Positive matrices

Def. $A \in \mathbb{R}^{m\times n}$ is called a **positive matrix** ($A > 0$), if $a_{ij} > 0$ for all i, j.
A is a **nonnegative matrix** ($A \geq 0$) if $a_{ij} \geq 0$ for all i, j.
We define **positive and nonnegativ vectors** analogously.
A column vector $\mathbf{v} \in \mathbb{R}^n$ is called a **stochastic vector** if

$$\mathbf{v} \geq 0 \text{ and } \sum v_i = 1$$

(i.e. $\mathbf{v} \geq 0$ and $\mathbf{1}^T\mathbf{v} = 1$, where $\mathbf{1}$ denotes the column vector whose every component is 1).
$A \in \mathbb{R}^{n\times n}$ is a **stochastic matrix** if all of its columns are stochastic vectors.

Example. The rank matrices of the PageRank algorithm are stochastic matrices. The modified link matrix $\hat{A}$ corresponding to the "impatient surfer" is a positive stochastic matrix.

Theorem (Perron). If $A \in \mathbb{R}^{n\times n}$ and $A > 0$, then
1) $\rho(A)$ is an eigenvalue of A.
2) The multiplicity of $\rho(A)$ in $k_A(x)$ is 1.
3) For every eigenvalue λ of A as of a complex matrix, if $\lambda \neq \rho(A)$, then $|\lambda| < \rho(A)$ (i.e. $\rho(A)$ is **the dominant eigenvalue** of A).
4) The 1-dimensional eigenspace for $\rho(A)$ is generated by a positive vector, and $\rho(A)$ is the only eigenvalue for which there is a nonnegative eigenvector.

Example.

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 3 & 3 & 2 \end{bmatrix} > 0, \quad k_A(x) = \begin{vmatrix} 1-x & 2 & 1 \\ 2 & 1-x & 1 \\ 3 & 3 & 2-x \end{vmatrix} = -x(x+1)(x-5)$$

So the dominant eigenvalue is $5 = \rho(A)$ (the others, 0 and -1, are smaller in absolute value).
The eigenvectors:

$$\text{for } \lambda = 5: \quad \begin{bmatrix} -4 & 2 & 1 \\ 2 & -4 & 1 \\ 3 & 3 & -3 \end{bmatrix} \mapsto\mapsto \begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \mathbf{x} = \frac{t}{2}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix},$$

so $(1, 1, 2)$ is a positive eigenvector for $\lambda = 5$;

$$\text{for } \lambda = 0: \quad \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 3 & 3 & 2 \end{bmatrix} \mapsto\mapsto \begin{bmatrix} 1 & 0 & \frac{1}{3} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \mathbf{x} = \frac{t}{3}\begin{bmatrix} -1 \\ -1 \\ 3 \end{bmatrix},$$

$$\text{for } \lambda = -1: \quad \begin{bmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \\ 3 & 3 & 3 \end{bmatrix} \mapsto\mapsto \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \mathbf{x} = t\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

showing that there are no positive (or nonnegative) eigenvectors for 0 and -1.

Cor.: Every real positive 2×2 matrix is diagonalizable in $\mathbb{R}^{2\times 2}$, since by Perron's theorem $\lambda_1 = \rho(A) \in \mathbb{R}$, and $\lambda_1 + \lambda_2 = \operatorname{tr} A \in \mathbb{R}$, so $\lambda_2 \in \mathbb{R}$. Furthermore, $\lambda_1 \neq \lambda_2$ by Perron/2, so

A is diagonalizable.

For example, $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ cannot be conjugated by a real matrix P to a positive matrix.

Proposition If A is a positive stochastic matrix, then $\rho(A) = 1$.

Proof. $\mathbf{1}^T A = \mathbf{1}^T \Rightarrow A^T \mathbf{1} = \mathbf{1}$, so the positive A^T matrix has a positive eigenvector for the eigenvalue 1, and then by Perron/4, $\rho(A) = \rho(A^T) = 1$.

Corollary. If $A \in \mathbb{R}^{n\times n}$ is a positive stochastic matrix, and $\mathbf{v}_0 \in \mathbb{R}^n$ is a stochastic vector, then $\lim\limits_{k\to\infty} A^k \mathbf{v}_0 = \mathbf{u}$ is the unique stochastic eigenvector for $\lambda = 1$.

