Vrije Universiteit Brussel

Faculty of Science
Master in Mathematics - specialization: Education

Similarity on Combinatorial Structures

Graduation thesis submitted in partial fulfillment of the requirements for
the degree of Master of Science in Mathematics - specialization: Education

Filip Moons

Promotor: Prof. Dr. P. Cara
Stagebegeleider: Promotor

Stagementor:

June 2015

Contents

Chapter 1

Preliminaries and notations

The Perron-Frobenius theorem states that a real square matrix with nonnegative entries has a unique largest real eigenvalue with an eigenvector that has only positive entries. The theorem was proved by Oskar Perron (1880-1975) in 1907 for strictly positive entries and extended by Ferdinand Georg Frobenius (1849-1917) to irreducible matrices with nonnegative entries.

1.1 Some families of matrices

In this section, we first introduce different kinds of matrices. We start with permutation matrices and their uses. With permutation matrices, we can introduce irreducible matrices. Also nonnegative and primitive square matrices are presented. After defining those, we look at the Perron-Frobenius theorem.

1.1.1 Permutation matrices

Definition 1.1.1. *Given a permutation π of n elements:*

$$\pi : \{1, \dots, n\} \to \{1, \dots, n\},$$

with:

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix}$$

the associated ***permutation matrix*** *P_π is the $n \times n$-matrix obtained by permuting the rows of the identity matrix I_n according to π. So:*

$$P_\pi = \begin{bmatrix} \mathbf{e}_{\pi(1)} \\ \mathbf{e}_{\pi(2)} \\ \vdots \\ \mathbf{e}_{\pi(n)} \end{bmatrix}.$$

where $\mathbf{e}_j$ is the j-th row of I_n.

Example 1.1.2. The permutation matrix P_π corresponding to the permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}$$

is:

$$P_\pi = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Note that $p_{ij} = 1$ if and only if $\pi(i) = j$.

Property 1.1.3. *A permutation matrix P satisfies:*

$$PP^T = I_n,$$

where P^T is the transpose and I_n is the identity matrix.

Proof. By direct computation, we get:

$$(PP^T)_{ij} = \sum_{k=1}^{n} P_{ik}P^T_{kj} = \sum_{k=1}^{n} P_{ik}P_{jk}$$

Assume $i \neq j$. Then for each k, $P_{ik}P_{jk} = 0$ since there is only one nonzero entry in the k-th row and $i \neq j$, P_{ik} and P_{jk} can't be both the nonzero entry. So, $(PP^T)_{ij} = 0$ when $i \neq j$.

When $i = j$, then there exists a $k' \in \{1, \ldots, n\}$ with $P_{ik'}P_{jk'} = 1$, since there is only one nonzero entry in the k-th row, this k' is unique, which results in $\sum_{k=1}^{n} P_{ik}P_{jk} = (PP^T)_{ij} = 1$.

In other words,

$$(PP^T)_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases},$$

this is exactly the formula for the entries of the identity matrix. □

Corollary 1.1.4. *The transpose of a permutation matrix P is its inverse:*

$$P^T = P^{-1}.$$

This can also more easily be concluded by the fact that a permutation matrix is clearly an orthogonal matrix (a real $n \times n$-matrix with orthonormal entries).

1.1.2 Nonnegative and primitive matrices

Definition 1.1.5. *Let A and B be two real $n \times r$-matrices. Then, $A \geq B$ (respectively $A > B$) if $a_{ij} \geq b_{ij}$ (respectively $a_{ij} > b_{ij}$) for all $1 \leq i \leq n, 1 \leq j \leq r$.*

Definition 1.1.6. *A real $n \times r$-matrix A is **nonnegative** if $A \geq 0$, with 0 the n-null matrix.*

Definition 1.1.7. *A real $n \times r$-matrix A is **positive** if $A > 0$, with 0 the n-null matrix.*

Since row vectors are $1 \times n$-matrices, we shall use the terms nonnegative and positive vector throughout.

Notation 1.1.8. *Let B be an arbitrary complex $n \times r$-matrix, then $|B|$ denotes the matrix with entries $|b_{ij}|$. This is not to be confused with the determinant of a square matrix B, which we denote by* $\det(B)$.

Definition 1.1.9. *A nonnegative square matrix A is called **primitive** if there is a $k \in \mathbb{N}_0$ such that all entries of A^k are positive.*

1.1.3 Irreducible nonnegative matrices

In developing the Perron-Frobenius theory, we shall first establish a series of theorems and lemmas on nonnegative irredicuble square matrices.

Definition 1.1.10. *A square matrix A is called* ***reducible*** *if there is a permutation matrix P such that*

$$PAP^T = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$

where B, and D are square matrices, each of size at least one and 0 is a zero matrix. A square matrix A is called ***irreducible*** *if it is not reducible.*

It follows immediately that a 1×1-matrix is always irreducible by definition. We now show a useful property to identify a reducible matrix.

Property 1.1.11. *Let A be an $n \times n$-matrix with $n \geq 2$. Consider a nonempty, proper subset S of $\{1, \dots, n\}$ with $a_{ij} = 0$ for $i \in S$, $j \notin S$. Then A is reducible.*

Proof. Let $S = \{i_1, i_2, \dots, i_k\}$, where we assume, without loss of generality, that $i_1 < i_2 < \dots < i_{k-1} < i_k$. Let S^c be the complement of S, consisting of the ordered set of elements $j_1 < j_3 < \dots < j_{n-k}$. Consider the permutation σ of $\{1, 2, \dots, n\}$ given by

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & k & k+1 & k+2 & \dots & n \\ i_1 & i_2 & \dots & i_k & j_1 & j_2 & \dots & j_{n-k} \end{pmatrix}$$

σ can be represented by the permutation matrix $P_\sigma = (p_{ij})$, where $p_{rs} = 1$ if $\sigma(r) = s$. We prove that

$$PAP^T = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$

where B and D are square matrices and 0 is a $k \times (n-k)$ zero matrix. Consider row c and column d, where $1 \leq c \leq k$ and $k+1 \leq d \leq n$:

$$(PAP^T)_{cd} = \sum_i \sum_j p_{ci} a_{ij} p_{dj}. \tag{1.1}$$

It is enough to show that each term in the summation is zero. Suppose $p_{ci} = p_{dj} = 1$. Thus $\sigma(c) = i$ and $\sigma(d) = j$. Since $1 \leq c \leq k$, then $i \in \{i_1, i_2, \dots, i_k\}$; similarly, since $k+1 \leq d \leq n$, we have $j \in \{j_1, j_2, \dots, j_{n-k}\}$. By assumption, for such a pair i, j, we have $a_{ij} = 0$. That completes the proof. $\square$

We know prove some equivalent definitions for a nonnegative, irreducible square matrix.

Theorem 1.1.12. *Let $A \geq 0$ be a nonnegative $n \times n$-matrix. Then the following conditions are equivalent:*

(1) A is irreducible.

(2) $(I + A)^{n-1} > 0$

(3) For any pair (i, j), with $1 \leq i, j \leq n$, there is a positive integer $k = k(i, j) \leq n$ such that $(A^k)_{ij} > 0$.

