NonNegative Matrices and Related Topics

1. The Perron-Frobenius Theorem
2. Graphs and Matrices
3. Stability
4. Applications and Extensions

1. The Perron-Frobenius Theorem

Nonnegative matrices are the main objects of this course. I chose to talk about such matrices because they enjoy lovely algebraic, geometric and combinatorial properties and have many important applications.

A matrix A is *nonnegative*, $A \geq 0$, if all its entries are nonnegative. A matrix A is *positive*, A>0, if all its entries are positive.

There are many books and surveys on nonnegative matrices. Here is a personal choice: [Bapat and Raghavan 1997], [Berman, Neumann and Stern 1989], [Berman and Plemmons 1979,1994], [Minc 1988], [Rothblum 2006] and [Senata 1981].

For a square matrix A we denote
ρ (A) = max $\{|\lambda|;\ \lambda$ is an eigenvalue of A$\}$
and
μ (A) = max $\{|\lambda|;\ \lambda$ is an eigenvalue of A, $\lambda \neq \rho$ (A)$\}$.
ρ (A) is called the *spectral radius* of (A).

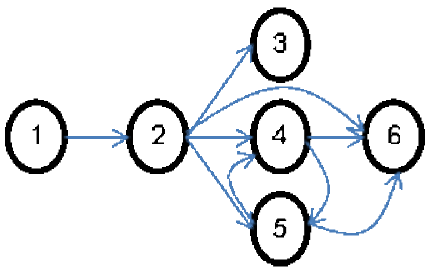
The seminal theorem on positive matrices was proved by Oscar Perron more than 100 years ago.

Theorem, [Perron 1907]
If A is a square positive matrix then
a) ρ (A) > 0,
b) ρ (A) is a simple eigenvalue of A,
c) to ρ (A) corresponds a positive eigenvector,
d) μ (A) < ρ (A),
e) $\lim_{m\to\infty}(A/(\rho(A))^m \equiv L = xy^T$, where
Ax= ρ (A)x, x>0; $A^T y = \rho(A) y$,y>0; $x^T y = 1$,
f) for every r, $\mu(A)/\rho(A) < r < 1$, there exists a constant C=C(r,A), such that for every m
$\|(A/\rho(A))^m - L\|_\infty \leq Cr^m$, where $\|A\|_\infty = \max|a_{ij}|$.

Application: The Google Matrix

When we google a concept, for example the Perron-Frobenius Theorem, we are offered a large number of possible websites. To make the search useful, the more important sites are offered first.
So, how are the sites ranked? In a similar way to the way that tennis players are ranked. A site is important if important sites point to it.
Suppose we have N pages that have to be ranked.

Example (N=6):



Let PR(i)=the rank of site i, denote the probability that a random surfer will be in site i.
Assumptions:
If k links go out from site i and if one of them goes to site j, then if the surfer follows the link he will choose j in probability 1/k.
If there are no links at all from site i, then the probability to go to j is 1/N.
The matrix P = (P_{ij}) is *stochastic*, i.e., nonnegative with row sums equal to 1.

In the example;

$$P=\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}$$

Another assumption:

The surfer uses one of the links (or artificial links) in probability α (around 0.85) or chooses an arbitrary (possibly the same) site in probability 1- α , so the final probability of moving from site i to site j is

$a_{ij} = \alpha P_{ij} + (1-\alpha)1/N.$

The new matrix A=(a_{ij}) is also stochastic. It is also positive.

In the example

$$A = \alpha P + (1-\alpha)\frac{ee^T}{N} = \begin{pmatrix} 1/40 & 7/8 & 1/40 & 1/40 & 1/40 & 1/40 \\ 1/40 & 1/40 & 19/80 & 19/80 & 19/80 & 19/80 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/40 & 1/40 & 1/40 & 9/20 & 1/40 & 9/20 \\ 1/40 & 1/40 & 1/40 & 9/20 & 1/40 & 9/20 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}$$

The matrix ee^T can be replaced by ve^T where v is a positive vector that is chosen to give preference to particular sites.

So how are the PR(i) s related?

$$\begin{aligned} PR(1) &= a_{11}PR(1) + a_{21}PR(2) + \ldots + a_{N1}PR(N) \\ PR(2) &= a_{12}PR(1) + a_{22}PR(2) + \ldots + a_{N2}PR(N) \\ &. \qquad\qquad\qquad\qquad . \\ &. \qquad\qquad\qquad\qquad . \\ &. \qquad\qquad\qquad\qquad . \\ PR(N) &= a_{1N}PR(1) + a_{2N}PR(2) + \ldots + a_{NN}PR(N) \end{aligned}$$

Denoting $x = (PR(1),\ldots PR(N))^T$ we get

$A^T x = x$

so x is an eigenvector of A^T .

Indeed 1 is the spectral radius of A^T , and by Perron's Theorem, this eigenvector is uniquely defined when the PR(i)s are positive numbers that sum to 1.

The Perron vector is computed by the power method, starting with a probability vector x and computing the limit of A^k x, which is given by part(e) of Perron's theorem. Using part(f) of the theorem one can show that the probability α determines the rate of convergence. For smaller α we have faster convergence. On the other hand for larger α we have better use of the hyperlink structure of the web so there is a tradeoff between the two.

Perron's Theorem can be generalized to primitive and to irreducible matrices.

Primitivity

A nonnegative matrix is *primitive* if for some natural number k, A^k is positive.
In Perron's Theorem, "positive" can be replaced by "primitive".

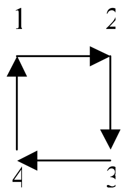
Irreducibility

To an nxn matrix A, corresponds a directed graph (digraph), D(A), with n vertices 1,2,...,n, and an arc from i to j iff $a_{ij} \neq 0$. Digraphs are surveyed in [Berman and Shaked-Monderer 2008].

Examples (x denotes a nonzero entry):

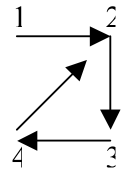
$$A = \begin{pmatrix} 0 & x & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & x \\ x & 0 & 0 & 0 \end{pmatrix}$$

D(A):



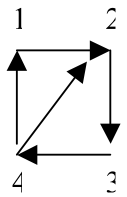
$$B = \begin{pmatrix} 0 & x & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & x \\ 0 & x & 0 & 0 \end{pmatrix}$$

D(B):



$$C = \begin{pmatrix} 0 & x & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & x \\ x & x & 0 & 0 \end{pmatrix}$$

D(C):

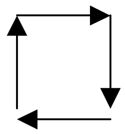


A directed graph D is *strongly connected* if for any two vertices i and j of D, D contains a path from i to j.
A square matrix A is *irreducible* if D(A) is strongly connected. Otherwise A is *reducible.*
In other words, an nxn, n>1, square matrix A is reducible iff it is permutationally similar to

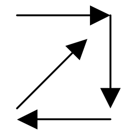
$$PAP^T = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

where B and D are square. If n=1, A is reducible iff it is a (1x1) zero matrix.

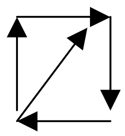
Examples:

$$A = \begin{pmatrix} 0 & x & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & x \\ x & 0 & 0 & 0 \end{pmatrix}$$ D(A): 

A is irreducible.

$$B = \begin{pmatrix} 0 & x & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & x \\ 0 & x & 0 & 0 \end{pmatrix}$$ D(B): 

B is reducible.

$$C = \begin{pmatrix} 0 & x & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & x \\ x & x & 0 & 0 \end{pmatrix}$$ D(C): 

C is irreducible.

Perron's Theorem was extended by Ferdinand Georg Frobenius. The resulting fundamental theorem is known as the Perron-Frobenius Theorem. We divide it into two parts.

The Perron-Frobenius Theorem [Frobenius 1912] - Part I
In the first three statements of Perron's Theorem, "positive" can be replaced by "nonnegative irreducible", namely:
If $A \geq 0$ is irreducible then

a) ρ (A) > 0,
b) ρ (A) is a simple eigenvalue of A,
c) to ρ (A) corresponds a positive eigenvector.

In addition, if $A \geq 0$ is irreducible (which of course includes A>0):

d) if Ax=λ x and x is positive, then $\lambda = \rho$ (A),
e) if B $\geq$ A, B $\neq$ A, then ρ (B) > ρ (A),
f) if B $\leq$ A, B $\neq$ A, then ρ (B) < ρ (A).

Corollary
If B$\neq$A is a principal submatrix of an irreducible nonnegative matrix A, then ρ(B) < ρ(A).

Another corollary
If A is a square nonnegative matrix,
$$\min_i \sum_{j=1}^{n} a_{ij} \le \rho(A) \le \max_i \sum_{j=1}^{n} a_{ij}$$
To introduce the second part of the PFT we need the definition of order of cyclicity.

The *order of cyclicity* of an irreducible nonnegative matrix is the number of its eigenvalues whose modulus is the spectral radius.

An irreducible nonnegative matrix is primitive iff its index of cyclicity is 1.

The Perron-Frobenius Theorem, Part II
If A $\geq$ 0 is irreducible and its index of cyclicity is k>1, then
a) the eigenvalues of A of modulus ρ (A) are $\rho(A)e^{2\pi i/k}$; i=0,1,...,k-1,
b) rotating the complex plane by $2\pi / k$ takes the spectrum of A onto itself,
c) A is permutationally similar to

$$PAP^T = \begin{pmatrix} 0 & A_{12} & 0 & . & . & 0 \\ 0 & 0 & A_{23} & . & . & . \\ . & . & . & . & . & . \\ . & & . & . & . & 0 \\ . & & & . & . & A_{k-1k} \\ A_{k1} & . & . & . & 0 & 0 \end{pmatrix}$$

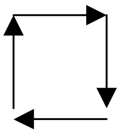
where the zero blocks on the diagonal are square.