Proof. Note that by Perron's theorem, in the Jordan form of A there is only a 1×1 block for 1, and all other blocks are for eigenvalues of absolute value smaller than 1, so by the convergence theorem of the previous lecture, A^k is convergent. Furthermore, $\lim A^k \mathbf{v}_0 \in V_1 = \mathrm{span}(\mathbf{u})$.

But if $\mathbf{1}^T \mathbf{v} = 1$, then $\mathbf{1}^T A\mathbf{v} = \mathbf{1}^T \mathbf{v} = 1$, so $\mathbf{v}_0,\ A\mathbf{v}_0,\ A^2\mathbf{v}_0, \ldots$ are all stochastic vectors, thus the limit vector is also stochastic. It is also clear that $\{\, c\mathbf{u} \mid c \in \mathbb{R}\,\}$ contains only one stochastic vector. □

So it follows that the distribution vector at the random surfer method indeed converges to the rank vector, which is the stochastic eigenvector of the modified link matrix, $\hat{A}$ for $\lambda = 1$. Perron's theorem also shows that in this case the eigenspace for 1 is one-dimensional, so the rank vector is essentially unique. We shall see that the requirement of positivity can be weakened.

A simplified example from economy. Let us try to study the housing condition of people in a given country. Qualify some apartments as being in good condition and some apartments as being in poor condition. Suppose that research shows that the lack of maintenance on some apartments will turn each year 5% of the apartments of good quality to poor ones; on the other hand, renovation of apartments of poor quality will turn 10% of them to good ones. What is the distribution of the good and poor apartments in the long run if the present situation is that 40% of the apartments are in good shape and 60% of them are in bad shape.

One can see that if $\mathbf{v}_0 = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$ is the initial distribution vector then distribution vector $\mathbf{v}_1$, describing the conditions in one year can be obtained as $\mathbf{v}_1 = A\mathbf{v}_0$, where $A = \begin{bmatrix} 0.95 & 0.1 \\ 0.05 & 0.9 \end{bmatrix}$ is the so called **transition matrix**. By the previous corollary, the limit of the distribution vectors is the stochastic eigenvector of A for eigenvalue 1.

$$A - I = \begin{bmatrix} -0.05 & 0.1 \\ 0.05 & -0.1 \end{bmatrix} \mapsto \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix} \Rightarrow \mathbf{x} = t \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Among the eigenvectors $\mathbf{x}$ is a stochastic vector if $2t + t = 1$, i.e. $t = \frac{1}{3}$, so $\mathbf{u} = \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$. So in the long run, 2/3 of the apartments will be in good condition, and 1/3 will be in poor condition.

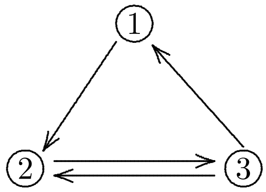
Nonnegative matrices

Def. A nonnegative matrix $A \in \mathbb{R}^{n\times n}$ is **primitive** if there is a $k > 0$ such that $A^k > 0$.

Lemma. If $A \geq 0$, and $A^k > 0$, then $A^{k+1} > 0$.
Proof. Note that the dot product of a positive and a nonzero nonnegative vector is positiv, so $A^k\mathbf{v} > 0$ if $\mathbf{0} \neq \mathbf{v} \geq 0$. On the other hand, $A^k > 0$ insures that no column of A can be $\mathbf{0}$, so $A^kA > 0$.

Corollary. The convergence theorem above holds for primitive stochastic matrices, as well.

Example. The link matrix of the following graph is primitive.



$$A = \begin{bmatrix} 0 & 0 & 1/2 \\ 1 & 0 & 1/2 \\ 0 & 1 & 0 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 0 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 1 & 0 & 1/2 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 1/2 & 0 & 1/4 \\ 1/2 & 1/2 & 1/4 \\ 0 & 1/2 & 1/2 \end{bmatrix},$$

$$A^4 = \begin{bmatrix} 0 & 1/4 & 1/4 \\ 1/2 & 1/4 & 1/2 \\ 1/2 & 1/2 & 1/4 \end{bmatrix}, \quad A^5 = \begin{bmatrix} 1/4 & 1/4 & 1/8 \\ 1/4 & 1/2 & 3/8 \\ 1/2 & 1/4 & 1/2 \end{bmatrix}$$

Actually, only the position of the positive entries in the matrix matter, not the actual values, when we want to decide if it has a positive power.

Def. To a nonnegative matrix $A \in \mathbb{R}^{n\times n}$, we can assign an oriented graph $G(A)$ on vertices $\{1, 2, \ldots, n\}$ so that there is an arrow from i to $j \Leftrightarrow a_{ij} > 0$.

