Appunti di Perron-Frobenius

Nikita Deniskin

13 July 2020

1 Prerequisiti e fatti sparsi

e_j è j-esimo vettore della base canonica e il vettore $\mathbf{1} = (1, 1, \dots 1)^\top$.
Spectrum: $\sigma(A)$
Spectral radius: $\rho(A)$
Autospazio generalizzato: $E_A(\lambda) = \mathrm{Ker}(A - \lambda\,\mathrm{I})$
Algebraic multiplicity: m_i of λ_i
Geometric multiplicity: d_i of λ_i
λ_i is simple if $m_i = 1$, semi-simple if $m_i = d_i$, defective if $d_i < m_i$.
$\lim\limits_k A_k$ è il limite puntuale delle matrici.
Teorema di Gelfand e su $\rho(A) = \lim(||A^k||)^{\frac{1}{k}}$
Cose su Jordan e scrittura astratta con il proiettore.
Si ha che $\lim\limits_k B^k = 0$ se e solo se $\rho(B) < 1$.
Inoltre il limite di B^k non diverge solo se $\rho(B) \leq 1$ e 1 è un autovalore semisemplice (guarda la forma di Jordan).
DA COMPLETARE

Definition 1.1 Let $A \in \mathbb{C}^{n\times n}$. Then:

- A is *zero-convergent* if $\lim\limits_{n\to\infty} A^n = 0$
- A is *convergent* if $\lim\limits_{n\to\infty} A^n = B \in \mathbb{C}^{n\times n}$.

Achtung! Some authors call zero-convergent matrix *convergent* and convergent matrix *semi-convergent*.

Theorem 1.2 Let $A \in \mathbb{C}^{n\times n}$, A is zero-convergent $\iff$ $\rho(A) < 1$.

Proof. Write $A = CJC^{-1}$ where J is the Jordan Normal Form and note that A is zero-convergent $\iff$ J is zero-convergent. □

Theorem 1.3 Let $A \in \mathbb{C}^{n\times n}$, A is convergent $\iff$ ALL the three following conditions apply:

1. $\rho(A) \leq 1$

2. If $\rho(A) = 1$, then $\lambda \in \sigma(A), |\lambda| = 1 \Rightarrow \lambda = 1$
3. If $\rho(A) = 1$, then the eigenvalue $\lambda = 1$ is semi-simple.

Or, in an equivalent formulation, if we are in one of the following two mutually exclusive cases:

1. $\rho(A) < 1$
2. $\rho(A) = 1$, the eigenvalue $\mu = 1$ is semi-simple, and if $\lambda \in \sigma(A), |\lambda| = 1 \Rightarrow \lambda = 1$.

Proof. Use the Jordan Normal Form $\square$

Remark 1.4 1 is a semi-simple eigenvalue of A if and only if $\text{rank}(I - A) = \text{rank}(I - A)^2 < n$.
This is useful since it's computationally easier to calculate the rank than all the eigenvalues of A.

2 Perron-Frobenius theorem

2.1 Positive and non-negative matrices

Definition 2.1 (**Positive and non-negative vectors**) Let $v \in \mathbb{R}^n$ be a vector

- v is *positive* if for all i, $1 \leq i \leq n$, we have $v_i > 0$. We will indicate that v is positive by writing $v > 0$.
- v is *non-negative* if for all i, $1 \leq i \leq n$, we have $v_i \geq 0$. We will indicate that v is non-negative by writing $v \geq 0$.

The set of all non-negative vector of dimension n is called the *non-negative orthant* of $\mathbb{R}^n$ and will be indicated with $\mathbb{R}^n_+$.

Definition 2.2 (**Positive and non-negative matrices**) Let $A \in \mathbb{R}^{n \times n}$ be a matrix

- A is *positive* if for all i, j, $1 \leq i, j \leq n$, we have $a_{ij} > 0$. We will indicate that A is positive by writing $A > 0$.
- A is *non-negative* if for all i, j, $1 \leq i, j \leq n$, we have $a_{ij} \geq 0$. We will indicate that A is positive by writing $A \geq 0$.

The set of all non-negative matrices of dimension $n \times n$ is called the *cone of non-negative matrices* and will be indicated with $\mathbb{R}^{n \times n}_+$.
We can define in similar way for $A \in \mathbb{R}^{m \times n}$ to be positive or non-negative. The space of all non-negative $m \times n$ matrices will be denoted by $\mathbb{R}^{m \times n}_+$.

Remark 2.3 A is a non-negative matrix if and only if $\forall\, x \in \mathbb{R}^n_+$, the product Ax is still a non-negative vector.

The definitions 1 and 2 can be used to state that a matrix (or a vector) is pointwise greater than 0. Similarly we can introduce an ordering on $\mathbb{R}^{n\times n}$ to state that a matrix is pointwise greater than another.

