix $A - \varepsilon I$ is still positive and, using Theorem 2.1, $\rho(A - \varepsilon I) = \rho(A) - \varepsilon$. On the other hand, if $\lambda \neq \rho(A)$ is another eigenvalue of A such that $|\lambda| = \rho(A)$, then $\lambda - \varepsilon$ is an eigenvalue of $A - \varepsilon I$ with $|\lambda - \varepsilon| > \rho(A - \varepsilon I)$, a contradiction. ■

It is possible to extend Theorem 2.7 to any matrix $A \geq O$ with the following property:

$$\text{There exists a positive integer } m \text{ such that } A^m > O. \tag{2}$$

In fact, if $\lambda \neq \rho(A)$ is an eigenvalue on the spectral circle of A then λ^m belongs to the spectral circle of A^m, thus violating Theorem 2.7.

Actually, Frobenius called *primitive* any matrix $A \geq O$ having $\rho(A)$ as the sole eigenvalue on the spectral circle $\{\lambda \in \mathbb{C} : |\lambda| = \rho(A)\}$, and proved that a matrix is primitive if and only if it fulfills the condition (2). Note that a primitive matrix is necessarily irreducible, but the converse is not true. Hence, a primitive, nonnegative matrix A has $\rho(A)$ as unique eigenvalue with largest modulus. All other eigenvalues are smaller in modulus.

Remark 2.8. *Theorem 2.7 allows us to conclude that if we apply the (normalized version of the) power method to a nonnegative, primitive matrix (starting from a positive vector, e.g., $\mathbf{1}$) then the method will converge to a Perron eigenpair. This fact is not true for a generic irreducible $A \geq O$. For example, examine the behaviour of the power method applied to the matrix*

$$A = \begin{pmatrix} 0 & 2 \\ 1/2 & 0 \end{pmatrix},$$

which is nonnegative and irreducible but not primitive (why?).

3 Complements: Bounding perturbations on Perron vectors

The forthcoming theorem provides a bound on the relative change in a Perron eigenvector when some matrix rows are changed. The result states that when a few rows of a nonnegative matrix are perturbed then the relative changes in the corresponding elements of the Perron vector bound the relative changes in the other elements:

Theorem 3.1 ([1]). *Let $A, \widehat{A}$ be irreducible, nonegative matrices, let $Ax = \rho x$ and $\widehat{A}\hat{x} = \hat{\rho}\hat{x}$ be corresponding Perron eigenpairs. Let $\mathcal{I}$ be the index set of unchanged rows:*

$$\mathcal{I} = \{i : A_{i,:} = \widehat{A}_{i,:}\}.$$

Hence,

$$\forall i \in \mathcal{I}, \qquad \frac{\rho}{\hat{\rho}} \min_{j=1\ldots n} \frac{\hat{x}_j}{x_j} \leq \frac{\hat{x}_i}{x_i} \leq \frac{\rho}{\hat{\rho}} \max_{j=1\ldots n} \frac{\hat{x}_j}{x_j}.$$

In particular, if $\hat{\rho} > \rho$ then $\max_{i\in\mathcal{I}} \frac{\hat{x}_i}{x_i} < \max_{j\notin\mathcal{I}} \frac{\hat{x}_j}{x_j}$ while if $\hat{\rho} < \rho$ then $\min_{i\in\mathcal{I}} \frac{\hat{x}_i}{x_i} < \min_{j\notin\mathcal{I}} \frac{\hat{x}_j}{x_j}$.

PROOF. Firstly, note that by hypotheses we have $x, \bar{x} > 0$. For any $i \in \mathcal{I}$ we have

$$\begin{aligned} \frac{\hat{x}_i}{x_i} = \frac{\hat{\rho}\hat{x}_i}{\hat{\rho}x_i} &= \frac{1}{\hat{\rho}x_i}\sum_k \widehat{A}_{ij}\hat{x}_j \\ &= \frac{1}{\hat{\rho}x_i}\sum_k A_{ij}x_j\frac{\hat{x}_j}{x_j} \\ &\le \frac{1}{\hat{\rho}x_i}\left[\max_j \frac{\hat{x}_j}{x_j}\right]\rho x_i = \frac{\rho}{\hat{\rho}}\left[\max_j \frac{\hat{x}_j}{x_j}\right]. \end{aligned}$$

The opposite inequality is obtained analogously. Furthermore, if $\rho/\hat{\rho} < 1$ then $\max_{i\in\mathcal{I}} \frac{\hat{x}_i}{x_i} < \max_j \frac{\hat{x}_j}{x_j}$, whence $\max_j \frac{\hat{x}_j}{x_j} = \max_{j\notin\mathcal{I}} \frac{\hat{x}_j}{x_j}$, and analogously for the other inequality when $\rho/\hat{\rho} > 1$. ■

Exercise 3.2. Prove the following result:[2] Let v be a nonnegative vector, let $B = A + e_i v^T$. If x, y are positive Perron vectors of A and B, respectively, then $y_i/x_i > y_j/x_j$ for $j \neq i$.