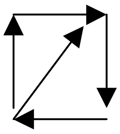
The Perron-Frobenius Theorem has many proofs. Here we mention only the classical proof given by Wielandt, [Wielandt 1950].

We conclude this subsection with a lovely way suggested by Romanovsky for computing the order of cyclicity.

Theorem [Romanovsky 1936]
Let A$\geq$0 be irreducible and let k_i be the greatest common divider of the lengths of cycles in D(A) that pass through i. Then all the k_i's are equal and are equal to the order of cyclicity of A.

Examples:

D(A):  $k_1 = k_2 = k_3 = k_4 = 4$

D(C):  $k_1 = k_2 = k_3 = k_4 = 1$

Reducible Matrices

If $A \geq 0$ is reducible (of order > 1) then there exists a permutation matrix P, so that A can be reduced to the form

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

where B and D are square matrices. If either B or D are reducible (of order > 1) they can be reduced in a similar manner so finally, by a suitable permutation, A can be reduced to a block triangular form*, the Frobenius normal form*

$$\begin{pmatrix} A_{11} & A_{12} & . & . & . & A_{1s} \\ 0 & A_{22} & . & & & . \\ . & . & & . & & . \\ . & & . & & . & . \\ . & & & . & & A_{s-1s} \\ 0 & . & . & . & 0 & A_{ss} \end{pmatrix}$$

where each block on the diagonal is square and irreducible or a (reducible) 1x1 zero matrix.
We will study the spectral properties of reducible matrices in greater depth in the end of this section.

Inverse Eigenvalue Problems

Question: Find a 4x4 nonnegative matrix with eigenvalues
$\sqrt{2}$, $\sqrt{2}$, 1, 1.

Answer :

$$\begin{pmatrix} \sqrt{2} & & & \\ & \sqrt{2} & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

Question: Find a 4x4 nonnegative matrix with eigenvalues
$\sqrt{2}$, $\sqrt{2}$, 1, -1.

Answer :

$$\begin{pmatrix} \sqrt{2} & & & \\ & \sqrt{2} & & \\ & & 0 & 1 \\ & & 1 & 0 \end{pmatrix}$$

Question: Find a 4x4 symmetric nonnegative matrix with eigenvalues $\sqrt{2}$, $\sqrt{2}$, i, -i.

Answer :
There is no such matrix since the eigenvalues of a real symmetric matrix are real.

Question: Find a 4x4 nonnegative matrix with eigenvalues $\sqrt{2}$, $\sqrt{2}$, i, -i.

The Nonnegative Inverse Eigenvalue Problem (NIEP)
Given n complex numbers $\lambda_1, \lambda_2, ..., \lambda_n$ and $|\lambda_1| \geq |\lambda_2| \geq ... \geq |\lambda_n|$,
is there an nxn nonnegative matrix A whose eigenvalues are $\lambda_1, \lambda_2, ..., \lambda_n$?

The Symmetric Nonnegative Inverse Eigenvalue Problem (SNIEP)
Given n real numbers
$\lambda_1, \lambda_2, ..., \lambda_n \quad |\lambda_1| \geq |\lambda_2| \geq ... \geq |\lambda_n|$,
is there an nxn symmetric nonnegative matrix A whose eigenvalues are $\lambda_1, \lambda_2, ..., \lambda_n$?

Necessary conditions for NIEP:
$|\lambda_1| = \lambda_1$,

$\{ \lambda_1, \lambda_2, ..., \lambda_n \} = \{ \bar{\lambda}_1, \bar{\lambda}_2, ..., \bar{\lambda}_n \}$

$S_k \geq 0$, k=1,2,...; where $S_k = \sum_{i=1}^{n} \lambda_i^k$,

$(S_k)^m \leq (n^{m-1}) S_{km}$;k,m=1,2,... .

([Loewy and London 1978], [Johnson 1981]).

Question: Is there a 4x4 nonnegative matrix with eigenvalues $\sqrt{2}$, $\sqrt{2}$, i, -i ?

Answer: No!
If there was such a matrix A it would be reducible, so for some permutation P

$$PAP^T = \begin{pmatrix} \sqrt{2} & x & x & x \\ 0 & & & \\ 0 & & & \\ 0 & & & B \end{pmatrix}$$

where B is a 3x3 nonnegative matrix with eigenvalues $\sqrt{2}$, i, -i but

$$(S_k)^m \le (n^{m-1})S_{km}$$

does not hold for k=1, m=2, n=3 as $\sqrt{2}^2$ > 3x0.

For n$\le$4, NIEP and SNIEP are solved. Also, for n$\le$4, a real solution of NIEP also solves SNIEP ([Johnson, Laffey and Loewy 1996]); However, for n>4 let S_t={3+t, 3, -2,-2,-2}, t>0.
The smallest value of t for which S_t solves SNIEP is t=1 (Hartwig and Loewy, see [Loewy and McDonald 2004]) but there exists 0<t<1 such that S_t solves NIEP (Meehan 1998).

The Boyle-Handelman Theorem ([Boyle and Handelman 1991])
Let $\lambda_1, \lambda_2, ..., \lambda_n$ be nonzero numbers such that $\lambda_1 > \max(2 \le i \le n, |\lambda_i|)$.
Then, for some N, $\lambda_1, \lambda_2, ..., \lambda_n$ are the eigenvalues of an (n+N)x(n+N) primitive matrix, iff
$\lambda_1 = |\lambda_1|$
$S_k \ge 0$; k=1,2,... (where $S_k = \sum_{i=1}^{n} \lambda_u^k$)
S_k >0=> S_{jk} >0; j,k=1,2,...

Example:
We saw that $\sqrt{2}$, i, -i are not the eigenvalues of a 3x3 nonnegative matrix. However, for every positive epsilon, $\sqrt{2+\varepsilon}$, i,-i are the nonzero eigenvalues of a (3+N)x(3+N) primitive matrix. However when $\varepsilon \to 0$, $N \to \infty$.

Suleimanova's type results

Let λ_1 >0 and $\lambda_i \leq 0$; i=2,...,n.
[Suleimanova 1949] showed that $\lambda_1, \lambda_2, ..., \lambda_n$ solve NIEP iff $S_1 \geq 0$.

The *companion matrix* C_f of the polynomial
$f(x) = x^n + a_1 x^{n-1} + ... + a_{n-1}x + a_n$
is

$$C_f = \begin{pmatrix} & & & & & -a_n \\ 1 & & & & & \\ & 1 & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & -a_2 \\ & & & & 1 & -a_1 \end{pmatrix}$$

[Friedland 1978] showed that in Suleimanova's result the nonnegative matrix can be chosen to be a companion matrix.

[Borobia, Moro and Soto 2004] showed that if $\lambda_1, \lambda_2, ..., \lambda_n$ satisfy
1) $\lambda_1 > 0$,
2) $\operatorname{Re}\lambda_i \leq 0$, i=2,…,n,
3){ $\lambda_1, \lambda_2, ..., \lambda_n$ } is closed under complex conjugation and
4) $|\operatorname{Re}\lambda_i| \geq |\operatorname{Im}\lambda_i|$, i=2,…,n,
then they solve NIEP iff $S_1 \geq 0$.
[Smigoc 2004} improved the result by replacing condition (4) by
$\sqrt{3}|\operatorname{Re}\lambda_i| \geq |\operatorname{Im}\lambda_i|$, i=2,…,n, and showed that her result is the best possible of this type.

Theorem ([Laffey and Smigoc 2006])
Let $\lambda_1, \lambda_2, ..., \lambda_n$ be nonzero complex numbers such that λ_1>0, $\operatorname{Re}(\lambda_i) \leq 0$; i=2,...,n. Then $\lambda_1, \lambda_2, ..., \lambda_n$ are the nonzero eigenvalues of an (n+N)x(n+N) nonnegative matrix if $\{\lambda_1, \lambda_2, ..., \lambda_n\}$ is closed under complex conjugation, $S_1 \geq 0$ and $S_2 > 0$.
The nonnegative matrix can be chosen to be of the form C+tI, where C is a companion matrix and t is a nonnegative number.

Unlike the Boyle-Handelman Theorem where N is not bounded, here the smallest nonnegative N needed is the smallest N that satisfies
$S_1^2 \le (n+N)S_2$.

Corollary ([Laffey and Smigoc 2006])
The numbers $\lambda_1, \lambda_2, ..., \lambda_n$, as in the theorem, solve NIEP iff $\{\lambda_1, \lambda_2, ..., \lambda_n\}$ is closed under complex conjugation, $S_1 \ge 0$, $S_2 \ge 0$ and $S_1^2 \le nS_2$.

Index of primitivity

Recall that a nonnegative matrix is primitive if for some natural number k, A^k is positive. Since this is equivalent to A being irreducible with order of cyclicity 1 it follows that A^l is also positive for all $l \ge k$.
The *index of primitivity* of a primitive matrix A is the smallest number $\gamma(A)$ such that $A^{\gamma(A)}$ is positive. The index for a primitive graph is defined in a similar way.

We now state some results on upper bounds for the index of primitivity.

[Wielandt 1950] proved that for an $n \times n$ primitive matrix $\gamma(A) \le n^2 - 2n + 2$ and showed that

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

is primitive with index primitivity $n^2 - 2n + 2$.
The upper bound can be reduced if more information on A is known.

Theorem
Let $A \ge 0$ be an $n \times n$ primitive matrix and suppose that for some natural number h, $A + A^2 + ... + A^h$ has at least d positive diagonal entries, then $\gamma(A) \le n - d + h(n-1)$.

A matrix A is combinatorially symmetric if a_{ij} is nonzero iff a_{ji} is.

Corollary

If A is an nxn combinatorial symmetric primitive matrix then its index of primitivity is less than or equal to 2(n-1).

Stochastic Matrices

Stochastic matrices are the matrices that appear in the study of Markov Chains.
A square matrix is *(row) stochastic* if it is nonnegative and if its row sums are equal to 1. A square matrix is *column stochastic* if it is nonnegative and if its column sums are equal to 1. A matrix is *doubly stochastic* if it is row stochastic and column stochastic.