Definition 2.4 (**Partial ordering**) Let $A, B \in \mathbb{R}^{n\times n}$, we say that

- $A > B$ if the matrix $A - B$ is positive, i.e. $A - B > 0$.
- $A \geq B$ if the matrix $A - B$ is non-negative, i.e. $A - B \geq 0$.

A useful observation is that $(\mathbb{R}^{n\times n}, \geq)$ is a partially ordered set: the relation $\geq$ is reflexive, transitive and antisymmetric.

Definition 2.5 (**Absolute value**)
Let $A \in \mathbb{C}^{m\times n}$, we will denote by $|A|$ the matrix that has as i, j^{th} element $|a_{ij}|$.
Similarly we can define the absolute value of a vector $|v|$.
By this definition, we have $|A| \in \mathbb{R}_+^{m\times n}$ and $|v| \in R_+^n$ for every $A \in \mathbb{C}^{m\times n}$ and $v \in \mathbb{C}^m$.

Proposition 2.6 Properties of the ordering $\geq$:

1. If $A_i \geq B_i$ are matrices, then $\sum_{i=1}^{m} A_i \geq \sum_{i=1}^{m} B_i$
2. If $A \geq B$ are matrices and $c \in R_+$ is a scalar, then $cA \geq cB$
3. If $A \geq B$ and $C \geq 0$ are matrices, then $AC \geq BC$ and $CA \geq CB$ if the products are defined
4. If $A_k \geq B_k$ and there exist a pointwise limit $\lim_k A_k = A$ and $\lim_k B_k = B$, then $A \geq B$
5. If $A \in \mathbb{C}^{m\times n}$, then $|A| \geq 0$ and $|A| = 0$ if and only if $A = 0$
6. If $A \in \mathbb{C}^{m\times n}$ and $\gamma \in \mathbb{C}$, $|\gamma A| = |\gamma|\,|A|$
7. If $A, B \in \mathbb{C}^{m\times n}$ then $|A + B| \leq |A| + |B|$.
8. If $A, B \in \mathbb{C}^{n\times n}$ then $|A \cdot B| \leq |A| \cdot |B|$.
   This is true also for every couple of matrices such that $A \cdot B$ is defined, i.e. $A \in \mathbb{C}^{n\times k}$ and $B \in \mathbb{C}^{k\times m}$.
9. If $A \in \mathbb{C}^{n\times n}$, and $k \in \mathbb{N}$, then $|A^k| \leq |A|^k$.

Remark 2.7 There are a couple of simple but very useful particular cases of property 3 and 8:

- If $A \geq 0$ is a matrix and $v \geq 0$ is a vector, then $Av \geq 0$.
- If $A > 0$ is a matrix, $v \geq 0$ is a vector, $v \neq 0$, then $Av > 0$.
- Let $A \in \mathbb{C}^{n\times n}$ and $x \in \mathbb{C}^n$. Then $|\,A\,x\,| \;\leq\; |A| \cdot |x|$

PROPOSITION 2.8 Let $A \in \mathbb{R}^{n\times n}$. Then:

1. $A \geq 0 \iff \forall\, v \in \mathbb{R}^n$, with $v > 0$, we have $Av \geq 0$.

2. $A > 0 \iff \forall\, v \in \mathbb{R}^n$, with $v \geq 0$, $v \neq 0$, we have $Av > 0$.

Proof. The $\Rightarrow$ implications follow easily from the above properties.
1) $\Leftarrow$ If there were $a_{i\,j} < 0$, then choosing v with $v_j = 1$ and other components $k \neq j, v_k = \varepsilon$ we have $[Av]_i = a_{i\,j}v_j + \sum_{k\neq j} a_{i\,k}\varepsilon \leq -a_{i\,j} + (n-1)\varepsilon \cdot \max_{k\neq j} |a_{i\,k}|$ which is negative for sufficiently small ε.
2) $\Leftarrow$ If there were $a_{i\,j} \leq 0$, it is sufficient to choose $v = e_j$, then $[Av]_i = a_{i\,j}v_j + 0 = a_{i\,j} \leq 0$. $\square$

2.2 Spectral radius bounds

THEOREM 2.9 (**Lappo-Danilevsky**)
Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{R}_+^{n\times n}$ such that $|A| \leq B$. Then $\rho(A) \leq \rho(|A|) \leq \rho(B)$

Proof. It is sufficent to prove that for all matrices A_1, B_1 that satisfy the hyphothesis of the theorem, it's true that $\rho(A_1) \leq \rho(B_1)$.
Then taking $A_1 = A$ and $B_1 = |A|$ we obtain $\rho(A) \leq \rho(|A|)$, and by taking $A_1 = |A|$ and $B_1 = B$ we obtain $\rho(|A|) \leq \rho(B)$.