Proof. (1) $\Rightarrow$ (2): Let $\mathbf{x} \geq 0, \mathbf{x} \neq \mathbf{0}$ be an arbitrary vector in $\mathbb{R}^n$. If a coordinate of $\mathbf{x}$ is positive, the same coordinate is positive in $\mathbf{x} + A\mathbf{x} = (I + A)\mathbf{x}$ as well. We claim that $(I + A)\mathbf{x}$ has fewer zero coordinates than $\mathbf{x}$ as long as $\mathbf{x}$ has a zero coordinate. If this claim is not true, then the number of zero coordinates must be at least equal, this means that for each coordinate j with $x_j = 0$ we would have that $x_j + (A\mathbf{x})_j = 0$. Let $J = \{j : x_j > 0\}$. For any $j \notin J, r \in J$, we have $(A\mathbf{x})_j = \sum_k a_{jk}x_k = 0$ and $x_r > 0$. It must be that $a_{jr} = 0$. It follows from Property 1.1.11 that A is reducible, which is a contradiction and the claim is proved. Thus $(I + A)\mathbf{x}$ has at most $n - 2$ zero coordinates. Continuing in this manner we conclude that $(I + A)^{n-1}\mathbf{x} > 0$. Let $\mathbf{x} = \mathbf{e}_i$, then the corresponding column of $(I + A)^{n-1}$ must be positive. Thus (2) holds.

(2) $\Rightarrow$ (3): We have $(I + A)^{n-1} > 0$, $A \geq 0$, so $A \neq 0$ and

$$A(I + A)^{n-1} = \sum_{k=1}^{n} \binom{n-1}{k-1} A^k > 0.$$

Thus for any i, j at least one of the matrices $A, A^2, \ldots, A^n$ has its (i, j)-th element entry positive.

(3) $\Rightarrow$ (1): Suppose A is reducible. Then for some permutation matrix P,

$$PAP^T = \begin{pmatrix} B_1 & 0 \\ C_1 & D_1 \end{pmatrix}$$

where B_1 and D_1 are square matrices. Furthermore, we know from Property 1.1.3 that $PAP^TPAP^T = PA^2P^T$, whence for some square matrices B_2, C_2 we have:

$$PA^2P^T = \begin{pmatrix} B_2 & 0 \\ C_2 & D_2 \end{pmatrix}$$

More generally, for some matrix C_t and square matrices B_t and D_t,

$$PA^tP^T = \begin{pmatrix} B_t & 0 \\ C_t & D_t \end{pmatrix}$$

Thus $(PA^tP^T)_{rs} = 0$ for $t = 1, 2, \ldots$ and for any r, s corresponding to an entry of the zero submatrix in PAP^T. Now, for $t = 1, \ldots, n$:

$$0 = (PA^tP^T)_{rs} = \sum_k \sum_l p_{rk} a_{kl}^{(t)} p_{st}$$

By using the same reasoning as in 1.1, choose k, l so that $p_{rk} = p_{sl} = 1$. Then $a_{kl}^{(t)} = 0$ for all t, contradicting the hypothesis. This completes the proof. $\square$

Corollary 1.1.13. *If A is irreducible then $I + A$ is primitive.*

Corollary 1.1.14. *A^T is irreducible whenever A is irreducible.*

Property 1.1.15. *No row or column of an irreducible matrix A can vanish. This means that A cannot have a row or a column of zeros.*

Proof. Suppose that A has a zero row, then it could be permuted to

$$PAP^T = \begin{pmatrix} 0 & 0\dots0 \\ c_1 & \\ \vdots & D \\ c_n & \end{pmatrix}$$

by some permutation matrix P. It follows from Definition 1.1.10 that A is reducible. Similarly, if A has zero column, it can be permuted to

$$PAP^T = \begin{pmatrix} & 0 \\ B & \vdots \\ & 0 \\ c_1\dots c_n & 0 \end{pmatrix}$$

again from Definition 1.1.10 we conclude that A is reducible.

□

1.2 Perron-Frobenius Theorem

1.2.1 Spectral radii of nonnegative matrices

Definition 1.2.1. *Let A be an $n \times n$-matrix with complex entries and eigenvalues λ_i, $1 \leq i \leq n$. Then:*

$$\rho(A) = \max_{1\leq i\leq n} |\lambda_i|$$

*is called the **spectral radius** of the matrix A.*

Geometrically, if all the eigenvalues λ_i of A are plotted in the complex plane, then $\rho(A)$ is the radius of the smallest disk $|z| \leq R$, with center at the origin, which includes all the eigenvalues of the matrix A.

We now establish a series of lemmas on nonnegative irreducible square matrices. These lemmas will allow us to prove the Perron-Frobenius at the end of this section.

If $A \geq 0$ is an irreducible $n \times n$-matrix and $\mathbf{x}$, a vector of size n with $0 \neq \mathbf{x} \geq 0$, let

$$r_{\mathbf{x}} = \min \left\{ \frac{\sum_{j=1}^{n} a_{ij} x_j}{x_i} \right\} \tag{1.2}$$

where the minimum is taken over all i for which $x_i > 0$. Clearly, $r_{\mathbf{x}}$ is a nonnegative real number and is the supremum of all $p \geq 0$ for which

$$A\mathbf{x} \geq p\mathbf{x} \tag{1.3}$$

We now consider the nonnegative quantity r defined by

$$r = \sup_{\substack{\mathbf{x}\geq 0 \\ \mathbf{x}\neq \mathbf{0}}} \{r_{\mathbf{x}}\} \tag{1.4}$$

As $r_{\mathbf{x}}$ and $r_{\alpha\mathbf{x}}$ have the same value for any scalar $\alpha > 0$, we need consider only the set B of vectors $\mathbf{x} \geq 0$ with $||\mathbf{x}|| = 1$, and we correspondingly let Q be the set of all vectors $\mathbf{y} = (I + A)^{n-1}\mathbf{x}$ where $\mathbf{x} \in B$. From Theorem 1.1.12 , Q consists only of positive vectors. Multiplying both sides of the inequality $A\mathbf{x} \geq r_{\mathbf{x}}\mathbf{x}$ by $(I + A)^{n-1}$, we obtain:

$$\forall \mathbf{y} \in Q : A\mathbf{y} \geq r_{\mathbf{x}}\mathbf{y},$$

and we conclude from (1.3) that $r_{\mathbf{y}} \geq r_{\mathbf{x}}$. Therefore, the quantity r of (1.4) can be defined equivalently as:

$$r = \sup_{\mathbf{y} \in Q}\{r_{\mathbf{y}}\} \tag{1.5}$$

As B is a compact set (in the usual topology) of vectors, so is Q, and as $r_{\mathbf{y}}$ is a continuous function on Q, we know from the extreme value theorem that there necessarily exists a positive vector $\mathbf{z}$ for which:

$$A\mathbf{z} \geq r\mathbf{z}, \tag{1.6}$$

and no vector $\mathbf{w} \geq 0$ exists for which $A\mathbf{w} > r\mathbf{w}$.

Definition 1.2.2. *We call all nonnegative, nonzero vectors* $\mathbf{z}$ *satisfying (1.6)* ***extremal vectors*** *of the matrix* A.