Theorem
The maximal eigenvalue of a stochastic matrix is 1. A nonnegative matrix A is stochastic iff e, the vector of ones, is an eigenvector of A corresponding to 1.

In the next subsection we will characterize those nonnegative matrices that posses a positive eigenvector that corresponds to the spectral radius. This class of matrices contains, of course, the irreducible matrices and by the last theorem also the stochastic matrices. In fact there is a similarity connection between stochastic matrices and this class and this connection is described in the next theorem.

Theorem.
If A≥0, $\rho(A) > 0$ and $Az = \rho(A)z$, then $A / \rho(A)$ is similar to a stochastic matrix.

The similarity matrix can be a diagonal matrix whose diagonal entries are the elements of z.

We conclude this subsection with three remarks on doubly stochastic matrices.

An *ordered vector* $x = (x_i)$; $x_1 \geq x_2 \geq ... \geq x_n$ *majorizes* an ordered vector $y = (y_i)$; $y_1 \geq y_2 \geq ... \geq y_n$ if $\sum_{i=1}^{k} x_i \geq \sum_{i=1}^{k} y_i$; $k = 1,..., n-1$; $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$
Theorem ([Hardy, Littlewood and Polya 1929])

An ordered vector x majorizes an ordered vector y iff for some doubly stochastic matrix A, y=Ax.

The second remark is a theorem of Birkhoff ([Birkhoff 1946])

Theorem
The set of all $n \times n$ doubly stochastic matrices is a convex polyhedron whose vertices are the permutation matrices.

The third remark is a classical conjecture of van der Waerden from 1926, more than fifty years later independently by Egorichev and Falikman .

Recall that the permanent of an nxn matrix A is a function similar to the determinant but without the minus sign.

$$Per(A) = \sum_{\sigma \in S_n} \overset{n}{\underset{i=1}{\pi}} a_{i\sigma(i)}$$

where S_n is the symmetric group of order n.

Theorem

If A is an $n \times n$ doubly stochastic matrix then

$$Per(A) \geq \frac{n!}{n^n}$$

and the minimum is obtained only for the matrix that all of its entries are equal (to 1/n).

The following example demonstrates what the permanent means:

Example:

Particles q_i are located on the vertices of a complete graph with n vertices c_i. If p_{ij} is the probability that at button push particle q_i moves from vertex c_i to vertex c_j and $P = (p_{ij})$ then the probability that after the button push there is precisely one particle at each vertex is per(P).

More on Reducible Matrices

To study the spectral properties of reducible matrices in greater depth we return to their directed graphs.

Let A be an $n \times n$ nonnegative matrix. Motivated by applications to Markov chains we say that in D(A) , vertex i *has an access* to vertex j if there is a path from i to j, and that i and j *communicate* if i has an access to j and j has an access to i. communication is an equivalence relation and we refer to the equivalence classes as the classes of A. A class α has an access to class β if for some $i \in \alpha$ and $j \in \beta$, i has an access to j.

A class is *final* if it has access to no other class. A class α is *basic* if $\rho(A[\alpha]) = \rho(A)$ where $A[\alpha]$ is the principal submatrix of A based on the indices in α and non basic if $\rho(A[\alpha]) < \rho(A)$.

The diagonal blocks in the Frobenius normal form of a matrix correspond to the classes of the matrix.

Example:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix};$$

$$\begin{tikzcd} \{1,2\} \arrow[r] \arrow[d] \arrow[dr] & \{3,4\} \\ \{5,6\} \arrow[r] & \{7\} \end{tikzcd}$$

Here {7} and {3,4} are final classes and {1,2} and {7} are basic classes.
A matrix A is irreducible iff it has only one (basic and final) class. Such a matrix has a positive eigenvector that corresponds to the spectral radius.
Stochastic matrices also have this property. An interesting question is: what are the matrices that have such a positive eigenvector?

Theorem
To the spectral radius of $A \geq 0$ there corresponds a positive eigenvector iff the final classes of A are exactly its basic ones.

Theorem
To the spectral radius of $A \geq 0$ a positive eigenvector of A and a positive eigenvector of A^T iff all the classes of A are basic and final.

Corollary
A is irreducible iff $\rho(A)$ is simple and positive vectors correspond to $\rho(A)$ both for A and for A^T.

For more information on reducible matrices the readers are referred to [Rothblum, 2006] and the references there.

2. Graphs and Matrices

Several matrices can be associated with a graph G. Similar definitions hold for digraphs (see Berman and Shaked-Monderer [2008]).

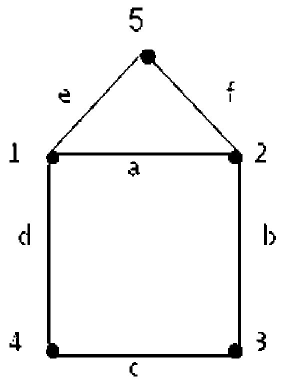
The *adjacency matrix* N(G) defined by

$$N(G) = \begin{cases} 1 & if \quad i \neq j \quad and \quad i \quad and \quad j \quad are \quad neighbors \\ 0 & otherwise \end{cases}$$

The *incidence matrix* C(G), where the rows are the vertices and the columns are the edges and $a_{ij} = 1$ if edge j contains vertex i, and 0 otherwise,
The *degrees matrix* D(G) – a diagonal matrix where d_{ii} is the degree of vertex i and the *Laplacian* L(G) = D(G)-N(G).

The matrices are of course related, $A = CC^T - D$ and $L = D - A = 2D - CC^T$.

Example



$$N(G) = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

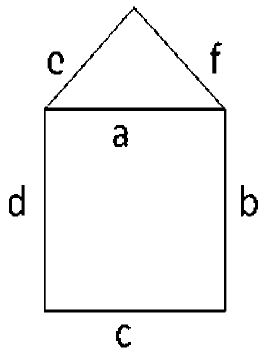
$$C(G) = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \qquad D(G) = \begin{pmatrix} 3 & & & & \\ & 3 & & & \\ & & 2 & & \\ & & & 2 & \\ & & & & 2 \end{pmatrix}$$

$$L(G)=\begin{pmatrix} 3 & -1 & 0 & -1 & -1 \\ -1 & 3 & -1 & 0 & -1 \\ 0 & -1 & 2 & -1 & 0 \\ -1 & 0 & -1 & 2 & 0 \\ -1 & -1 & 0 & 0 & 2 \end{pmatrix} = 2D(G)-C(G)C(G)^T \quad .$$

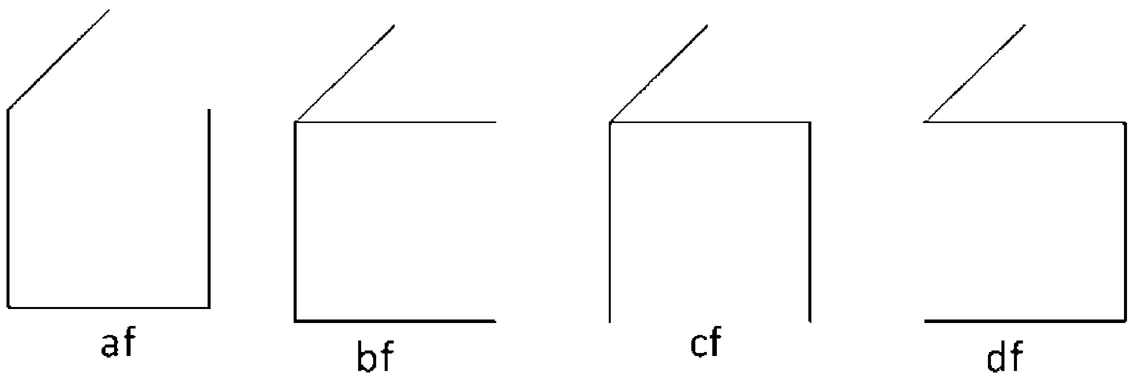
The Laplacian matrix is positive semi definite and singular. Its second smallest eigenvalue is called the *algebraic connectivity* of G, as it is positive iff G is connected. Bounds on the algebraic connectivity are related to bounds on $\mu(A)$ of a nonnegative matrix A and are studied in computer science in connection with expanders.

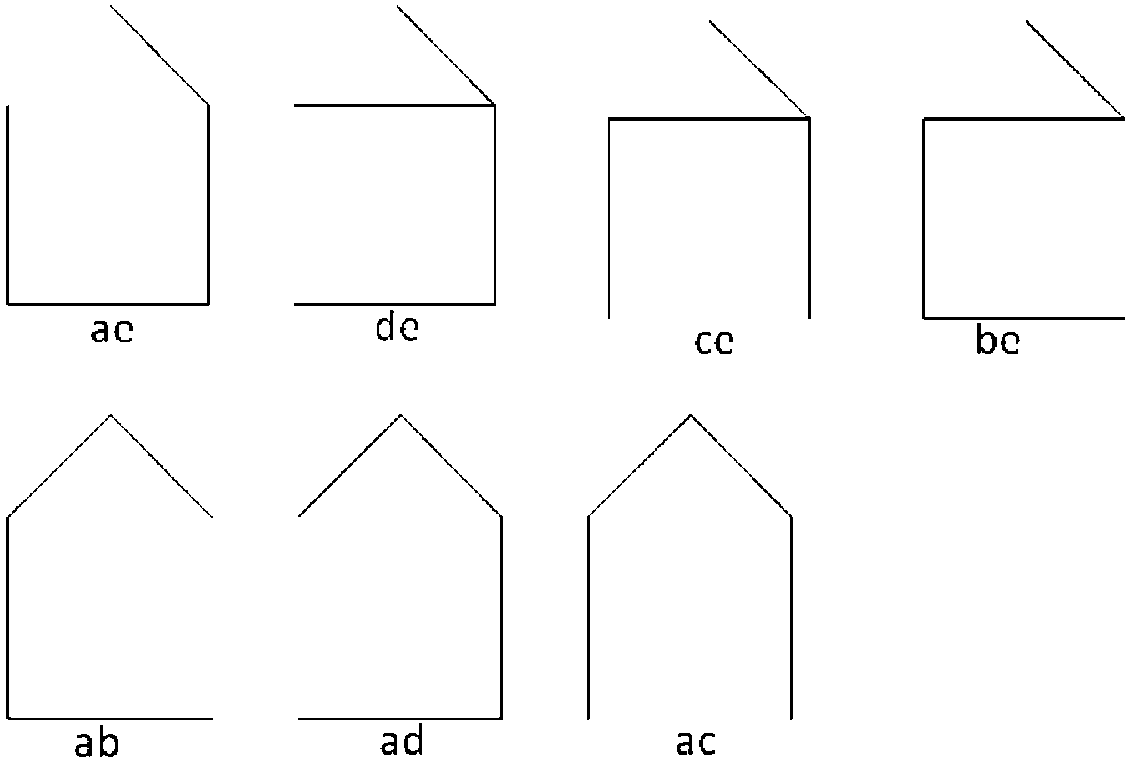
Theorem
Adj L = kJ where J is a matrix of ones and k is the number of spanning trees of G.