Let's prove the claim by contradiction, so assume $\rho(B) < \rho(A)$. Then there exist a scalar $\gamma \in \mathbb{R}_+$ such that

$$\gamma\rho(B) < 1 < \gamma\rho(A)$$

Which is the same as

$$\rho(\gamma B) < 1 < \rho(\gamma A)$$

This means that the $\lim_k (\gamma B)^k = 0$, while the sequence of matrices $\{(\gamma A)^k\}_{k\in\mathbb{N}}$ does not converge to any matrix (if it were to converge, then the spectral radius has to converge to a real number). On the other hand, we have $0 \leq |A| \leq B$, so $0 \leq |\gamma A| \leq \gamma B$ and $0 \leq |\gamma A|^k \leq (\gamma B)^k$ for all $k \in \mathbb{N}$. Since $(\gamma B)^k$ converges to the 0 matrix, it follows that also $|\gamma A|^k$ and $(\gamma A)^k$ has to converge to the 0 matrix, which is a contradiction. $\square$

Alternative proof. For every $x \in \mathbb{C}^n$, we have:

$$|A \cdot x| \leq |A| \cdot |x| \leq B \cdot |x|$$

Let $||\cdot||$ be an induced matrix norm from the vector norm $||\cdot||_v$ (for example, the 2-norm). Then:

$$||\, Ax\,||_v \leq ||\,|A| \cdot |x|\,||_v \leq ||\,B \cdot |x|\,||_v$$

By taking the sup on all $x \in \mathbb{C}^n$ of unitary norm, $||x||_v = 1$, we obtain the matrix norm:

$$|| \, A \, || \leq || \, |A| \, || \leq || \, B \, ||$$

Since $A \leq |A| \leq B$, we have $A^k \leq |A|^k \leq B^k$ for every integer $k > 0$; then applying the above method gives us

$$|| \, A^k \, || \leq || \, |A|^k \, || \leq || \, B^k \, ||$$

$$|| \, A^k \, ||^{\frac{1}{k}} \leq || \, |A|^k \, ||^{\frac{1}{k}} \leq || \, B^k \, ||^{\frac{1}{k}}$$

We can take the limit for $k \to \infty$ and apply Gelfand's theorem to obtain:

$$\rho(A) \leq \rho(|A|) \leq \rho(B)$$

□

THEOREM 2.10 Let $A \in \mathbb{R}_+^{n \times n}$, let $R_i = \sum_{j=1}^{n} a_{i\,j}$ be the sum of entries of the i^{th}-row and $C_j = \sum_{i=1}^{n} a_{i\,j}$ be the sum of the j^{th}-column.Then:

$$\min_{1 \leq i \leq n} R_i \;\leq\; \rho(A) \;\leq\; \max_{1 \leq i \leq n} R_i$$

$$\min_{1 \leq j \leq n} C_j \;\leq\; \rho(A) \;\leq\; \max_{1 \leq j \leq n} C_j$$

Proof. The maximum row-sum can be expressed as the infinity norm, since A is non-negative

$$\max_{1 \leq i \leq n} R_i = ||A||_\infty$$

Furthermore, this is an induced norm, so $\rho(A) \leq ||A||_\infty$ and we have the upper bound inequality. For the column sums the argument is similar, using the 1-norm instead of the infinity norm.

$$\rho(A) \leq ||A||_1 = \max_{1 \leq j \leq n} C_j$$

For the lower bound, let $m = \min_{1 \leq i \leq n} R_i$ and $M = \max_{1 \leq i \leq n} R_i$. We can construct two matrices $B, C \geq 0$ such that the row sums of B are all m, and those of C are all M; we can take A and multiply every element in the i^{th} row by $\frac{m}{R_i} \leq 1$ for B and $\frac{M}{R_i} \geq 1$ for C, obtaining $B \leq A \leq C$. By ?.9 we have that $\rho(B) \leq \rho(A) \leq \rho(C)$. $\rho(B) \leq ||B||_\infty = m$ and $B \cdot e = m\, e$, which implies that $\rho(B) = m$; similarly we get that $\rho(C) = M$.
We have obtained that $m \leq \rho(A) \leq M$ as desired. The case with column sums is analogous.

□

COROLLARY 2.11 If $A > 0$ then $\rho(A) > 0$.

Since A is positive, all the row sums R_i are strictly greater than 0, which implies that $\rho(A)$ cannot be 0.