Lemma 1.2.3. *If* $A \geq 0$ *is an irreducible* $n \times n$*-matrix, the quantity* r *of (1.4) is positive.*

Proof. If $\mathbf{x}$ is the positive vector whose coordinates are all unity, then since the matrix A is irreducible, we know from Property 1.1.15 that no row of A can vanish, and consequently no component of $A\mathbf{x}$ can vanish. Thus, $r_{\mathbf{x}} > 0$, proving that $r > 0$. □

Lemma 1.2.4. *If* $A \geq 0$ *is an irreducible* $n \times n$*-matrix, each extremal vector* $\mathbf{z}$ *is a positive eigenvector of* A *with corresponding eigenvalue* r *of (1.4), i.e.,* $A\mathbf{z} = r\mathbf{z}$ *and* $\mathbf{z} > 0$.

Proof. Let $\mathbf{z}$ be an extremal vector with $A\mathbf{z} - r\mathbf{z} = \mathbf{t}$. If $\mathbf{t} \neq \mathbf{0}$, then some coordinate of $\mathbf{t}$ is positive; multiplying through by the matrix $(I + A)^{n-1}$, we have:

$$A\mathbf{w} - r\mathbf{w} > 0, \text{ with } \mathbf{w} = (I + A)^{n-1}\mathbf{z}$$

from Theorem 1.1.12 we know that $\mathbf{w} > 0$. It would then follow that $r_{\mathbf{w}} > r$, contradicting the definition of r in (1.5). Thus $A\mathbf{z} = r\mathbf{z}$, and since $\mathbf{w} > 0$ and $\mathbf{w} = (1 + r)^{n-1}\mathbf{z}$, then we have $\mathbf{z} > 0$, completing the proof. □

Lemma 1.2.5. *Let* $A \geq 0$ *be an irreducible* $n \times n$*-matrix, and let* B *be an* $n \times n$*- complex matrix with* $|B| \leq A$. *If* β *is any eigenvalue of* B, *then*

$$|\beta| \leq r, \tag{1.7}$$

where r *is the positive quantity of (1.4). Moreover, equality is valid in (1.7), i.e.,* $\beta = re^{i\phi}$, *if and only if* $|B| = A$, *and where* B *has the form:*

$$B = e^{i\phi}DAD^{-1}, \tag{1.8}$$

and D *is a diagonal matrix whose diagonal entries have modulus unity.*

Proof. If $\beta\mathbf{y} = B\mathbf{y}$ where $\mathbf{y} \neq \mathbf{0}$, then

$$\beta y_i = \sum_{j=1}^{n} b_{ij} y_i, \text{ with } 1 \leq i \leq n.$$

Using the hypotheses of the lemma and the notation of Definition 1.1.8, it follows that:

$$|\beta||\mathbf{y}| \leq |B||\mathbf{y}| \leq A|\mathbf{y}|, \tag{1.9}$$

which implies that $|\beta| \leq r_{|\mathbf{y}|} \leq r$, proving (1.7). If $|\beta| = r$, then $|\mathbf{y}|$ is an extremal vector of A. Therefore, from Lemma 1.2.4, $|\mathbf{y}|$ is a positive eigenvector of A corresponding to the positive eigenvalue r. Thus,

$$r|\mathbf{y}| = |B||\mathbf{y}| = A|\mathbf{y}|, \tag{1.10}$$

and since $|\mathbf{y}| > 0$, we conclude from (1.10) and the hypothesis $|B| \leq A$ that

$$|B| = A \tag{1.11}$$

For the vector $\mathbf{y}$, where $|\mathbf{y}| > 0$, let

$$D = \operatorname{diag}\left\{\frac{y_1}{|y_1|}, \ldots, \frac{y_n}{|y_n|}\right\}.$$

It is clear that the diagonal entries of D have modulus unity, and

$$\mathbf{y} = D|\mathbf{y}|. \tag{1.12}$$

Setting $\beta = re^{i\phi}$, then $B\mathbf{y} = \beta\mathbf{y}$ can be written as:

$$C|\mathbf{y}| = r|\mathbf{y}|, \tag{1.13}$$

where

$$C = e^{-i\phi} D^{-1} B D. \tag{1.14}$$

From (1.10) and (1.13), equiting terms equal to $r|\mathbf{y}|$ we have

$$C|\mathbf{y}| = |B||\mathbf{y}| = A|\mathbf{y}|. \tag{1.15}$$

From the definition of the matrix C in (1.14), $|C| = |B|$. Combining with (1.11), we have:

$$|C| = |B| = A. \tag{1.16}$$

Thus, from (1.15) we conclude that $C|\mathbf{y}| = |C||\mathbf{y}|$, and as $|\mathbf{y}| > 0$, it follows that $C = |C|$ and thus $C = A$ from (1.16). Combining this result with (1.14), gives the desired result that $B = e^{i\phi} DAD^{-1}$. Conversely, it is obvious that if B has the form in (1.8), then $|B| = A$, and B has an eigenvalue β with $|\beta| = r$, which completes the proof. □

Corollary 1.2.6. *If $A \geq 0$ is an irreducible $n \times n$-matrix, then the positive eigenvalue r of Lemma 1.2.4 equals the spectral radius $\rho(A)$ of A*

Proof. Setting $B = A$ in Lemma 1.2.5 immediately gives us this result. □

In other words, if $A \geq 0$ is an irreducible $n \times n$-matrix, its spectral radius $\rho(A)$ is positive, and the intersection in the complex plane of the circle $|z| = \rho(A)$ with the positive real axis is an eigenvalue of A.

Definition 1.2.7. *A **principal square submatrix** of an $n \times n$-matrix A is any matrix obtained by crossing out any j rows and the corresponding j columns of A, with $1 \leq j \leq n$.*

Lemma 1.2.8. *If $A \geq 0$ is an irreducible $n \times n$-matrix, and B is any principal square submatrix of A, then $\rho(B) < \rho(A)$.*

Proof. If B is any principal submatrix of A, then there is an $n \times n$-permutation matrix P such that $B = A_{11}$ where

$$C = \begin{pmatrix} A_{11} & 0 \\ 0 & 0 \end{pmatrix}; PAP^T = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \tag{1.17}$$

Here, A_{11} and A_{22} are, respectively, $m \times m$ and $(n-m) \times (n-m)$ principal square submatrices of PAP^T, $1 \leq m \leq n$. Clearly, $0 \leq C \leq PAP^T$, and $\rho(C) = \rho(B) = \rho(A_{11})$, but as $C = |C| \neq PAP^T$, the conclusion follows immediately from Lemma 1.2.5 and Corollary 1.2.6. ☐

The following lemma is used to prove that $\rho(A)$ is a simple eigenvalue of A in the Perron-Frobenius theorem. The proof uses the extension of the product rule of derivation for multilinear functions $M(a_1, \ldots, a_k)$. Suppose $x_1, .., x_k$ are differentiable vector functions, then $M(x_1, \ldots, x_k)$ is differentiable and:

$$\frac{\mathrm{d}}{\mathrm{d}t} M(x_{1,k}) = M(\frac{\mathrm{d}}{\mathrm{d}t} x_1, x_2, \ldots, x_k) + M(x_1, \frac{\mathrm{d}}{\mathrm{d}t} x_2, \ldots, x_k) + \ldots + M(x_1, x_2, \ldots, \frac{\mathrm{d}}{\mathrm{d}t} x_k)$$