Example



Here there are 11 spanning trees.
The letters denote the deleted edges.

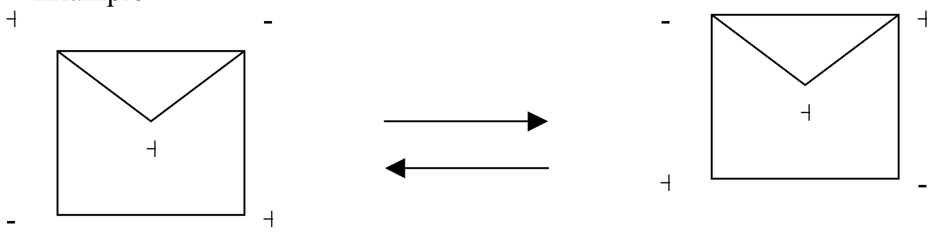




We now return to the adjacency matrix.
Here are two simple examples.

Neural Networks
At each vertex of a graph G there is a light bulb.
Some of the lights are lit. Some are off. Their status changes according to the following majority rule:
If at time t, a bulb has more neighbors that are on, it will be on at time t+1. If at time t, it has more neighbors that are off, it will be off at time t+1. In a case of a tie, there is no change.

Example



Neural Networks Theorem

For every graph and for every initial states, there is T such that for all $t \geq T$, the states of the lights at time t+2 are the same as the states at time t.

Proof.
Let N be the adjacency matrix of G and let A = (1/2)I + N.

Let $x(t)_i = \begin{cases} 1 & if \quad bulb \quad i \quad is \quad on \\ -1 & if \quad bulb \quad i \quad is \quad off \end{cases}$

We have to show that for $t \geq T$, x(t+2)=x(t).
Ax(t) and x(t+1) have the same signs.
Let $f(t) = x(t+1)^T Ax(t) = ! \max_{y_i \in \{\pm 1\}} y^T Ax(t) = x(t)^T Ax(t+1)$
$f(t+1) = x(t+2)^T Ax(t+1) = ! \max_{y_i \in \{\pm 1\}} y^T Ax(t+1)$ so $f(t+1) \geq f(t)$

f can attain only a finite number of values so from some T, f(t+1)=f(t) and thus, x(t+2)=x(t).

The "all lights on" problem.

At each vertex of a graph there is a light bulb and a knob.
Pressing a knob activates it (changes the state of the corresponding bulb) and also activates its neighbors.

Prove: Given any graph with all lights off it is possible to simultaneously lit all the bulbs.

Proof.
Here let A=I+N.
Consider the columns as representing the knobs pressed and the rows as representing the rows activated.
We have to show that Ax=e is solvable over Z_2. Suppose it is not. Then some rows of (A/e) sum to (0...0/1), so the number of these rows is odd. Thus A has a submatrix based on an odd number of rows and on all the columns such that the number of ones in each column is odd.
The same is true for the principal submatrix based on these rows and on the columns with the same indices.
Contradiction!

Completely Positive Matrices and Graphs

This section is based on [Berman & Shaked-Monderer 2003]

Question: Given n vectors $x_1, x_2, ..., \ x_n$ in a (m- dimensional) vector space V, can they be imbedded in a nonnegative orthant? In other words, is there a natural number k and an isometry T such that $Tx_1, Tx_2, ..., \ Tx_n \in R_+^k$?

Answer:
They can iff the Gram matrix $A = (< x_i, x_j >)$ is a Gram matrix of nonnegative vectors.
An $n \times n$ matrix A is completely positive (CP) iff there exists an $n \times k$ (not necessarily square) nonnegative matrix B such that:
$A = BB^T$, or equivalently
$A = b_1 b_1^T + b_2 b_2^T + ... + b_k b_k^T ; \ \ b_i \in R_+^n ; i = 1, ...k$

Example:
If a>0 and det $A \geq 0$, then

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \sqrt{a} & 0 \\ \frac{b}{\sqrt{a}} & \sqrt{\frac{\det A}{a}} \end{pmatrix} \left(\quad \right)^T = \begin{pmatrix} \sqrt{a} \\ \frac{b}{\sqrt{a}} \end{pmatrix} \left(\quad \right)^T + \begin{pmatrix} 0 \\ \sqrt{\frac{\det A}{a}} \end{pmatrix} \left(\quad \right)^T$$

Applications:
Block designs, [Hall 1958, 1967, 1986],
Modeling energy demand, [Gray and Wilson, 1980],
Markovian model for DNA evolution, [Kelly, 1994],
Probability[Diaconis, 1994],
Clustering, [Linial and Samorodinsty, 1998],
Image processing, [Li, Kummert and Frommer, 2004],

$A = A^T$ is *copositive* if
$x \geq 0 \Rightarrow x^T A x \geq 0$

Applications:
The linear complementarity problem, quadratic optimal control.

The dual cone S^* of a set S is an inner product vector space V is the set
$\{ v \in V ; < v, s > \geq 0, \forall s \in S \}$

Theorem, ([Hall and Newman, 1993])

The $n \times n$ copositive matrices and the $n \times n$ completely positive matrices are closed convex cones in the space of $n \times n$ real symmetric matrices, and each is the dual of the other with respect to the inner product <A,B>=trace AB.

Questions:
Given a matrix A is it CP?
Given an $n \times n$ CP matrix what is the smallest k such that $A = BB^T$, $B \geq 0 \quad B \in R^{n \times k}$?

Here we will briefly discuss the first question.

Necessary conditions:
$A \in CP \quad \Rightarrow$ A is positive definite (PSD) and element wise nonnegative. We refer to having the two properties as being doubly nonnegative (DNN).

Sufficient Conditions:

Theorem (Kaykobad, 1987)
If $A = A^T$ is *diagonally dominant* $(a_{ii} \geq \sum_{j=1} a_{ij}, \quad \forall i)$ then A is CP.

Example:

$$\begin{pmatrix} 5 & 2 & 2 \\ 2 & 4 & 1 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \sqrt{2} & \sqrt{2} & 0 \\ 0 & \sqrt{2} & 0 & 1 \\ 0 & 0 & \sqrt{2} & 1 \end{pmatrix} \begin{pmatrix} & & & \\ & & & \\ & & & \end{pmatrix}^T$$

The *comparison matrix* of A, M(A), is defined by
$M(A)_{ii} = |a_{ii}|, \quad M(A)_{ij} \underset{i \neq j}{=} -|a_{ij}|$

Theorem (Drew, Johnson and Loewy , 1996)

$A = A^T \geq 0, \quad M(A) \quad is \quad PSD \Rightarrow$ A is CP
This follows from the fact that there exists a positive diagonal matrix D such that DAD is diagonally dominant.

Combining the necessary of the sufficient conditions we have for a nonnegative of symmetric matrix A:
$M(A) \quad is \quad PSD \Rightarrow$ A is CP $\Rightarrow A \quad is \quad PSD$

The sufficient condition is not necessary.

Example:

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

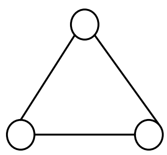
The necessary condition is not sufficient.
Example:

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 & 3 \end{pmatrix}$$

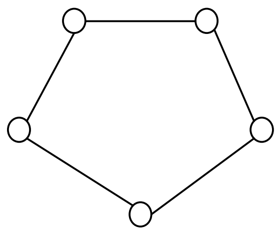
Theorem (Maxfield and Minc 1962)
For $n \le 4$ ($A = A^T \in R_+^{n\times n}$), CP $\Leftrightarrow$ DNN
with an $n \times n$ symmetric matrix A we associate a graph G(A):
V(G(A))={1,2,…,n};
(i,j) $\in$ E(G(A)) iff i$\neq$j and $a_{ij} \neq 0$.
A is a *matrix realization* of G(A)).

Examples:

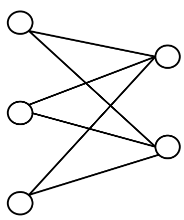
$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$, G(A):

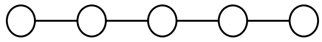


$$A = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 & 3 \end{pmatrix}$$, G(A):



$$A = \begin{pmatrix} 2 & 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 \\ 1 & 1 & 1 & 3 & 0 \\ 1 & 1 & 1 & 0 & 3 \end{pmatrix}$$, G(A):



$$A=\begin{pmatrix}2&1&0&0&0\\1&2&1&0&0\\0&1&2&1&0\\0&0&1&2&1\\0&0&0&1&2\end{pmatrix}$$, G(A): 

When is the sufficient condition necessary?