COROLLARY 2.12 If $\widehat{A} \in \mathbb{R}^{k\times k}$ is a principal submatrix of $A \in \mathbb{R}_+^{n\times n}$, then $\rho(\widehat{A}) \leq \rho(A)$.
Consider the matrix $B \in \mathbb{R}^{n\times n}$ which has the corresponding entries equal to $\widehat{A}$ and all the other are zeros. Then $\rho(\widehat{A}) = \rho(B)$ and apply Lappo-Danilevsky's theorem.

COROLLARY 2.13 If $A \geq 0$, then $\rho(A) \geq \max\limits_i a_{i\,i}$.

Let $x > 0$ be a vector and $D = \operatorname{diag}(x_1, \ldots, x_n)$, with all diagonal entries strictly greater than 0. Then D is non-singular, it's inverse is $D^{-1} = \operatorname{diag}(x_1^{-1}, \ldots, x_n^{-1})$. Let $\widehat{A} = D^{-1}AD$, calculating its entries we obtain $\widehat{a}_{i\,j} = \frac{x_j}{x_i}\, a_{i\,j}$.
$\widehat{A}$ is similar to A, so they have the same eigenvalues and $\rho(\widehat{A}) = \rho(A)$. We can use theorem (2.10) with $\widehat{A}$ to obtain new bounds for $\rho(A)$.
The row sums of $\widehat{A}$ are $\widehat{R} = \dfrac{1}{x_i}\left(\sum\limits_{j=1} a_{i\,j}\, x_j\right)$, so applying theorem (2.10) leads us to:

THEOREM 2.14 (**Collatz - Wielandt**) Let $A \geq 0$, for all positive vectors $x > 0$ we have:

$$\min_{1\leq i\leq n} \frac{1}{x_i}\left(\sum_{j=1} a_{i\,j}x_j\right) \;\leq\; \rho(A) \;\leq\; \max_{1\leq i\leq n} \frac{1}{x_i}\left(\sum_{j=1} a_{i\,j}x_j\right)$$

This is true if we use columns instead of rows. However, the following theorems allows us to give even better bounds.

THEOREM 2.15 (**Collatz 1**) Let $A \geq 0$ be a $n \times n$ matrix, $x > 0$ a vector and $\sigma, \tau \geq 0$ real numbers.
If $\sigma x \leq Ax \leq \tau x$ (as vectors), then $\sigma \leq \rho(A) \leq \tau$.

Proof. By the above remark, we have

$$\rho(A) \geq \min_{1\leq i\leq n} \frac{1}{x_i}\left(\sum_{j=1} a_{i\,j}x_j\right) = \min_{1\leq i\leq n} \frac{1}{x_i}[Ax]_i$$

By hypothesis $Ax \geq \sigma x$, so $[Ax]_i \geq \sigma x_i$.

$$\rho(A) \geq \min_{1\leq i\leq n} \frac{1}{x_i}[Ax]_i \geq \min_{1\leq i\leq n} \frac{1}{x_i}\sigma x_i = \sigma$$

The proof for $\rho(A) \leq \tau$ is analogous. $\square$

COROLLARY 2.16 If $A \geq 0$ and $p > 0$ is a positive vector. Suppose that p is also an eigenvector for A: $Ap = \mu p$. Then necessarily $\mu = \rho(A)$.

Proof. First we note that since $A \geq 0$ and $p > 0$, $Ap \geq 0$ is a non-negative vector. This implies that the eigenvalue μ is a real number. Since $\mu p = Ap \geq 0$, we have that also $\mu \geq 0$.
Applying theorem 2.15 with $\sigma = \tau = \mu$ and $x = p$, we obtain that $\mu \leq \rho(A) \leq \mu$, as desidered. □

THEOREM 2.17 (**Collatz 2**) Let $A \geq 0$ and suppose there is a positive vector $p > 0$ which is an eigenvector for A. Then:

$$\max_{x>0} \min_{1\leq i\leq n} \frac{1}{x_i}\left(\sum_{j=1} a_{i\,j} x_j\right) = \rho(A) = \min_{x>0} \max_{1\leq i\leq n} \frac{1}{x_i}\left(\sum_{j=1} a_{i\,j} x_j\right)$$

Where $\max\limits_{x>0}$ and $\min\limits_{x>0}$ taken for all possible values of $x \in \mathbb{R}^n_+$. They should be thought as sup and inf, but they actually obtain maximum for some x.

Proof. For the above corollary, p has to be an eigenvector relative to $\rho(A)$, so $A\,p = \rho(A)\,p$, which implies:

$$\rho(A) = \frac{1}{p_i}\left(\sum_{j=1} a_{i\,j} p_j\right) \qquad \forall\ i$$

Which means that the inequalities stated in theorem ?.15 is obtained with $x = p$, both the lower and upper. □

In this theorem it's crucial to have A with a positive eigenvector (which happens if $A > 0$ or $A \geq 0$ and A is irreducible, by Perron-Frobenius results); otherwise the min/max have to be replaced with inf/sup.