The most important application of this rule is for the derivative of the determinant:

$$\frac{\mathrm{d}}{\mathrm{d}t} \det(x_1, \ldots, x_k) = \det(\frac{\mathrm{d}}{\mathrm{d}t} x_1, x_2, \ldots, x_k) + \det(x_1, \frac{\mathrm{d}}{\mathrm{d}t} x_2, \ldots, x_k) + \ldots + \det(x_1, x_2, \ldots, \frac{\mathrm{d}}{\mathrm{d}t} x_k)$$

Lemma 1.2.9. *Let A be an $n \times n$-matrix over the complex numbers and let $\phi(A, \lambda) = \det(\lambda I_n - A)$ be the characteristic polynomial of A. Let B_i be the principal submatrix of A formed by deleting the i-th row and column of A and let $\phi(B_i, \lambda)$ be the characteristic polynomial of B_i. Then:*

$$\phi'(A, \lambda) = \frac{\mathrm{d}\phi(A, \lambda)}{\mathrm{d}\lambda} = \sum_i \phi(B_i, \lambda)$$

Proof. The proof is immediately done by direct computation:

$$\phi(A, \lambda) = \det \begin{bmatrix} \lambda - a_{1,1} & -a_{1,2} & \ldots & -a_{1,n} \\ -a_{2,1} & \lambda - a_{2,2} & \ldots & -a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n,1} & -a_{n,2} & \ldots & \lambda - a_{n,n} \end{bmatrix}.$$

Using the extension of the product rule of derivation for multilinear functions

$$\phi'(A,\lambda) = \det\begin{bmatrix} 1 & 0 & \dots & 0 \\ -a_{2,1} & \lambda - a_{2,2} & \dots & -a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n,1} & -a_{n,2} & \dots & \lambda - a_{n,n} \end{bmatrix} + \det\begin{bmatrix} \lambda - a_{1,1} & -a_{1,2} & \dots & -a_{1,n} \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n,1} & -a_{n,2} & \dots & \lambda - a_{n,n} \end{bmatrix} + \dots$$

$$+ \det\begin{bmatrix} \lambda - a_{1,1} & -a_{1,2} & \dots & -a_{1,n} \\ -a_{2,1} & \lambda - a_{2,2} & \dots & -a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = \sum_i \phi(B_i, \lambda).$$

□

We now collect the above results into the following main theorem: we finally arrived at the Perron-Frobenius Theorem:

Theorem 1.2.10. ***(Perron-Frobenius theorem)*** *Let $A \geq 0$ be an irreducible $n \times n$-matrix. Then,*

1. *A has a positive real eigenvalue equal to its spectral radius.*
2. *To $\rho(A)$ there corresponds an eigenvector $\mathbf{x} > 0$.*
3. *$\rho(A)$ increases when any entry of A increases.*
4. *$\rho(A)$ is a simple eigenvalue of A.*
5. *If $A\mathbf{x} = \rho(A)\mathbf{x}$ where $\mathbf{x} > 0$ and $\mathbf{x}$ is a normalized vector, then $\mathbf{x}$ is unique.*

Proof. (1) and (2) follow immediately from Lemma 1.2.4 and Corollary 1.2.6.

(3) Suppose we increase some entry of the matrix A, giving us a new irreducible matrix $\tilde{A}$ where $\tilde{A} \geq A$ and $\tilde{A} \neq A$. Applying Lemma 1.2.5, we conclude that $\rho(\tilde{A}) > \rho(A)$.

(4) $\rho(A)$ is a simple eigenvalue of A, i.e., $\rho(A)$ is a zero of multiplicity one of the characteristic polynomial $\phi(\lambda) = \det(\lambda I_n - A)$, we make use of Lemma 1.2.9 by using the fact that $\phi'(\lambda)$ is the sum of the determinants of the principal $(n-1) \times (n-1)$ submatrices of $\lambda I - A$. If A_i is any principal submatrix of A, then from Lemma 1.2.8, $\det(\lambda I - A_i)$ (with I the identity matrix with the same size as the principal submatrix A_i) cannot vanish for any $\lambda \geq \rho(A)$. From this it follows that:

$$\det(\rho(A)I - A_i) > 0,$$

and thus

$$\phi'(\rho(A)) > 0.$$

Consequently, $\rho(A)$ cannot be z zero of $\phi(\lambda)$ of multiplicity greater than one and thus $\rho(A)$ is a simple eigenvalue of A.

(5) If $A\mathbf{x} = \rho(A)\mathbf{x}$ where $\mathbf{x} > 0$ and $||x|| = 1$ ($||x||$ denotes the standard Euclidean norm), we cannot find another eigenvector $\mathbf{y} \neq s\mathbf{x}$, with s a scalar, of A with $A\mathbf{y} = \rho(A)\mathbf{y}$, so that the eigenvector $\mathbf{x}$, meaning that the normalized eigenvector $\mathbf{x}$ is uniquely determined. □

1.2.2 Example

To check wether a matrix with nonnegative entries is primitive, irreducible or neither, we just have to replace all nonzero entries by 1 since this does not affect the classification. The matrix

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

is strictly positive and thus primitive. The matrices

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

both have 1 as a double eigenvalue hence can not be irreducible. The matrix $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ satisfies:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

and hence is primitive. The same goes for

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

this matrix is irreducible but not primitive. Its eigenvalues are 1 and -1.

1.3 Graphs

After introducing different kinds of matrices and proving the Perron-Frobenius theorem, we now take a closer look at graphs. Here too, we'll look at different families of graphs and prove some relevant properties about them. We also link the concept of graphs with different kinds of matrices, deepening our insight of some theorems of the previous section.

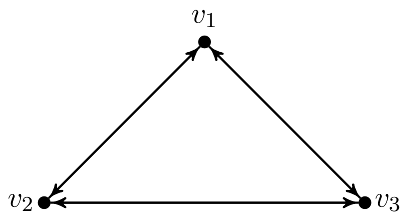
The definitions and results in this section are mainly based on the course 'Discrete Mathematics' by P. Cara [CARA].

1.3.1 General definitions

Definition 1.3.1. *A* ***graph*** *is a an ordered pair $(V, \to)$ where V is a set and $\to$ is a relation. The elements of V are called* ***vertices*** *and $\to$ is called the* ***adjacency relation****. Let $u, v \in V$, then the couple (u, v) belonging to $\to$ is called an* ***arc*** *or* ***edge*** *and we write $u \to v$. We also say that u is* ***adjacent*** *to v. When $v \to v$ (with $v \in V$) we say that the graph has a* ***loop*** *in v. A graph is most of the time denoted by calligraphic letters $\mathscr{G}, \mathscr{H}, \ldots$ Therefore we denote $\mathscr{G}(V, \to)$ for the graph $\mathscr{G}$ with vertex set V and adjancency relation $\to$. When handling multiple graphs, $V(\mathscr{G})$ denotes the set of vertices of $\mathscr{G}$.*

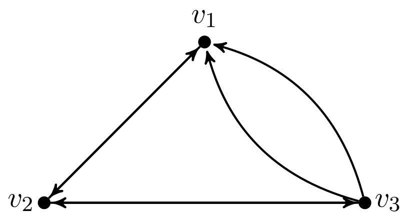
When the relation $\to$ is symmetric, we call the graph **undirected**, in this case we often write $\sim$ instead of $\to$.