Theorem (Drew, Johnson and Loewy, 1994)
If A is CP and G(A) does not contain a triangle then M(A) is PSD.

When is the necessary condition sufficient?

A graph G is *CP* if for every symmetric nonnegative matrix realization A of G,
A is CP iff it is PSD.

Examples
Small graphs ($n \le 4$), trees, forests, (Berman and Hershkowitz, 1987)

Theorem (Berman and Hershkowitz, 1987)
If G contains an odd cycle of length greater than 4, then it is not CP.

Here we will show the proof for the case that G is an odd cycle of length k, k=5,7,….
The result for a general graph follows by continuity since the cone of CP matrices is closed.
So let B be the $k\times(k-1)$ matrix

$$B=\begin{pmatrix}1&0&.&.&.&0\\1&1&.&&&.\\0&&&.&&.\\.&.&&&.&.\\.&.&.&&&\\.&&.&.&&0\\0&.&.&0&1&1\\1&-1&.&.&\pm1&0\end{pmatrix}$$

Then $A = BB^T = \begin{pmatrix} 1 & 1 & 0 & . & . & . & 0 & 1 \\ 1 & 2 & . & & & & & 0 \\ 0 & . & . & & & & & . \\ . & . & . & . & & & & . \\ . & & . & . & . & & & . \\ 1 & & & . & . & . & & 0 \\ 0 & . & . & . & 0 & 1 & 2 & 1 \\ 1 & . & . & . & . & . & 1 & k-2 \end{pmatrix}$

G(A) is a k-cycle. A is PSD and for odd k it is also nonnegative.

$$M(A) = \begin{pmatrix} 1 & -1 & 0 & . & . & . & 0 & -1 \\ -1 & 2 & . & & & & & 0 \\ 0 & . & . & & & & & . \\ . & . & . & . & & & & . \\ . & & . & . & . & & & . \\ . & & & . & . & . & & 0 \\ 0 & . & . & . & . & . & 2 & -1 \\ -1 & 0 & . & . & . & 0 & -1 & k-2 \end{pmatrix}$$

and det M(A)= k-2-1-2-(k-1)=-4, so for k>3, it follows from the no-triangle theorem that A is DNN but not CP.

Theorem (Berman and Grone, 1988)
Bipartite graphs are CP.

This result follows from Sylvester's law of Inertia for real symmetric matrices since

$$\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\begin{pmatrix} D_1 & C \\ C^T & D_2 \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} = \begin{pmatrix} D_1 & -C \\ -C^T & D_2 \end{pmatrix}$$

In the module of stability we will mention the Hermitian version of Sylvester's Theorem.

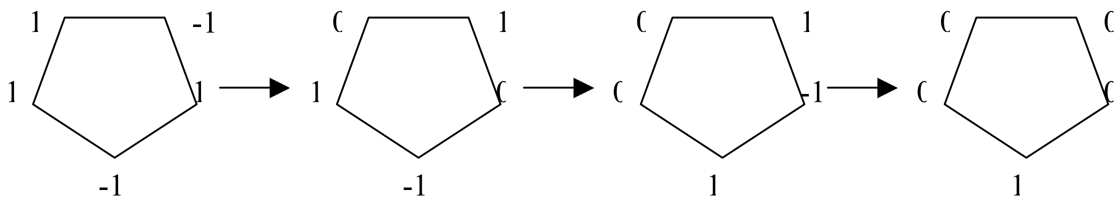
Finally, the complete characterization of CP graphs is:

Theorem (Kogan and Berman, 1988, 1993)
A graph G is CP iff it does not contain an odd cycle of length greater than 4.

A Game of Numbers

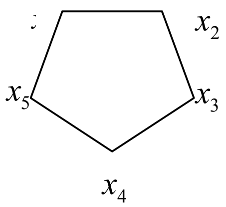
The following problem was given in the International Olympiad in Mathematics in Poland in 1986:
In every vertex of a pentagon there is an integer. The sum of the five numbers is positive. If one of them is negative, the player can choose one of the negative numbers, add it to its neighbors and multiply it by -1. This is continued as long as there is a negative number. Prove that the game must terminate.

Example:



Here is a solution:
With a state

$(x_1 x_2 x_3 x_4 x_5)$



We associate the number
f(x) $= f(x) = (x_1 - x_3)^2 + (x_2 - x_4)^2 + (x_3 - x_5)^2 + (x_4 - x_1)^2 + (x_5 - x_2)^2$.
If, for example, $x_1 < 0$ and we choose the corresponding vertex then f (x_{new}) – f(x) = $2x_1(x_1 + x_2 + x_3 + x_4 + x_5) < 0$ and since f is a sum of squares, the process must terminate.

[Shahar Mozes 1990] extended the problem in the following way:
At each vertex of a graph (non directed, connected and simple) there is a real number (not necessarily integer).
The sum of the numbers is not necessarily positive.
If some of the numbers are negative, such a number is chosen, added to its neighbors and multiplied by -1. The problem is treated matrix- theoretically in [Eriksson 1992].

Question: What are the initial states for which the process will terminate?

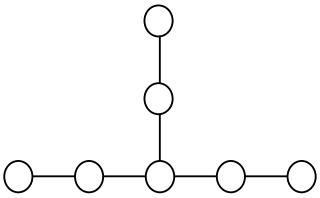
There are 3 possibilities.

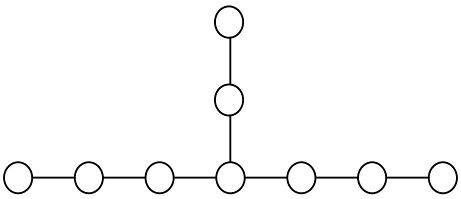
The game converges, i.e., ends in a finite number of states,
The game loops, i.e., it is periodical and thus does not end,
The game does not end and no state is repeated.

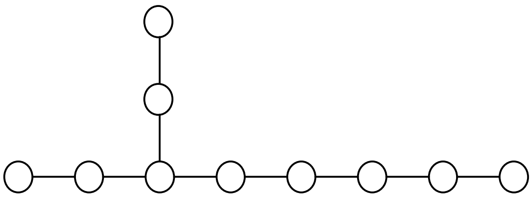
We say that a graph is a *looper* if there is an initial state from which it loops.

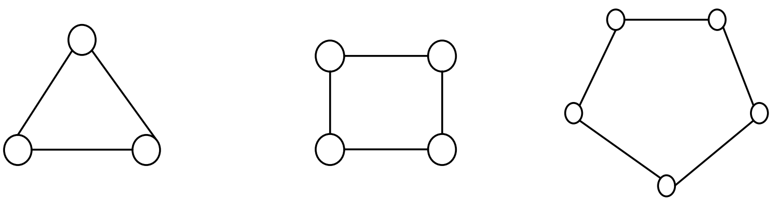
Theorem

a) If a game converges, the final state and the number of steps do not depend on the choice of vertices.
b) If $\rho(N(G))<2$, every game converges.
c) G is a looper iff $\rho(N((G))=2$.
d) If $\rho(N(G))>2$, G is not a looper and there are initial states from which the game never loops.
e) The loopers are exactly the following graphs:

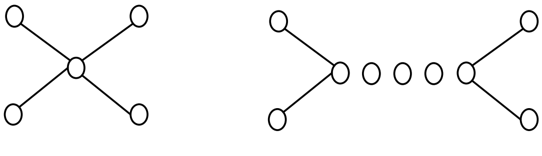






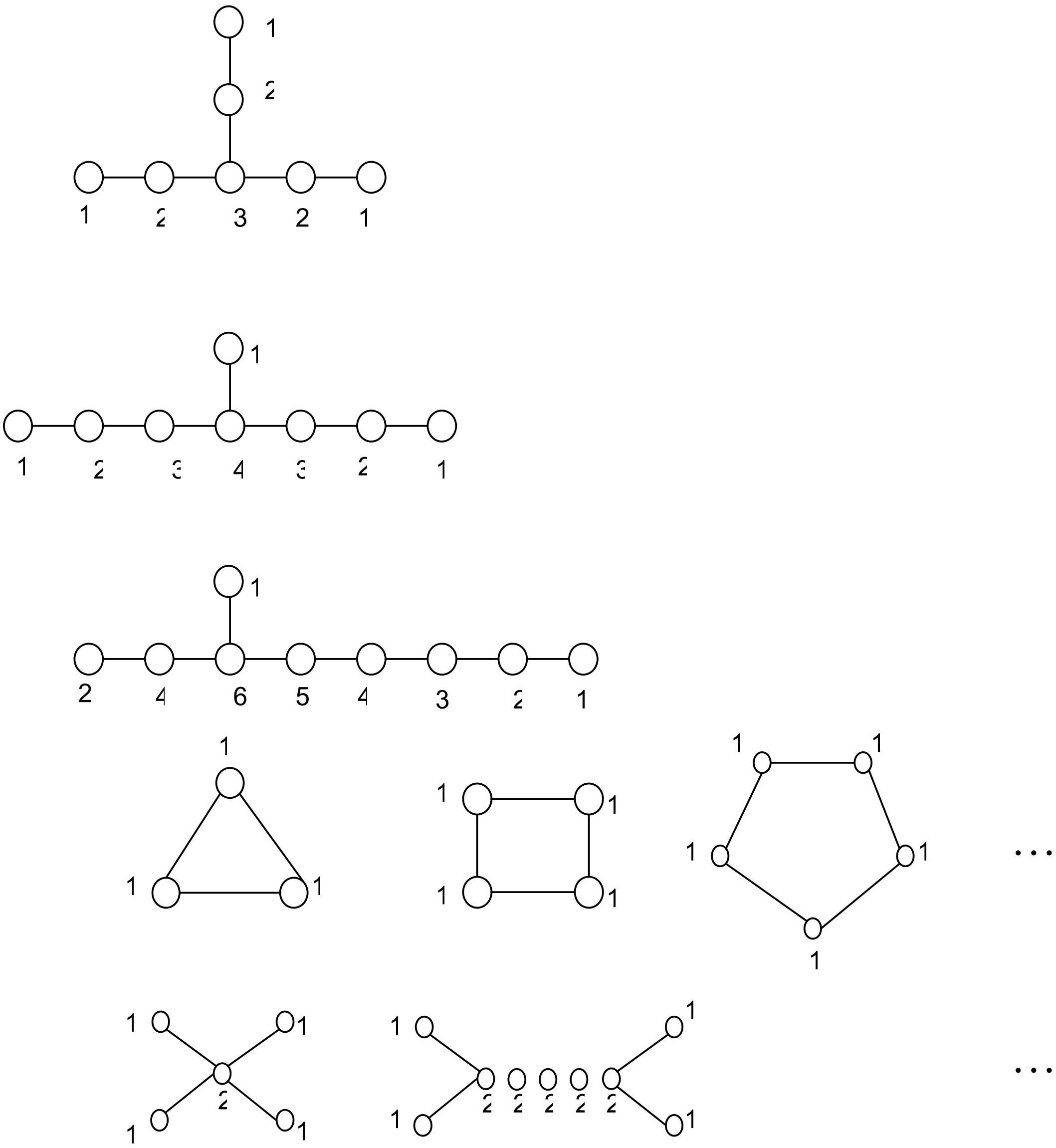


. . .