2.3 Perron's Theorem

THEOREM 2.18 (**Perron**) Let $A \in \mathbb{R}^{n\times n}_+$ be a strictly positive matrix, $A > 0$, and let $\rho = \rho(A)$ be the spectral radius of A. Then:

1. ρ is an eigenvalue of A and $\rho > 0$.

2. There exists a positive vector $p > 0$ which is an eigenvector for ρ: $A\,p = \rho\,p$. Similarly, there exists a left-eigenvector $q^\top$ for A with eigenvalue ρ and $q^\top > 0$.

3. If $x > 0$ is another positive eigenvector for A, so $Ax = \lambda x$, then necessarily $\lambda = \rho$ and $x = \alpha p$ with $\alpha > 0$ a positive number. Similarly, a positive left-eigenvector is a positive multiple of a $q^\top$.

4. ρ is a simple eigenvalue of A

5. For each $\lambda \in \sigma(A)$, $\lambda \neq \rho$, we have $|\lambda| < \rho$.

6. Let $B = \dfrac{1}{\rho} A$, so $\rho(B) = 1$. Then there exists the limit $\lim\limits_{k\to\infty} B^k = L = pq^\top$

7. For $x \geq 0$ vector, $x \neq 0$, we have $\lim_{k\to\infty} B^k x$ is a positive multiple of p.

8. For $x \geq 0$ vector, $x \neq 0$, let $\gamma_k = [A^k\, x]_1^{-1}$ be a scalar, such the vector $b_k = \gamma_k A^k\, x$ has the first entry equal to 1. Then there exists the limit $\lim_{k\to\infty} b_k = \theta p$ with $\theta = \frac{1}{p_1}$.

Proof. Notice that for small $\varepsilon > 0$, $A > \varepsilon\,\mathrm{I}$, which implies by Lappo-Danilevsky ?.9 that $\rho(A) \geq \varepsilon > 0$. We will prove the first 4 statements for $B = \dfrac{1}{\rho}A$. Since B is a positive scalar multiple of A, they have the same eigenvalues and eigenvectors (with the same sign).

Step 1. $\rho(B) = 1$, so there exist an eigenvector x, $Bx = \lambda x$ with $|\lambda| = 1$. Then

$$|x| = |\lambda x| = |Bx| \leq |B| \cdot |x| = B \cdot |x|$$

Call $y = B|x| - |x| = (B - \mathrm{I}) \cdot |x|$, the above results means $y \geq 0$. If it were $y = 0$, then $B|x| = |x|$ and so B has an eigenvector with eigenvalue 1, as desired. By contradiction, assume that $y \neq 0$; then from $y \geq 0$ and $B > 0$ we have $By > 0$.
Let $z = B|x| > 0$, then for a sufficiently small $\varepsilon > 0$ we have $By > \varepsilon z$ (observe that this is false if we require only $By \geq 0$).

$$\varepsilon z < By = B(B - \mathrm{I}) \cdot |x| = (B - \mathrm{I})B \cdot |x| = (B - \mathrm{I})z$$

So we obtained $(B - \mathrm{I})z > \varepsilon z$, or $Bz > (1 + \epsilon)z$. By theorem ?.15, this means that $1 = \rho(B) > 1 + \varepsilon$, which is absurd.
This means that necessarily $y = 0$ and $1 \in \sigma(B)$.

Step 2. We have proved that if $Bx = \lambda x$ with $|\lambda| = 1$, then $B|x| = |x|$. Surely $|x| \geq 0$, but we do not know if $|x|$ is strictly positive.
Let us look at the i^{th} entry of $Bx = \lambda x$:

$$\sum_{j=1}^{n} b_{i\,j}\, x_j = \lambda x_j \quad \Longrightarrow \quad |\sum_{j=1}^{n} b_{i\,j}\, x_j\,| = |\lambda|\,|x_j| = |x_j|$$

Also, by looking at the i^{th} entry of $B|x| = |x|$

$$\sum_{j=1}^{n} b_{i\,j}\,|x_j| = |x_j|$$

These two equations give us the following, were the last equality is true because $b_{i\,j} > 0$:

$$|\sum_{j=1}^{n} b_{i\,j}\, x_j\,| = |x_j| = \sum_{j=1}^{n} b_{i\,j}\,|x_j| = \sum_{j=1}^{n} |b_{i\,j}\, x_j|$$

The sum of absolute values of complex numbers can be equal to the absolute value of the sum only if all numbers have the same argument, i.e. there exists $\alpha \in [0, 2\pi]$ such that $b_{i\,j}x_j = b_{i\,j}\,|x_j|e^{i\alpha}$ and thus $x_j = |x_j|e^{i\alpha}$.