Example 1.3.2. The graph



is an undirected graph with vertices v_1, v_2, v_3. The adjacency relation $\to$ equals $\{(v_1, v_2), (v_2, v_1), (v_2, v_3), (v_3, v_2), (v_3, v_1), (v_1, v_3)\}$.

There is a small problem with our definition, because not all graphs are taken into account, by example the graph below is not a graph following our definition because you can not define multiple edges between vertices in a relation.



Therefore we introduce a more general definition and introduce the concept of multiplicity of an edge:

Definition 1.3.3. *A graph is an ordered pair (V, μ) with V a set and $\mu : V \times V \to \mathbb{N}$ a function that gives the **multiplicity** of an edge. The function is defined as follows:*

- *when $\mu(u, v) = 0$ we say that u and v are not adjacent;*
- *when $\mu(u, v) = k > 0$ we say that there are k edges from u to v.*

It is clear that our previous definition fits perfectly in this more general definition, by constructing μ in this case as follows:

$$\mu(u, v) = \begin{cases} 1 & \text{if } (u, v) \in \to \\ 0 & \text{otherwise.} \end{cases}$$

Definition 1.3.4. *The **neighbourhood** of a vertex v of a graph $\mathscr{G}(V, \mu)$ is the induced subgraph $\mathscr{G}_v$ with vertex set V' consisting of all vertices adjacent to v without v itself and with the multiplicity function μ', which is the restriction of μ to the vertices in V'. A vertex with a neighbourhood equal to the empty graph (a graph with an empty set of vertices) is called* isolated.

Definition 1.3.5. *The **order of a finite graph** $\mathscr{G}$ is the number of vertices of $\mathscr{G}$ and is denoted by $|\mathscr{G}|$.*

Definition 1.3.6. *The **degree of a vertex** v **in a graph** $\mathscr{G}$ is the number of edges containing v and is denoted by* $\deg(v)$*, so:*

$$\deg v = |\mathscr{G}_v|$$

Definition 1.3.7. *A* ***walk*** *in a graph $\mathscr{G}$ is a sequence of vertices*

$$v_0, v_1, \dots, v_k$$

such that $v_{i-1} \to v_i$ for each $i \in \{1, \dots, k\}$. The ***length*** *of the walk is k, one less than the number of vertices.*

Definition 1.3.8. *If all edges are distinct in a walk in a graph $\mathscr{G}$, we call the walk a* ***path****.*

Definition 1.3.9. *A* ***cycle*** *is a walk from v_0 to v_0 in which all vertices except v_0 are distinct.*

Definition 1.3.10. *A graph $\mathscr{G}$ is* ***connected*** *if it possible to establish a path from any vertex to any other vertex.*

Product graphs

Definition 1.3.11. *Take two graphs $\mathscr{G}(U, \to), \mathscr{H}(V, \to')$, the* ***product graph*** *$\mathscr{G} \times \mathscr{H}$ is the graph with $|\mathscr{G}|.|\mathscr{H}|$ vertices and that has an edge between vertices (u_i, v_j) and (u_k, v_l) if there is an edge between u_i and u_k in $\mathscr{G}$ and there is an edge between v_j and v_l in $\mathscr{H}$.*

Adjacency matrices

We now represent an undirected graph in the form of an adjacency matrix. This matrix gives a lot of useful information about the graph and vice versa.

Definition 1.3.12. *Let $\mathscr{G}(V, \sim)$ be an undirected graph of order n and define a numbering on the vertices $v_1, \dots, v_n$. Then the* ***adjacency matrix*** *$A_{\mathscr{G}}$ of $\mathscr{G}$ is the real $n \times n$-matrix with a_{ij} equal to the number of edges between i and j.*

Theorem 1.3.13. *Let $k > 0$. The element on place (i, j) in $A_{\mathscr{G}}^k$ contains the number of walks of length k from i to j.*

Proof. By induction on k.

For $k = 1$ we count the walks of length 1. These are edges and the result follows immediately from the defintion of $A_{\mathscr{G}}$.

Let v_l be a vertex of $\mathscr{G}$. If there are b_{ij} walks of length k from i to l and a_{lj} walks of length 1 (edges) from v_l to v_j, then there are $b_{il}a_{lj}$ walks of length $k+1$ from v_i to v_j passing vertex v_l. Therefore, the number of walks of length $k+1$ between v_i and v_j is equal to:

$$\sum_{l \in V(\mathscr{G})} b_{il}a_{lj} =: c_{ij}.$$

By the induction hypothesis we now that b_{il} equals the element on place (i, l) in $A_{\mathscr{G}}^k$ so c_{ij} is exactly the element on place (i, j) in the matrix product

$$A_{\mathscr{G}}^k A_{\mathscr{G}} = A_{\mathscr{G}}^{k+1}.$$

□

Example 1.3.14. The adjacency matrix of the graph in Example 1.3.2 is:

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

1.3.2 Simple graphs

We now introduce simple graphs and show that the adjacency matrix of a simple graph $\mathscr{G}$ is irreducible if and only if $\mathscr{G}$ is connected.

Definition 1.3.15. *A* ***simple graph*** *is an undirected graph $\mathscr{G}(V,\mu)$ containing no loops and for all vertices $v_i, v_j \in V$, we have that the multiplicity $\mu(v_i, v_j)$ is at most 1.*

Theorem 1.3.16. *Let $\mathscr{G}$ be a simple graph with n vertices and adjacency matrix A. Then $\mathscr{G}$ is connected if and only if A is irreducible.*

Proof. A path between two vertices in the simple graph $\mathscr{G}$ has at most $n-1$ edges. So, $\mathscr{G}$ is connected if and only if $\forall i, j \in V(\mathscr{G})$ there exists a $k \leq n-1$ with a path of length k from i to j. So $\forall i, j \in V(\mathscr{G}) : \exists k < n : (A^k_{\mathscr{G}})_{ij} > 0$. Because $(I_n + A)^{n-1} = \sum_{k=0}^{n-1} \binom{n-1}{k} A^k$, we know from Theorem 1.1.12 that A is irreducible. $\square$

1.3.3 Directed graphs

In this section, we take a closer look at *directed graphs.* We already showed that the adjacency matrix of a simple graph $\mathscr{G}$ is irreducible if and only if $\mathscr{G}$ is connected. But this is not appropriate for checking the irreducibility of matrices using graphs, because the only matrices we can check are the ones containing entries equal to 0 or 1. Can we find a more general method for checking irreducibility of a matrix using graphs? Luckily, we can and that's where directed graphs come in: we will introduce a method for turning a matrix in to a directed graph and vice versa and link the irreducibility of this matrix to a property of the directed graph.