. . .

f) Let x denote the vector of the initial state and let c denote the vector of the numbers at the vertices



then

1) If $c^T x$ <0, the game is not periodical and does not converge,
2) If $c^T x$ =0, the game loops,
3) If $c^T x$ <0, the game loops (this is the Olympic example).

3. Stability

This section is based on Chapter 2 of [Horn and Johnson 1991].

Consider the first-order linear system of n ordinary differential equations

$$\frac{dx}{dt} = A(x(t) - \hat{x}), A \in R^{n \times n}$$

If at time $\hat{t}$, $x(\hat{t}) = \hat{x}$, x(t) will cease changing at $t = \hat{t}$, so $\hat{x}$ is an equilibrium for the system. When A is nonsingular, x(t) will cease changing only when it has reached this equilibrium.

If x(t) converges to $\hat{x}$ for all choices of the initial data x(0), we say the system is *globally stable* and A is *negative stable*. (The term negative will soon become clear.)

The unique solution x(t) of the system is

$$X(t) = e^{At}(x(0) - \hat{x}) + \hat{x}$$

where $e^{At} = \sum_{k=0}^{\infty} \frac{1}{k!} A^k t^k$. (This series converges for all t and all A).

e^{At} → 0 iff each eigenvalue λ of A satisfies Re $\lambda < 0$, Thus A is negative stable iff all its eigenvalues lie in the open left half of the complex plane.
For those who prefer positivity, we define A to be *positive stable* if Re $\lambda > 0$ for every eigenvalue λ of A.

The *inertia* i(A) of $A \in C^{n \times n}$ is the triple $(i_+(A), i_-(A), i_0(A))$
where

$i_+(A)$ is the number of eigenvalues, including multiplicities, with positive real part,
$i_-(A)$ is the number of eigenvalues, including multiplicities, with negative real part,
$i_0(A)$ is the number of eigenvalues, including multiplicities, with zero real part.

Thus, $A \in C^{n \times n}$ is positive stable if $i_+(A) = n$, or in other words the inertia of A is (n,0,0).

Sylvester's law of inertia.
B=SAS* , for some nonsingular matrix S iff i(B)=i(A).

If $A \in C^{n \times n}$ is positive stable then so are:

a. aA+bI, $a \geq 0, b \geq 0, a + b > 0$

b. A^{-1}
c. A^{*}
d. A^{T}

If $A \in C^{n\times n}$ is positive stable, then Re tr A>0,
If $A \in R^{n\times n}$ is positive stable, then tr A>0, det A>0 (for n=2 this is iff).
Matrix stability and the problem of location of the roots of the polynomial are related through the companion matrix of the polynomial: Since f is the characteristic (and the minimal) polynomial of C_f , the roots of f lie in the open right half plane iff C_f is positive stable.

Lyapunov's Theorem
$A \in C^{n\times n}$ is positive stable if and only if there exists a positive definite matrix G such that $GA + A^{*}G$ is positive definite.
Furthermore, if for some positive definite matrix H, there exists an Hermitian matrix G, such that:
$GA + A^{*}G = H$
Then A is positive stable iff G is positive definite.

Given matrices $A, B \in C^{n\times n}$ it is known that the matrix equation
$AX - XB = C$
Has a unique solution X for every C iff no eigenvalue of A is an eigenvalue of B
$(\sigma(A) \cap \sigma(B) = \varnothing)$
Thus
$GA + A^{*}G = H$
has a unique solution G for every given right hand side H iff $\sigma(A^{*}) \cap \sigma(-A) = \varnothing$ a condition that holds when A is positive stable, so if A is positive stable, then for every H, there is a unique G such that
$GA + A^{*}G = H$
Furthermore, if H is Hermitian so is G and if H is positive definite then G is also positive definite. In particular, H can be chosen as the identity matrix and then Lyapunov theorem can be stated as
A is positive stable iff there is a positive definite matrix G such that
$GA + A^{*}G = I$
If A is positive stable, the above equation has a unique solution G and G is positive definite. Conversely, if for a given A, there exists a positive definite solution G, then A is positive stable and G is the unique solution to the equation. So there is a way to check the stability of a matrix A:
Solve the equation $GA + A^{*}G = I$ If a solution does not exist or if a solution exists but is not positive definite, A is not positive stable. If there is a positive definite solution then (it is unique and) A is positive stable.
This theorem and check is lovely but costly. For real matrices there is a less costly algorithm given the Routh-Hurwitz conditions:

Let $A \in R^{n\times n}$ and let E_k denote the sum of all $\begin{pmatrix} n \\ k \end{pmatrix}$ principal minor of A,

$(E_1 = trace \quad A, E_n = \det \quad A)$.

The Routh-Hurwitz matrix of A is the $n\times n$ matrix $\Omega = \Omega(A)$ where $w_{ii} = E_i$, in the entries above E_i are $w_{i-1,i} = E_{i+1}, \quad w_{i-2,i} = E_{i+2}, ...,$ up to the first row $w_{1,i}$ or to E_n, whichever comes first. The entries above E_n are zero. The entries below E_i are the first n-i elements of the sequence $E_{i-1}, \quad E_{i-2}, ..., E_1, 1, 0, ..., 0$ for example, for n=5

$$\Omega = \begin{pmatrix} E_1 & E_3 & E_5 & 0 & 0 \\ 1 & E_2 & E_4 & 0 & 0 \\ 0 & E_1 & E_3 & E_5 & 0 \\ 0 & 1 & E_2 & E_4 & 0 \\ 0 & 0 & E_1 & E_3 & E_5 \end{pmatrix}$$

The Routh-Hurwitz Theorem

$A \in R^{n\times n}$ is positive stable iff the n leading principal minors of $\Omega(A)$ are positive.

Example:

We know that a 2x2 real matrix A is positive stable iff E_1 and E_2 are positive. Indeed, by the Routh-Hurwitz conditions

$$\Omega = \begin{pmatrix} E_1 & 0 \\ 1 & E_2 \end{pmatrix}$$

and the leading principal minors are E_1 and $E_1 E_2$

M- matrices

A *Z- Matrix* is a real matrix of the form $A = \alpha I - B$, where B is a nonnegative matrix, i.e. a matrix whose off diagonal entries are non positive.

A matrix $A = \alpha I - B, \quad B \geq 0$ is an *M- matrix* (M for Minkowski) if $\alpha > \rho(B)$. (If $\alpha = \rho(B)$, A is a singular M- matrix).

By the Perron-Frobenius theorem, a Z- matrix is an M- matrix iff it is positive stable.

The following theorem is a partial list of conditions a Z- matrix that are equivalent to being M- matrix. For a longer list and some of the proofs , see [Berman and Plemmons, 1979,1994, Chapter 6]

Theorem

Let A be a Z-matrix, $A = \alpha I - B$; α real, B nonnegative.

Then the following conditions are equivalent:

a) $\alpha > \rho(B)$, i.e., A is an M-matrix
b) A is *inverse positive*, i.e. A^{-1} exists and is nonnegative.
c) A is *monotone*, i.e. $Ax \geq 0 \Rightarrow x \geq 0$.

d) A is *diagonally stable*, i.e. there exists a positive diagonal matrix D such that $DA + A^T D$ is positive definite.
e) A is positive stable.
f) A is a *P-matrix*, i.e. all the principal minors of A are positive.
g) A does not reverse the sign of any vector, i.e., if $x \neq 0$ and y=Ax, then for some i, $x_i y_i > 0$.
h) Every real eigenvalue of a principal submatrix of A is positive.
i) A+tI is nonsingular for all $t \geq 0$.
j) Every real eigenvalue of A is positive.

Proof.
We will show that

1) For every $A \in R^{n \times n}$
$(b) \Leftrightarrow (c)$ and

$$\begin{array}{ccccccc}
(d) & & \Rightarrow & & (e) & & \\
\Downarrow & & & & & & \\
(f) & \Rightarrow & (g) & & \Downarrow & & \\
\nwarrow & & \swarrow & & & & \\
& (h) & & \Rightarrow & (i) & \Leftrightarrow & (j)
\end{array}$$

2)
$$\begin{array}{l}
(a) \Rightarrow (b) \\
\Downarrow \\
(d)
\end{array}$$

3)
$$\begin{array}{l}
(c) \Rightarrow (a) \\
\textit{and} \\
(i) \Rightarrow (a)
\end{array}$$

So here we go:

$(b) \Rightarrow (c)$: Ax $\geq 0 \Rightarrow A^{-1}$Ax ≥ 0 => $x \geq 0$.