We want to show that $p = |x|$ is the positive eigenvector we're looking for. We already have $B\,p = p$, but $B > 0$ and $p \geq 0$, so $Bp > 0$ and thus $p > 0$.
By applying the result to $A^\top$ we find an eigenvector q such that $A^\top q = \rho q$, which means $q^\top A = \rho q$.

Also from $x = e^{i\alpha}p$ and $Bx = \lambda x$ we obtain

$$\lambda e^{i\alpha} p = \lambda x = Bx = e^{i\alpha} Bp = e^{i\alpha} p$$

Which implies that $\lambda = 1$. In other words, every eigenvalue of absolute value 1 in $\sigma(B)$ has to be equal to 1, so there are no eigenvalues of A of absolute value ρ other than ρ itself, which is point 5.

By corollary ?.16, every positive eigenvector y such that $By = \mu y$, has necessarily $\mu = 1$. Then

Step 3. We will prove that 1 is a semi-simple, and then simple, eigenvalue of B.
From $Bp = p$ it follows that $B^k p = p$ for every $k > 0$. Let $b_{i\,j}^{(k)} = [B^k]_{i\,j}$, then this equality can be expanded as:

$$\sum_{j=1}^{n} b_{i\,j}^{(k)}\, p_j = p_i \qquad \forall\, k > 0, \;\; \forall i$$

This means that $b_{i\,j}^{(k)}\, p_j < p_i$ and $b_{i\,j}^{(k)} < \frac{p_i}{p_j}$ for every i, j, k, which is well defined since $p > 0$. So the entries of B^k are bounded by above by $\max\limits_{i,j} \frac{p_i}{p_j}$.
Let $J = XBX^{-1}$ be the Jordan normal form of B, and suppose that there is a Jordan block for eigenvalue $\lambda = 1$ of size $d \geq 2$. Then:

$$J = \left(\begin{array}{ccccc|c} 1 & 1 & & & & \\ & 1 & 1 & & & \\ & & 1 & \ddots & & \\ & & & \ddots & 1 & \\ & & & & 1 & \\ \hline & & & & & \widehat{B} \end{array}\right) \qquad J^k = \left(\begin{array}{ccccc|c} 1 & k & \binom{k}{2} & \cdots & \binom{k}{d} & \\ & 1 & k & & & \\ & & 1 & \ddots & \vdots & \\ & & & \ddots & k & \\ & & & & 1 & \\ \hline & & & & & \widehat{B}^k \end{array}\right)$$

This shows that the entries of J^k can be arbitrarily large, but $J^k = XB^kX^{-1}$ and the entries of B^k are bounded, so we have a contradiction and $\lambda = 1$ must be a semi-simple eigenvalue.

To prove that is simple, assume by contradiction that there exists $x, y \in \mathrm{Ker}(B - \mathrm{I})$ which are linearly independent.Then it is possible to choose $\alpha, \beta \neq 0$ such that $z = \alpha x + \beta y$ has one entry equal to zero.
We have that $Bz = z$ and $B\,|z| \geq |Bz| = |z|$. The proof of Step 1 gives us that there can't inequality and necessarily $B\,|z| = |z|$. Since $B > 0$ and $|z| \geq 0$, it means that $B\,|z| > 0$, which is absurd because $|z|$ has a zero entry.

We've proved point 4). Point 3) follows easily: let p be the positive eigenvector relative to ρ and x another generic eigenvector. For corollary ?.16 x must have eigenvalue ρ, but ρ is simple so x has to be a multiple of p.

Step 4. We will prove point 6). Since 1 is a simple eigenvalue of B, the Jordan normal form of B is:

$$J = XBX^{-1} = \left(\begin{array}{c|c} 1 & \\ \hline & \widehat{J} \end{array}\right)$$

Where $\widehat{J} \in \mathbb{C}^{(n-1)\times(n-1)}$ and $\rho(\widehat{J}) < 1$. This means that the limit of J^k and A^k is

$$\lim_{k\to\infty} J^k = \left(\begin{array}{c|c} 1 & \\ \hline & 0 \end{array}\right) \qquad L = \lim_{k\to\infty} B^k = X^{-1} \left(\begin{array}{c|c} 1 & \\ \hline & 0 \end{array}\right) X$$

Which is a matrix of rank one. Furthermore L is the product of the first column of X^{-1} and the first row of X, which are eigenvectors p and $q^\top$ respectively.
This means that $L = pq^\top > 0$.

The limit $\lim_{k\to\infty} B^k x = Lx = pq^\top x = (q^\top x)p$ is a multiple of p by the scalar $q^\top x$, as desired.