Definition 1.3.17. *A graph $\mathscr{G}(V,\rightarrow)$ is* ***directed*** *when the relation $\rightarrow$ is not symmetric. A directed graph $\mathscr{G}(V,\rightarrow)$ is* ***connected*** *if the underlying undirected graph (remove all arrows on the edges) is connected. $\mathscr{G}$ is* ***strongly connected*** *if for any two vertices v_i and v_j there exists a* ***directed path****:*

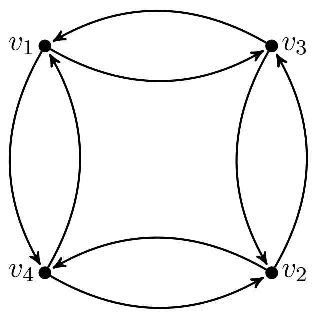
$$v_i \rightarrow v_{l_1}, v_{l_1} \rightarrow v_{l_2}, \ldots, v_{l_{r-1}} \rightarrow v_{l_r=j}.$$

We now link directed graphs with $n \times n$-matrices and vice versa. Consider any $n \times n$-matrix A, and consider any n distinct points $v_1, v_2, \ldots, v_n$ in the plane, which will be the nodes of the directed graph. For every nonzero entry a_{ij} of the matrix, we connect the node v_i to the node v_j by means of an arc $v_i \rightarrow v_j$ directed from v_i to v_j. In this way, with every $n \times n$-matrix A can be associated a **directed** graph $\mathscr{G}(A)$.

Example 1.3.18. Consider the matrix:

$$B = \begin{pmatrix} 0 & 0 & 89 & 7 \\ 0 & 0 & 2 & 2 \\ 123 & 9 & 0 & 0 \\ 14 & 89 & 0 & 0 \end{pmatrix}.$$

We get the directed graph:



Notice that this graph is strongly connected.

In the next proof, we study the equivalence of the matrix property of irreducibility of Definition 1.1.10 with the concept of the strongly connected directed graphs of a matrix:

Theorem 1.3.19. *An $n \times n$-matrix A is irreducible if and only if its directed graph $\mathscr{G}(A)$ is strongly connected.*

Proof. First assume that A is reducible. Consider:

$$PAP^T = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$

where D is a $k \times k$-submatrix and $1 \leq k \leq n-1$, and B is a square matrix of size at least 1. Consider any directed path from vertex v_i with $k < i$. The last segment of the path is determined by the presence of a nonzero element PAP^T_{ij} in the ith row of PAP^T. This row has zeros in the last k entries, so it is possible to make a connection from node i to node j only if $j < k$. Similarly, the path can be extended from node j only to another node smaller than k. Continuing in this way, we found that a directed path from node i with $i < k$ cannot be connected to a node greater than $k + 1$. Hence the directed graph $\mathscr{G}(PAP^T)$ is not strongly connected. The directed graph $\mathscr{G}(A)$ is also not connected: observe that the graph of PAP^T is obtained from that of A just by renumbering the nodes, and this operation does not affect the connectedness of the graph.

Now assume that $\mathscr{G}(A)$ is not strongly connected. Then there are distinct vertices v_i and v_j of $\mathscr{G}(A)$ for which there is no directed path from v_i to v_j. Let W_1 consist of v_j and all vertices from which there is directed path to v_j, let W_2 consist of v_i and all vertices to which there is a directed path from v_i. The sets W_1 and W_2 are nonempty and disjoint. Let W_3 be the set consisting of those vertices which belong to neither W_1 nor W_2 ($W_3 = V \backslash (W_1 \bigcup W_2)$). We now permute A so that the rows and columns corresponding to the vertices in W_2 come first followed by those corresponding to the vertices in W_3:

$$\begin{array}{c} \\ W_2 \\ W_3 \\ W_1 \end{array} \begin{array}{c} \begin{array}{ccc} W_2 & W_3 & W_1 \end{array} \\ \begin{pmatrix} B_1 & B_2 & O_1 \\ B_3 & B_4 & O_2 \\ C_3 & C_4 & D \end{pmatrix} \end{array}$$

Since there is no directed walk from v_i to v_j there is no arc from a vertex in W_2 to a vertex in W_1. Also there is no arc from a vertex v_k in W_3 to a vertex in W_1, because such an arc implies that v_k belongs to W_1. Hence $O_1 = 0$ and $O_2 = 0$, D is a square matrix and so is

$$B = \begin{array}{c c} & \begin{array}{cc} W_2 & W_3 \end{array} \\ \begin{array}{c} W_2 \\ W_3 \end{array} & \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} \end{array}.$$

Hence, A is reducible. $\square$

Chapter 2

Similarity on graphs

In the previous chapter all the basic terminology and results were introduced, now we take an extensive look at the concept of similarity on graphs. In the first section, we introduce similarity on directed graphs. More precisely, we define a similarity matrix S of to matrixes $\mathscr{G}, \mathscr{H}$ whose real entries s_{ij} express how similar a vertex v_j of $\mathscr{G}$ is to vertex v_i' of $\mathscr{H}$. In the second section, we take a closer look at similarity on colored graphs. Both sections conclude with an extensive application of the results.

2.1 Similarity on directed graphs

This section summarizes the paper '*A Measure of Similarity between Graph Vertices: Applications to Synonym Extraction and Web Searching*'[BLONDEL] of V. D. Blondel e.a. An overview of the idea and results of the paper is presented, but the proofs are developed far more extensively than they are are in te paper. The paper concludes with a detailed example about the automatic extraction of synonyms in a monolingual dictionary, but this example is left out and replaced by an own example about .... .

2.1.1 Introduction

The method of Kleinberg

The concept of similarity between directed graphs arises as a generalization of hubs and authorities introduced by Kleinberg[KLEINBERG]. Web searching engines like Google are using graphs with vertices and edges that represent the links between pages on the web. To get information out af such a graph, the idea is to identify in a set of pages relevant to a query search, the subset of pages that are good *hubs* and the subset of pages that are good *authorities*. For example, searching for 'super markets', the web-pages of Carrefour, Colruyt,... and other super markets are good authorities, whereas web-pages that point to these home-pages (like Test Aankoop, websites with recipes,...) are good hubs. A good hub points to good authorities and vice versa. The whole idea of Kleinberg was to derive an iterative method that assigns an 'authority score' and a 'hub score' to every vertex of a given graph. These scores can be obtained as the limit of a converging iterative process. We will now describe this iterative process.

Let $\mathscr{G}(V, \to)$ be a graph and let h_j and a_j be the hub and authority scores of vertex v_j. These scores must be initialized by some positive start values and then updated simultaneously for all vertices. This leads to a *mutually reinforcing relation* in which the hub score of v_j is set equal to the authority scores of all vertices pointed to by v_j and in an equal manner the authority score of v_j is set equal to the sum of the hub scores of all vertices pointing to v_j.

$$\begin{cases} h_j := \sum_{i:(v_j,v_i)\in\to} a_i, \\ a_j := \sum_{i:(v_i,v_j)\in\to} h_i. \end{cases}$$

Let B be the adjacency matrix of $\mathscr{G}$ and let h and a be the vectors of hub and authority scores. The mutually reinforcing relation can now be rewritten in:

$$\begin{pmatrix} h \\ a \end{pmatrix}_{k+1} = \begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} h \\ a \end{pmatrix}_k, k = 0, 1, \dots,$$

In compact form, we denote

$$x_{k+1} = Mx_k, k = 0, 1, \dots,$$

where

$$x_k = \begin{pmatrix} h \\ a \end{pmatrix}_k, M = \begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix}$$

Because adjacency matrices are nonnegative by definition, the matrix M is nonnegative too. M is also clearly a symmetric matrix. Now we make a sequence of normalized vectors because only the relative scores do matter[1].