Lemma. If $A \geq 0$, then $(A^k)_{ij} > 0 \Leftrightarrow$ there is an oriented walk of length k from i to j in $G(A)$.
Proof. $(A^2)_{ij} = \sum_t a_{it}a_{tj} > 0 \Leftrightarrow \exists t$: $a_{it} > 0$ and $a_{tj} > 0 \Leftrightarrow \exists t$: $i \to t \to j$, i.e. there is a walk of length 2 from i to j in $G(A)$. With induction on k, we get the statement of the lemma.

Def. We call a nonnegative square matrix **irreducible** if $G(A)$ is strongly connected, i.e. from every i to every j there is an oriented path.

Theorem. A nonnegative matrix is primitive $\Leftrightarrow$ it is irreducible, and the greatest common divisor of the lengthes of the oriented cycles in $G(A)$ is 1.
Proof of the $\Rightarrow$ direction. If A is not irreducible then there are $i \neq j$ such that $(A^k)_{ij} = 0$ for all k, so A is not primitive. If the gcd of the cycle lengthes is $d > 1$, then the diagonal elements of A^{kd+1} are zero for every k, so A is not primitive.

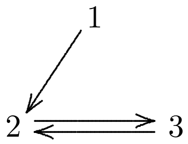
Example. The graph assigned to the previous link matrix is the same as the original, only with reversed arrows. This is strongly connected since it has a cycle going through all vertices, and it has a 3-cycle, and a 2-cycle as well, so the gcd of the cycle lengthes is 1.

Exercise. Determine, which of the following nonnegative matrices are irreducible, and which are primitive.

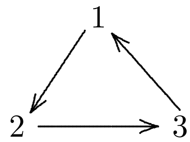
$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 5 \\ 1 & 0 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 5 \\ 1 & 0 & 3 \end{bmatrix}$$

Solution. The graphs assigned to A and B are:

$G(A)$:



$G(B)$:



In $G(A)$ there is no path from 2 to 1, so A is not irreducible, and then it is not primitive, either. B is irreducible, because we can go along the 3-cycle from any vertex to any vertex, but it is not primitive since the gcd of the cycle lengthes is 3. $G(C)$ only differs from $G(B)$ by having an extra loop at 3. But this means that the gcd of the cycle lengthes became 1 (while the graph is still strongly connected), so C is primitive. (Note that if a matrix is irreducible and at least one of its diagonal elements is positive, then it is always primitive.)

Theorem. Perron's theorem holds for primitive matrices.
Proof. If A is primitive then there is a $k > 0$ such that A^k is positive.
It follows from the Jordan normal form that if the complex eigenvalues of A are $\lambda_1, \ldots, \lambda_n$ (with absolute values $\rho(A) = |\lambda_1| \geq \cdots \geq |\lambda_n|$), then A is similar to an upper triangular matrix T with diagonal elements $\lambda_1, \ldots, \lambda_n$, then $A^k \sim T^k$, whose diag. elements are $\lambda_1^k, \ldots, \lambda_n^k$, so these are the eigenvalues of T^k, and then of the positive matrix A^k, as well, with decreasing absolute values.
So $|\lambda_1^k| = \rho(A^k)$, and by Perron's theorem, $\lambda_1^k = \rho(A^k) > |\lambda_i^k|$ for all $i > 1$. This implies that $\rho(A) = |\lambda_1| > |\lambda_i|$ for all $i > 1$
($\Rightarrow$ Perron/2,3 for A.)
The eigenspace V_{λ_1} of A is in the eigenspace of A^k for λ_1^k, and the latter is 1-dimensional spanned by a positive vector $\mathbf{u}$. So $V_{\lambda_1} = \mathrm{span}(\mathbf{u})$.
($\Rightarrow$ the first half of Perron/4 for A.)
$A\mathbf{u} \geq 0$, and $\mathbf{u} > 0$, so λ_1 is real, and > 0, proving that $\lambda_1 = |\lambda_1| = \rho(A)$.
($\Rightarrow$ Perron/1 for A.)
If there would be a nonnegative eigenvector of A for some other eigenvalue, then it would be true for the positive matrix A^k, as well, contrary to Perron's theorem.
($\Rightarrow$ the second half of Perron/4. for A.)

Remark. Most of the statements of Perron's theorem also hold for irreducible matrices (Perron–Frobenius theorem).