Step 5. We will prove point 8).
We know that $\lim_{k\to\infty} \left(\frac{1}{\rho}A\right)^k x = \lim_{k\to\infty} B^k x = Lx = \alpha p$ where α is a positive scalar.
From $\gamma_k A^k x = b_k$ we have $\left(\frac{1}{\rho}A\right)^k x = \frac{1}{\gamma_k \rho^k} b_k$; since $b_k(1) = 1$, by taking the limit we see that $\lim_{k\to\infty} \frac{1}{\gamma_k \rho^k} = \alpha p(1)$. This implies:

$$\lim_{k\to\infty} b_k = \lim_{k\to\infty} \gamma_k \rho^k B^k x = \frac{1}{\alpha\, p(1)} \alpha p = \frac{1}{p(1)}\, p$$

□

We will present another proof of Perron's theorem, based on Brouwer's fixed-point theorem.

Theorem 2.19 (**Brouwer's fixed-point**) Let $S \subseteq R^n$ be convex and compact, and let $f : S \to S$ be a continous map. Then there exists a fixed point $x \in S$, i.e. such that $f(x) = x$.
DA COMPLETARE Apply this theorem to the set of all non-negative vectors $\mathbb{R}^n_+ \subseteq \mathbb{R}^n$.

2.4 Non-negative Matrices

Passiamo ora a matrici non negative $A \geq 0$

Theorem 2.20 (**Frobenius 1**) Let $A \geq 0$ be a non-negative matrix, and $\rho = \rho(A)$ it's spectral radius. Then:

1. ρ is an eigenvalue of A
2. There exists right and left **non-negative** eigenvectors for A.

Proof. DA COMPLETARE Approximate from above with positive matrices. □

2.5 Graphs and Frobenius

Definition 2.21 (**reducible and irreducible matrix**)
$A \in \mathbb{R}^{n\times n}$ is called *reducible* if there exists a permutation matrix P that transforms A in a block triangular matrix: $PAP^\top = \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix}$.
If A is not reducible then it's called *irreducible*

Proposition 2.22 Let $A > 0$, $A \in \mathbb{R}_+^{n\times n}$. Then A is irreducibile $\iff$ $(\mathrm{I}_n + A)^{n-1} > 0$.

Proof. $\Leftarrow$ If A were reducible, there would exists a permutation P such that $PAP^\top$ is block triangular. Write:

$$
\begin{aligned}
P(\mathrm{I}+A)^{n-1}P^\top =& (P(\mathrm{I}+A)P^\top)^{n-1} = (\mathrm{I}+(PAP^\top))^{n-1} = \\
=& \mathrm{I} + \binom{n-1}{1}(PAP^\top) + \binom{n-1}{2}(PAP^\top)^2 + \ldots + (PAP^\top)^{n-1}
\end{aligned}
$$

The RHS is the sum of block triangular matrices with the same zero block, this means that $P(\mathrm{I}+A)^{n-1}P^\top$ has a zero block and thus $(\mathrm{I}+A)^{n-1}$ cannot be a positive matrix.

$\Rightarrow$ Basta dimostrare che $(\mathrm{I}+A)^{n-1}x > 0$ per ogni $x \geq 0$. DA COMPLETARE □

Definition 2.23 (**Associated graph**) For every $A \in \mathbb{C}^{n\times n}$, we can associate a directed graph denoted by $\mathcal{G}(A)$. It has $V = \{1, 2, \ldots n\}$ as set of vertices, and there is an edge $E(i,j)$ between vertices i and j if and only if $A_{ij} \neq 0$.
A is called the *adjacency matrix* of $\mathcal{G}(A)$.

Proposition 2.24 Let $A \geq 0$ and $\mathcal{G}(A)$ be its associated graph. For every pair of vertices i, j, there is a path from i to j of length k if and only if $[A^k]_{ij} > 0$.
If we construct $\mathcal{G}(A)$ as a weighted graph, then $[A^k]_{ij}$ is the sum of all weighted paths from i to j of length k.

Proposition 2.25 $A \in \mathbb{C}^{n\times n}$ is irreducible $\iff$ it's associated graph $\mathcal{G}(A)$ is strongly connected

Theorem 2.26 (**Frobenius 2**) Let $A \geq 0$ be an irreducible matrix and let $\rho = \rho(A)$ be its spectral radius. Then:

1. $\rho > 0$ and ρ is an eigenvalue of A

2. ρ is a simple eigenvalue

3. There exists a **positive** vector $p > 0$ which is an eigenvector for ρ: $A\,p = \rho\,p$. Similarly for the left eigenvalue.

Proof. DA COMPLETARE □

PROPOSITION 2.27 Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{R}^{n\times n}$, $B \geq 0$ be matrices such that $|A| \leq B$. Assume also that B is irreducible and $\rho(A) = \rho(B)$. Then necessarily $|A| = B$.