$$z_0 = x_0 > 0, z_{k+1} = \frac{Mz_k}{||Mz_k||_2}, k = 0, 1, \dots \tag{2.1}$$

We are now interested in the limit of the sequence z_k and take this as the definition for the hub and authority scores, but this would be a little bit problematic. The sequence doesn't always converge in the first place. We will prove that sequences associated with symmetric, nonnegative matrices M oscillate between the limits:

$$z_{\text{even}} = \lim_{k\to\infty} z_{2k} \ \&\ z_{\text{odd}} = \lim_{k\to\infty} z_{2k+1}.$$

A second problem is that the limit vectors z_{even} and z_{odd} do in general depend on the initial vector z_0 and that there is no clear natural choice for z_0. The set of all limit vectors obtained when starting from any positive initial vector is given by:

$$Z = \{z_{\text{even}}(z_0), z_{\text{odd}}(z_0) : z_0 > 0\},$$

and we have to select one vector in Z. A good choice is the vector z_{even} obtained for $z_0 = \mathbf{1}$, because it has several nice properties that make it a good choice[2]. First, it is easy to compute and second, it has the extremal property of being the unique vector in Z with the largest possible Manhattan norm[3]. This will also be proved later in this section. Because of these properties, we take the subvectors of $z_{\text{even}}(\mathbf{1})$ as the definitions for the hub and authority

[1] $||.||_2$ is the Euclidean vector norm.

[2] $\mathbf{1}$ is a matrix, or vector, whose entries are all equal to 1.

[3] The $||.||_1$-norm or the Manhattan norm of a vector $\mathbf{x}$ is $||\mathbf{x}_1|| = \sum_{i=1}^{n} |x_i|$.

scores.

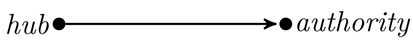
Finally, notice that the second power of the matrix M has the form:

$$M^2 = \begin{pmatrix} BB^T & 0 \\ 0 & B^TB \end{pmatrix},$$

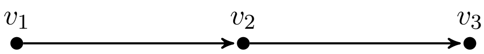
meaning that if the invariant subspaces associated with BB^T and B^TB have dimension 1, then the normalized hub and authority scores are given by the eigenvectors of BB^T and B^TB. Also notice that when the invariant subspace has dimension 1, any positive vector z_0 would give the same result as starting vector $\mathbf{1}$.

Developing a generalization of Kleinbergs method

We introduce the concept of similarity on directed graphs by generalizing the construction of the previous paragraph. The authority score of vertex v_j of a graph $\mathscr{G}$ can be thought of as a similarity score between v_j of $\mathscr{G}$ and vertex *authority* of the graph



and, conversely, the hub score of vertex v_j of $\mathscr{G}$ can be seen as a similarity score between v_j and vertex *hub*. We call the hub-authority graph a *structure graph* for this mutually reinforcing relation. Now the mutually reinforcing updating iteration used in the previous paragraph, can be generalized to graphs with other structure graphs. We start with an example. In our example, we use as structure graph a path graph with three vertices v_1, v_2, v_3.



Let $\mathscr{G}(W, \to)$ be a graph. With each vertex w_j of $\mathscr{G}$ we now associate three scores x_{i1}, x_{i2} and x_{i3}, one for each vertex of the structure graph. We initialize these score with a positive value and then update them according to the mutually reinforcing relation:

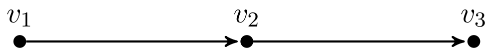
$$\begin{cases} x_{i1} := & & \sum_{j:(w_i,w_j)\in\to} x_{i2}, & \\ x_{i2} := & \sum_{j:(w_j,w_i)\in\to} x_{i1} & & + \sum_{j:(w_i,w_j)\in\to} x_{i3}, \\ x_{i3} := & & \sum_{j:(w_j,w_i)\in\to} x_{i2}, & \end{cases}$$

or, in matrix form ($\mathbf{x_j}$ denotes the column vector with entries x_{ij}),

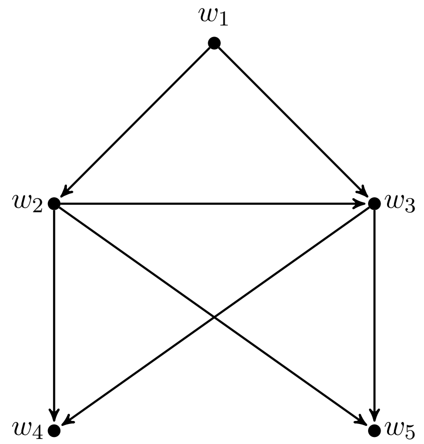
$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_{k+1} = \begin{pmatrix} 0 & B & 0 \\ B^T & 0 & B \\ 0 & B^T & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_k$$

which we again can denote by $x_{k+1} = Mx_k$. The example is now exactly the same as the previous example with hubs and authorities. The matrix M is symmetric and nonnegative, the normalized even and odd iterates converge and the limit $z_{\text{even}}(\mathbf{1})$ is, among all possible limits, the unique vector with the largest possbile Manhattan norm. The three components of the extremal limit $z_{\text{even}}(\mathbf{1})$ are now defined as the *similarity scores* s_1, s_2, s_3 and the *similarity matrix* is defined by $\mathbf{S} = [s_1\ s_2\ s_3]$. We now give a numerical example.

Example 2.1.1. Take as structure graph again the path graph with three vertices v_1, v_2, v_3:



Let $\mathscr{G}(W, \to)$ be the following graph:



Then the adjacency matrix B is:

$$B = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

By using the described mutually reinforcing updating iteration we become the following similarity matrix (a numerical algorithm to calculate this is presented later on in this section together with some proofs that facilitate the calculation):

$$S = \begin{pmatrix} 0.4433 & 0.1043 & 0 \\ 0.2801 & 0.3956 & 0.0858 \\ 0.0858 & 0.3956 & 0.2801 \\ 0.2216 & 0.0489 & 0.2216 \\ 0 & 0.1043 & 0.4433 \end{pmatrix}$$

The similarity score of w_4 with v_2 of the structure graph is equal to 0.0489.

The general case

We now construct the general case. Take two directed graphs $\mathscr{G}(U, \to)$ and $\mathscr{H}(V, \to')$ with $n_{\mathscr{G}}$ and $n_{\mathscr{H}}$ the order of the graphs. We think of $\mathscr{G}$ as the structure graph (such as the graphs hub $\to$ authority and the graph $1 \to 2 \to 3$ in the previous paragraphs). We consider the real scores x_{ij} for $i = 1, \ldots, n_{\mathscr{H}}$ and $j = 1, \ldots, n_{\mathscr{G}}$ and we get the following mutually reinforcing updating iteration with the following updating equations:

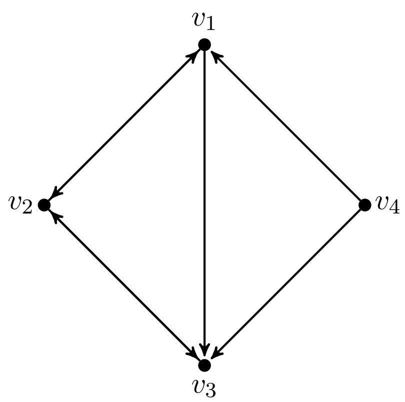
$$x_{ij} := \sum_{r:(v_r,v_i)\in\to',\, s:(u_s,u_j)\in\to} x_{rs} + \sum_{r:(v_i,v_r)\in\to',\, s:(u_j,u_s)\in\to} x_{rs}$$