(c) $\Rightarrow$ (b): The nullspace of A is a subspace so if only nonnegative vectors are mapped to zero, A must be nonsingular. Multiplying both sides of (c) by A^{-1} yields $x \geq 0 \Rightarrow A^{-1} x \geq 0$, so A^{-1} is nonnegative.

(d) $\Rightarrow$ (e) $\Rightarrow$ (i): trivial.

(d) $\Rightarrow$ (f): If A is diagonally stable then so are its principal submatrices. Thus all the principal submatrices of A are stable and, being real, have positive determinants.

(f) $\Rightarrow$ (g): First we observe that if A is a P-matrix and D is a nonnegative diagonal matrix, then
det(P+D) $\geq$ det(P) (if D is nonzero the inequality is strict).
This follows from the fact that for any matrix A
$\frac{\partial}{\partial a_{ij}} \det A = (-1)^{i+j} \det A_{ij}$, the i,j cofactor of A.
Now suppose that $x \neq 0$ and that $x \circ Ax \leq 0$ ($\circ$ denotes the *Hadamard product*; $A \circ B_{ij} = a_{ij} b_{ij}$. Let s be the set of indices i s.t. $x_i \neq 0$, and let A[s] and x[s] denote the corresponding principal submatrix and subvector.
Then
x[s] $\circ$ (A[s]x[s]) ≤ 0
so there is a nonnegative diagonal matrix D for which
A[s]x[s]=-Dx[s] so (A[s]+D)x[s]=0 so A[s]+D is singular. This contradicts the fact that A[s] is also a P-matrix and thus
det(A[s]+D) $\geq$ detA[s]>0.

(g) $\Rightarrow$ (h): Let k be an index such that $x_k (Ax)_k > 0$. Then there exists ε>0 such that
$x_k (Ax)_k = \varepsilon \sum_{i \neq k} x_i (Ax)_i > 0$.
Let D be a diagonal matrix with d_k =1 and all other diagonal entries equal to ε. Then
$x^T DAx > 0$.
Now let λ be a real eigenvalue of a principal submatrix A[s] of A,
$A[s]y = \lambda y, \quad y \neq 0, \quad \lambda \quad real.$
We want to show that $\lambda > 0$.
Let x be a vector such that x[s]=y and all other entries of x are zeros, and let D be a nonnegative diagonal matrix for which $x^T DAx > 0$.
Then
$0 < x^T DAx = (Dx)^T Ax = (D[s]y)^T A[s]y = (D[s]y)^T \lambda y = \lambda (y^T D[s]y)$ proving that $\lambda > 0$.

(h) $\Rightarrow$ (f): The spectrum of a real matrix is closed under complex conjugation, and since the real eigenvalues of every principal submatrix are positive, all the principal minors are positive.

(h) $\Rightarrow$ (i): trivial.

(i) $\Rightarrow$ (j) since λ is an eigenvalue of A iff $A - \lambda I$ is singular.

(a) $\Rightarrow$ (b): Since $\alpha > \rho(B)$, A is nonsingular. A^{-1} is nonnegative since

$(A/\rho(A))^{-1} = I + (A/\rho(A)) + (A/\rho(A))^2 + ...$

(a) $\Rightarrow$ (d): Since $\rho(B) \geq b_{ii}$, $\alpha > b_{ii}$ so the diagonal entries of A are positive.
Since A^{-1} is nonnegative, $x = A^{-1}e > 0$. Let E=diag(x). Then AEe=Ax=e, and since the diagonal entries of AE are positive, it is strictly row diagonally dominant. $(AE)^{-1}$ is also nonnegative so $y^T = e^T(AE)^{-1} > 0$ so for D=diag(y), $e^T DAE = y^T AE = e^T$, so DAE is strictly column diagonally dominant. It is also strictly row diagonally dominant since it was obtained from such a matrix by multiplication on the left by a positive diagonal matrix. By Gersgorin, $DAE + (DAE)^T$ is positive definite. By Sylvester Law of Inertia $E^{-1}(DAE + (DAE)^T)E^{-1}$ is also positive definite, so $(E^{-1}D)A + A^T(E^{-1}D)$ is positive definite, meaning that A is diagonally stable.

(c) $\Rightarrow$ (a): Let v be a Perron vector for B, $Bv = \rho(B)v$; $v \geq 0$, $v \neq 0$. If $\alpha \leq \rho(B)$ then $A(-v) = (\rho(B) - \alpha)v \geq 0$, but –v is not nonnegative, contradicting (c).

(i) $\Rightarrow$ (a): $\alpha - \rho(B)$ is a real eigenvalue of A so by (i) it is positive.

We conclude this section with few additional remarks on M-matrices and stability.

D- stability

A is *D-stable* if DA is positive stable for every positive diagonal matrix D.
D-stable matrices appear in models in economics, for example in study of stability of prices in multiple markets.
Diagonal stability implies D stability, for if D and E are positive diagonal matrices and $EA + A^T E = B$ is positive definite then
$(ED^{-1})DA + DA^T(ED^{-1}) = EA + A^T E = B$.

A matrix A is said to be a P_0^+ *matrix* if all its principal minors are nonnegative and for every order, one of them is positive.

There are D-stable matrices that are not P-matrices , for example $A = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$ but every D-stable matrix is P_0^+.

DIAGONAL STABILITY

Diagonally stable matrices appear in Predator-Prey systems. They are also called *Volterra-Lyapunov stable*)
They were characterized in [Barker, Berman and Plemmons 1978] as follows:

An $n \times n$ matrix is diagonally stable iff for every nonzero $n \times n$ positive semi definite matrix B, BA has a positive diagonal entry.

H-matrices

A is an *H-matrix* if its comparison matrix, defined in the section on completely positive matrices, is an M-matrix.

An H-matrix with positive diagonal entries is diagonally stable and thus also D-stable.

Irreducible M-matrices

If A is an irreducible M-Matrix then its inverse is positive (not only nonnegative).
If A is an $n \times n$ irreducible singular M-matrix then rank A =n-1, every proper principal submatrix of A is a (nonsingular) M-matrix and A is *almost monotone*, i.e.
$Ax \geq 0 \Rightarrow Ax=0$.

4. Applications and Extensions

Operators that leave a proper cone invariant

An $n \times n$ A nonnegative matrix can be characterized by
$x \in R_+^n \Rightarrow Ax \in R_+^n$
Where R_+^n is the nonnegative orthant of R^n.
This orthant is an example of a proper cone (that we will define soon) and the Perron-Frobenius theory has natural extensions to maps that leave a proper cone invariant. We conclude the lectures with some examples (where K is a convex cone in R^n)
First some definitions:
A convex cone k is *pointed* if the intersection of K and –K is $\{0\}$ and *solid* if int K, the interior of K is not empty, A closed, pointed and solid convex cone is called a *proper* cone.
Let K be a proper cone in R^n, A matrix A is *K-nonnegative* if A K $\subseteq$K; A is *K-positive* if A(K-$\{0\}$) $\subseteq$ int K; A is *K-irreducible* if it is K-nonnegative and no eigenvector of A lies on bd K (the boundary of K); and A is *K-primitive* if it is K-nonnegative and the only nonempty subset of bd K which is left invariant by A is $\{0\}$.

Theorem

Let $A \in R^{n \times n}$ of bd K be a proper cone in R^n. Then:

a. If A is K-nonnegative, then $\rho(A)$ is an eigenvector and K contains a corresponding eigenvector;
b. If A is K-positive, then $\rho(A)$ is greater than the absolute value of any other eigenvalue and an eigenvector corresponding to $\rho(A)$ lies in int K.
c. If A is K-irreducible then $\rho(A)$ is a simple eigenvalue and any other eigenvalue with the same modulus is also simple; there is an eigenvector corresponding to $\rho(A)$ in int K, and no other eigenvector (up to scalar multiples) lies in K;
d. A K-nonnegative matrix is K-primitive iff for some natural number m, A^m is K-positive.

The theory of K-nonnegative matrices is much richer. For an extensive study of them the reader is referred to [Tam and Schneider 2006].

There are other directions in which the Perron-Frobenius Theory can be extended, for example to matrices that are not necessarily real and to nonlinear operators. This will not be done in these lectures.
In the course we saw several applications of the theory of nonnegative matrices. In particular we were motivated by the application to ranking of web sites. We conclude the notes with another application to ranking and another internet application.

Tournaments

In a *tournament* every player plays any other player. In each game there is a clear win. A tournament can be described by a *tournament digraph* in which there is an arc from i to j iff i beats j. Thus , for two different players, there is an arc from i to j, iff there is no arc from j to i.
This means that the adjacency matrix (*tournament matrix*) A is a (0,1) matrix satisfying $A+A^T$ = I+J.
The idea here is that players who beat players with high rank should have a high rank.
The vector of scores is Ae.
The vector that sums the scores of those defeated by the players is
A^2 e, …

Recall that for primitive matrices $\lim_{m\to\infty}(A/(\rho(A))^m \equiv L = xy^T$, where

Ax= ρ (A)x, x>0; $A^T y = \rho(A) y$,y>0; $x^T y = 1$, Therefore if the adjacency matrix of the tournament is primitive, Le is a positive multiple of the Perron vector of A that naturally gives the ranking of the players.

For the general case we are lucky to have the following theorem.

Theorem

An $n \times n$ irreducible tournament is primitive iff n is greater than 3.

The irreducible case n=3 happens only when a beats b, b beats c and c beats a, and in this case it is natural to rank all three equally.

In the more general case when the tournament matrix is reducible, look at its Frobenius normal form. Since it is a tournament matrix all its entries above the diagonal blocks are equal to 1 which means that for i>j the players in block i should be ranked lower than those in block j. An irreducible tournament block is at least of order 3. If a block is 3x3 all its players are ranked equal and if it has more than 3 players, the Perron vector can be used.