4 Applications: Epidemics on graphs

If the graph G represents a computer network, or a social network, and $A = A_G$, then the number $\rho(A)$ plays an important role in modelling (computer or biologic, respectively) virus propagation in G. The smaller $\rho(A)$, the larger the robustness of the network against the spread of viruses. Hereafter, I present a simple virus propagation model which has been discussed e.g., in [5, §2.7] and [4].[3]

Consider a virus spreading on G, where at each time step, a contagious node may infect its susceptible neighbors with probability μ (virus birth rate). At each time step, an infected node may also be cured with probability β (virus curing rate). If the number $p_i(t)$ measures the amount of infection of node i at time t, then the model is

$$p_i(t) = (1-\beta)p_i(t-1) + \mu \sum_{j:j\to i} p_j(t-1),$$

where $p(t) = (p_1(t), \ldots, p_n(t))^T$ and $p(0)$ is the initial state of infection. In matrix form,

$$p(t) = ((1-\beta)I + \mu A)p(t-1).$$

Exercise 4.1. Prove that $p(t) \to 0$ as $t \to \infty$ for all initial states $p(0)$ if and only if $\rho(A) < \beta/\mu$.
You can assume that A is irreducible, but the claim is true in the most general case.

5 Exercises and problems

Exercises marked with a star ($\star$) are requested for the final evaluation.

1. Let $A \geq O$ be irreducible. Prove that if $z \geq 0$ is an eigenvector of A then it is the Perron eigenvector of A.

   Hint: Let $y : y^T A = \rho(A) y^T$. Evaluate $\rho(A) y^T z$.

2. Let $A \geq O$. Suppose $(\lambda I - A)^{-1}$ exists and is nonnegative. Then $\lambda > \rho(A)$.

   Hint: Let (μ, x) be a Perron eigenpair of $(\lambda I - A)^{-1}$. Deduce $Ax = (\lambda - 1/\mu)x$ and prove that x is a Perron vector of A.

3. ($\star$) A graph $G = (V, E)$ is called *bipartite* if $V = V_1 \cup V_2$, $V_1 \cap V_2 = \varnothing$, and every edge belongs to either $V_1 \times V_2$ or $V_2 \times V_1$.

   (a) Prove that A_G is not primitive.

[2] Found in L. Elsner, C. Johnson, M. Neumann; *Czech. Math. J.* 32 (1982), 99–109.

[3] Another model of virus propagation in computer networks, leading to the same conclusion concerning $\rho(A)$, has been developed in Van Mieghem P., Omic J., Kooij R., Virus spread in networks, *IEEE/ACM Trans. Netw.* 17 (2009) 1–14.

(b) Assume that G is undirected. Find all eigenpairs (λ, x) of A_G with $|\lambda| = \rho(A_G)$.

4. ($\star$) Let $A \geq O$ be irreducible. Prove that if there exists an index i such that $A_{ii} > 0$ then A is primitive.

5. ($\star$) Let $O \leq A \lneqq B$ and let B be irreducible. Prove that $\rho(A) < \rho(B)$.

References

[1] Erik Dietzenbacher. Perturbations of matrices: a theorem on the Perron vector and its applications to input-output models. *J. of Economics*, 48 (1988), 389–412.

[2] Karl Meyer. *Matrix Analysis and Applied Linear Algebra.* SIAM, 2000.

[3] Gilbert W. Stewart. Perron–Frobenius Theory: A New Proof of the Basics. Technical report TR-94-79, University of Maryland, 1994. `ftp://thales.cs.umd.edu/pub/reports/pftnpb.ps`

[4] Wang Y., Chakrabarti D., Wang C., Faloutsos C. Epidemic spreading in real networks. *ACM Trans. Inf. System Security* 10 (2008), no. 13.

[5] D. Cvetković, S. Simić. Graph spectra in Computer Science. *Lin. Algebra Appl.*, 434 (2011), 1545–1562.

`Last update: August 29, 2014`