Consider the product graph $\mathscr{G} \times \mathscr{H}$. The above updating equation is equivalent to replacing the scores of all vertices of the product graph by the sum of the scores of the vertices linked by an incoming or outgoing edge. Equation 2.1 can also be rewritten in a a more compact matrix form. Let X_k be the $n_{\mathscr{H}} \times n_{\mathscr{G}}$ matrix of entries x_{ij} at iteration k, and A and B are the adjacency matrices of $\mathscr{G}$ and $\mathscr{H}$. Then the updating equations can be written as:

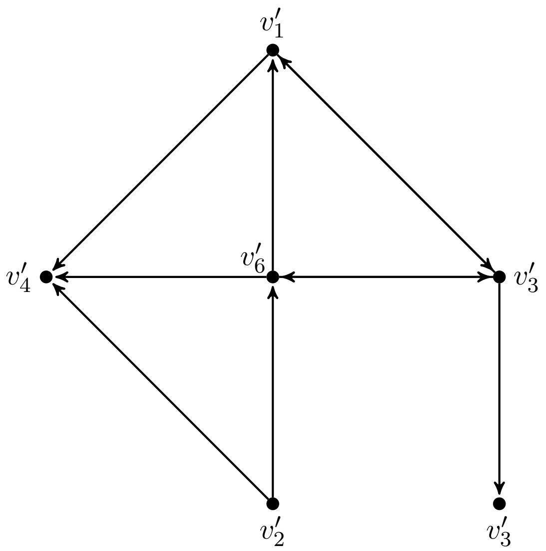
$$X_{k+1} = BX_kA^T + B^TX_kA, \qquad k = 0, 1, \ldots, \tag{2.2}$$

We'll prove that the normalized even and odd iterates of this updating equation converge and that the limit $z_{\text{even}}(\mathbf{1})$ is the limit with the largest Manhattan norm. This limit is the definition of the similarity matrix. The following example shows a calculated similarity matrix of two directed graphs.

Example 2.1.2. Let $\mathscr{G}_A(V, \to)$ be the following graph:



Let $\mathscr{G}_B(V', \to')$ be the following graph:



We become the following similarity matrix (a numerical algorithm to calculate this matrix is introduced later in this section):

$$S = \begin{pmatrix} 0.2636 & 0.2786 & 0.2723 & 0.1289 \\ 0.1286 & 0.1286 & 0.0624 & 0.1268 \\ 0.2904 & 0.3115 & 0.2825 & 0.1667 \\ 0.1540 & 0.1701 & 0.2462 & 0 \\ 0.0634 & 0.0759 & 0.1018 & 0 \\ 0.3038 & 0.3011 & 0.2532 & 0.1999 \end{pmatrix}$$

We see for example, that vertex v_2 of $\mathscr{G}_A$ is most similar to vertex v_3' in $\mathscr{G}_B$ because the similarity score s_{32} is the highest among the similarity scores in s_2.

2.1.2 Convergence of the sequence z_k

In the introduction, we depend on the result that the sequence in Equation 2.1 converges for even and odd iterates. We now investigate this hypothesis and prove a theorem about this convergence with the help of the Perron-Frobenius theory of the first chapter. Before we can start with this theorem, we first need some more corollaries of the Perron-Frobenius theorem for matrices that are not only nonnegative, but also symmetric.

Theorem 2.1.3. *Let M be a symmetric nonnegative matrix with spectral radius ρ. Then the algebraic and geometric multiplicity of the Perron root ρ are equal; there is a nonnegative matrix X whose columns span the invariant subspace associated with the Perron root; and the elements of the orthogonal projector Π on the vector space associated with the Perron root of M are all nonnegative.*

Proof. Ik heb nog enkele vragen over dit bewijs en heb de uitwerking hier daarom nog even achterwege gelaten. □

Theorem 2.1.4. *Let M be a symmetric nonnegative matrix of spectral radius ρ. Let $z_0 > 0$ and consider the sequence*

$$z_{k+1} = \frac{Mz_k}{||Mz_k||_2}, k = 0, 1, \ldots$$

Two convergence cases can occur depending on whether or not $-rho$ is an eigenvalue of M. When $-\rho$ is not an eigenvalue of M, then the sequence of z_k simply converges to $\frac{\Pi z_0}{||\Pi z_0||_2}$, where Π is the orthogonal projector on the invariant subspace associated with the Perron root ρ. When $-\rho$ is an eigenvalue of M, then the subsequences z_{2k} and z_{2k+1} converge to the limits

$$z_{even}(z_0) = \lim_{k\to\infty} z_{2k} = \frac{\Pi z_0}{||\Pi z_0||_2} \ \& \ z_{odd}(z_0) = \lim_{k\to\infty} z_{2k+1} = \frac{\Pi M z_0}{||\Pi M z_0||_2}.$$

and the vector $z_{even}(\mathbf{1})$ is the unique vector of largest possible Manhattan norm in that set.

Proof. Ik heb nog enkele vragen over dit bewijs en heb de uitwerking hier daarom nog even achterwege gelaten. □

2.1.3 Similarity matrices

2.1.4 Approximation algorithm

Mag ik hier van de complexiteit van de klassieke power methode uitgaan als voorkennis? Is het opportuun mijn Pythoncode van dit algoritme hier af te drukken of is dit beter als bijlage of beter helemaal niet?

2.1.5 Structure graphs

Zwaar stuk.

2.1.6 Self-Similarity

Gemakkelijk deel.

2.1.7 Similarity of regular graphs

Voor dit stuk moet de definitie van een reguliere directe graf nog toegevoegd worden aan het eerste deel. Voor de rest zeer gemakkelijk deel.

2.1.8 Application to...

Bibliography

[EVES] H. Eves, *Elementary Matrix Theory*, Dover Publications, 2012.

[BAPAT] . B. Bapat, T.E.. Raghavan, *Nonnegative Matrices and Applications*, Encyclopedia of Mathematics and its Applications 64, Cambridge University Press, 1997.

[STERN2010] S. Sternberg, *The Perron-Frobenius theorem*, Chapter 9 in Dynamical Systems, Dover Publications, 2010.

[NOUT2008] D. Noutsos, *Perron-Frobenius theory and some extensions* (lecture notes), Department of Mathematics, University of Ioannina, May 2008.

[VARGA] . S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, 1962.

[1] . A. Brualdi, H. J. Ryser, *Combinatorial Matrix Theory*, Cambridge University Press, 1991

[2] . Lancaster, M. Tismenetsky, *The Theory of Matrices: With Applications*, Academic Press, 1985

[CARA] P. Cara, *Discrete Wiskunde*, Vrije Universiteit Brussel - Dienst Uitgaven, 2011

[BLONDEL] V. D Blondel, A. Gajardo, M. Heymans, P. Senellart, P. Van Dooren, *A Measure of Similarity between Graph Vertices: Applications to Synonym Extraction and Web Searching*, Society for Industrial and Applied Mathematics, 2004.

[KLEINBERG] J. M. Kleinberg, *Authoritative sources in a hyperlinked environment*, J. ACM, 1999.