This ranking is known as *the Kendall-Wei ranking.*

TCP

TCP stands for Transmission Control Protocol.
A good reference on internet congestion control is the book [Srikant 2003] where the models used are fluid dynamic models. Recently, algebraic models, using the theory of Nonnegative Matrices, were developed in the Hamilton Institute in Ireland. Here I will describe these models.

A wireline network consists of sources and sinks that communicate via networks links (wires) and routers (queues). Packets are acknowledged or lost because of congestion. Source i has a window size w_i which is the number of packets sent till the first is acknowledged or a congestion occurs.
TCP uses an Additive Increase Multiplicative Decrease congestion algorithm (see, for example [Berman, Shorten and Leigh 2004]:

When source i receives an acknowledgement it increases its window size

$w_i \to w_i + \alpha'_i / w_i ; \alpha_i > 0$.

When a packet is dropped, the window size is decreased

$w_i \to \beta_i w_i ; 0 \prec \beta_i \prec 1$

In standard TCP, $\alpha_i = 1$, $\beta_i = 0.5$.

Notation: $\alpha_i = \alpha'_i / (\sum \alpha_i)$.

Thus, α_i is positive and $\sum \alpha_i = 1$.

A synchronized model

Under a synchronization assumption, that all the sources simultaneously decrease their window sizes in case of congestion

$\omega(k+1) = A\omega(k)$
Where $\omega(k)$ is the vector of window sizes at event k and

$$A = D + xy^T = \begin{pmatrix} \beta_1 & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \beta_n \end{pmatrix} + \begin{pmatrix} \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_n \end{pmatrix} (1-\beta_1, \ldots, 1-\beta_n)$$

$$(0 < \alpha_i, \ \ \beta_i < 1; \ \ \sum_{i=1}^{n} \alpha_i = 1)$$

A TCP matrix is positive column stochastic (PSC) so the positive dynamical system $\omega(k+1) = A\omega(k)$
Possesses a unique stationary point

$$\left(\frac{C\alpha_1}{1-\beta_1}, \ldots, \frac{C\alpha_n}{1-\beta_n} \right)^T, C > 0$$

Which is a multiple of the Perron vector of A.
Thus, all sources get a fair share of the system iff $\frac{\alpha_i}{(1-\beta_i)}$ does not depend on i, i.e. when A is symmetric.
When A is PCS, $\rho(A) = 1 = \lambda_1$ and the rate of convergence to the stationary point is bounded by $\mu(A) < 1$.

A TCP matrix $A = D + xy^T$ is diagonally similar to diag $\left\{ \frac{\sqrt{y_i}}{\sqrt{x_i}} \right\}$ $D + xy^T$ diag $\left\{ \frac{\sqrt{x_i}}{\sqrt{y_i}} \right\}$ which is a (positive) symmetric rank 1 perturbation of $D = diag \ \ \{\beta_1, \ldots, \ \beta_n\}$. Thus the eigenvalues of A, except 1, interlace the β's. In particular, they are positive and $\mu(A)$ lies between the two largest β's.

Recall the tradeoff in choosing α in the Google page rank discussion. Here there is a similar tradeoff; higher β's give better use of the network but yield a slower convergence to the equilibrium state.

Here we discussed only the synchronized model but we must admit that it is quite restrictive (it is valid in some long distance networks). The matrix model can be generalized also to the unsynchronized case, see for example [Berman, Laffey, Leizarowitz and Shorten 2006]. A major difference between the synchronized and the unsynchronized models is that while in the first one studies convergence of homogeneous matrix products A^k where A is a nonnegative matrix, in the latter one looks at nonhomogeneous matrix products $A_k \; A_{k-1} \; \ldots \; A_1$ where $A_1 \; A_2 \; \ldots \; A_k$ are nonnegative matrices. An important reference on such products is [Hartfiel 2002].

References

R. B. Bapat and T. E. S. Raghavan, Nonnegative Matrices and Applications, Encyclopedia of Mathematics and its Applications 64, Cambridge University Press, 1997.

G. P. Barker, A. Berman and R. J. Plemmos, Positive diagonal solutions to the Lyapunov equation, Linear and Multilinear Algebra 5: 249-256, 1978.

A. Berman and R. Grone, Completely positive bipartite matrices, Math. Proc. Cambridge Philos. Soc. 103:269-276, 1998.

A. Berman and D. Hershkowitz, Combinatorial results on completely positive matrices, LAA 95: 111-125, 1987.

A. Berman, T. Laffey, A. Leizarowitz and R. Shorten, On the second eigenvalue of matrices associated with TCP, LAA 416: 175-184, 2006.

A. Berman, M. Neumann and R. J. Stern, Nonnegative Matrices in Dynamic Systems, Wiley-Interscience Series in Pure and Applied Mathematics, John Wiley and Sons, 1989.

A. Berman and R. J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, Academic Press, 1979.

A. Berman and R. J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, 2nd Edition, Classics in Applied Mathematics, SIAM, 1994.

A. Berman and N. Shaked-Monderer, Completely Positive Matrices, Worlsd Scientific, 2003.

A. Berman and N. Shaked-Monderer, Nonnegative matrices and digraphs, Encyclopedia of Complexity and System Science, 2008.

A. Berman, R. Shorten and D. Leigh, Positive matrices associated with synchronized communication networks, LAA 393:47-54, 2004.

G. Birkhoff, Tres observaciones sobre el algebra lineal, Univ. Nac. Tucuman Rev. Ser. A5: 147-150, 1946.

A. Borobia, J. Moro and R. Soto, Negativity compensation in the nonnegative inverse eigenvalue problem, LAA 393: 73-89, 2004.

M. Boyle and D. Handelman, The spectra of nonnegative matrices via symbolic dynamics, Ann. Math. 133: 249-316, 1991.

P. Diaconis, Lecture in IMA, 1994.

J. H. Drew, C. R. Johnson and R. Loewy, Completely positive matrices associated with M-matrices, Linear and Multilinear Algebra 37:303-310, 1994.

G.P. Egorichev, The solution of van der Waerden's problem on permanents, Advances in Math., 42: 299-305, 1981.

K. Eriksson, Convergence of Mozes' game of numbers, LAA 166: 151-165, 1992.

D. I. Falikman, Proof of the van der Waerden's conjecture on the permanent of a doubly stochastic matrix, Mat. Zametki 29, 6: 931-938, 957, 1981, (in Russian).

G. Frobenius, Uber matrizen aus nichtnegativen elemente, Sitzungsber Kon Preuss Acad. Wiss. Berlin, 456-457, 1912.

L. J. Gray and C. R. Wilson, Nonnegative factorization of positive semidefinite nonnegative matrices, LAA 31: 119-127, 1980.

G. H. Hardy, J. E. Littlewood and G. Polya, Some simple inequalities satisfied bu convex functions, Messenger of Math. 58, 145-152, 1929.

D. J. Hatfiel, Nonhomogeneous Matrix Products, World Scientific, 2002.

M. Hall Jr., Combinatorial Theory, 2nd ed. Wiley, 1986.

R. Horn and C. A. Johnson, Matrix Analysis, Cambridge University Press, 1985.

R. H. Horn and C. A. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991.

C. A. Johnson, Row stochastic matrices similar to doubly stochastic matrices, Linear Multilinear Algebra 10: 113-130, 1981.

C. A. Johnson, T. Laffey and R. Loewy, The real and the symmetric nonnegative eigenvalue problems are different, Proc. AMS 24:3647-3651, 1996.

M. Kaykobad, On nonnegative factorization of matrices, LAA 241-243: 655-657, 1987

C. Kelly, A test of the Markovian model of DNA evolution, Biometrics 50: 653-664, 1994.

N. Kogan and A. Berman, Characterization of completely positive graphs, Discrete Math 114: 298-304, 1993.

T. Laffey and H. Smigoc, Nonnegative realization of spectra having negative real parts, LAA 416:m 148-159, 2006

Y. Li, A. Kummert and A. Frommer, A linear programming based analysis of the cp-rank of completely positive matrices, Int. J. Appl. Math. Comut. Sci. 1: 25-31, 2004

N. Linial. Private communication, 1998.

R. Loewy and D. London, A note on an inverse problem for nonnegative matrices, Linear Multilinear Algebra 6:83-90, 1978

R. Loewy and J. J. McDonald, The symmetric nonnegative inverse eigenvalue problem for 5x5 matrices, LAA 393: 275-298, 2004.

J. E. Maxfield and H. Minc, On the Matrix equation X'X = A, Proc. Edinburgh Math, Soc. 13: 125-129, 1962.

M. E. Meehan, Some results on matrix spectra, Ph.D Thesis, University College Dublin, 1998.

H. Minc, Nonnegative Matrices, Wiley, 1988.

S. Mozes, Reflection processes on graphs and Weyl groups, J. Com. Theory. Ser.A 53(1): 128-142, 1990

O. Perron, Zur theorie der uber matrizen, Math Ann., 64:248-263, 1907.

V. Romanovsky, Recherche sur les chains de Markoff, Acta Math 66, 147-251, 1936.

U. Rothblum, Nonnegative matrices and stochastic matrices, in L. Hogben, editor, Handbook of Linear Algebra, CRC Press, 2006.

E. Senata, Nonnegative Matrices and Markov Chains, Springer Series in Statistics, Springer-Verlag, 1981. 64, Cambridge University Press, 1997.

H. Smigoc, The inverse eigenvalue problem for nonnegative matrices, LAA 393: 365-374, 2004.

R. Srikant, The Mathematics of Internet Congestion Control, Birkhauser, 2003.

B. S. Tam and H. Schneider, Matrices that leave a cone invariant, in L. Hogben, editor, Handbook of Linear Algebra, CRC Press, 2006.

H. Wielandt, Unzerlegbare, nicht negative Matrizen, Math Z. 52, 642-648, 1950.