Proof. By Lappo-Danilevsky's theorem ?.9, we already know that necessarily $\rho(A) \leq \rho(|A|) \leq \rho(B)$. So now assume that $\rho(A) = \rho(B) = \rho$.
Let $\mu \in \mathbb{C}$ be an eigenvalue of A of maximum absolute value, i.e. $|mu| = \rho$, and let x be the corresponding eigenvector, $Ax = \mu x$.
By taking the absolute value of this expression, we obtain:

$$\rho\,|x| = |\rho\,x| = |\mu x| = |Ax| \leq |A| \cdot |x| \leq B \cdot |x|$$

Since B is irreducible, by Frobenius theorem there exists a left positive eigenvector z, such that $z^\top B = \rho z^\top$. Multiplying this by $|x|$ we obtain: $z^\top B|x| = \rho z^\top |x|$, so $z^\top\left(B|x| - \rho|x|\right) = 0$. But since $z^\top > 0$, necessarily the vector $B|x| - \rho|x|$ is zero, so $B|x| = \rho|x|$. Then $|x|$ is a right Perron eigenvector for B, so $|x| > 0$.
Since $\rho|x| \leq |A|\,|x| \leq B|x| = \rho|x|$, we obtain that $B|x| = |A|\,|x|$, so $(B - |A|)|x| = 0$. Since $B - |A| \geq 0$ and $|x| > 0$, we have that necessarily $B - |A| = 0$, so $B = |A|$ as desired. □

COROLLARY 2.28 Let $A, B \geq 0$ be matrices such that $A \leq B$, $A \neq B$ and B is irreducible. Then $\rho(A) < \rho(B)$.
Esempio che irriducibilità di B è necessaria. DA COMPLETARE

DEFINITION 2.29 (**index of imprimitivitiy**) Let $\mathcal{G}$ be a directed, strongly connected graph. The *index of imprimitivity* of $\mathcal{G}$ is $k = \gcd\{\text{Length of closed walk in } \mathcal{G}\}$.
If $k = 1$, then $\mathcal{G}$ is called *primitive*.

THEOREM 2.30 Let $\mathcal{G} = (V, E)$ be a strongly connected graph, $|V| = n$ and with index of imprimitivity k. Then:

1. $\forall\, v \in V$, k is the gcd of the lengths of all closed walks starting from v.

2. $\forall\, v, w \in V$, if there are two walks from v to w of length L_1, L_2, then $L_1 \equiv L_2 \pmod{k}$.

3. V can be partitioned in k disjoint sets $V_1, V_2, \ldots, V_k$, such that every arc starting from a vertex in V_i ends in a vertex in V_{i+1}, with $V_{k+1} = V_1$.

4. Take $v_i \in V_i$, $v_j \in V_j$ and a path $v_i \to v_j$ of length L. Then $L \equiv j - i \pmod{k}$.

The vertex sets V_i are called the *imprimitivity sets* of $\mathcal{G}$.
Forma normale di Frobenius (diagonale a blocchi shiftata) per matrici imprimitive.DA COMPLETARE

PROPOSITION 2.31 Let $A \in \mathbb{C}^{m\times l}$, $B \in \mathbb{C}^{l\times m}$, so $AB \in \mathbb{C}^{l\times l}$ and $BA \in \mathbb{C}^{m\times m}$. Then AB and BA have the same non-zero eigenvalues with the same multiplicity.

Proof. The following identity holds:

$$\left(\begin{array}{c|c} \lambda\,\mathrm{I}_l - AB & A \\ \hline 0 & \lambda\,\mathrm{I}_m \end{array}\right)\left(\begin{array}{c|c} \mathrm{I}_l & 0 \\ \hline B & \mathrm{I}_m \end{array}\right) = \left(\begin{array}{c|c} \lambda\,\mathrm{I}_l & A \\ \hline \lambda B & \lambda\,\mathrm{I}_m \end{array}\right) = \left(\begin{array}{c|c} \mathrm{I}_l & 0 \\ \hline B & \mathrm{I}_m \end{array}\right)\left(\begin{array}{c|c} \lambda\,\mathrm{I}_l & A \\ \hline 0 & \lambda\,\mathrm{I}_m - BA \end{array}\right)$$

By equating the determinants and using Binet's formula, we obtain:

$$\det(\lambda\,\mathrm{I}_l - AB)\cdot\lambda^m = \lambda^l\cdot\det(\lambda\,\mathrm{I}_m - BA)$$

This means that AB and BA have the same characteristic polynomial up to a λ^{m-l} factor, and thus have the same eigenvalues with same multiplicity. □

Proposizione sullo shift di autovettori per radice di unità. DA COMPLETARE

DEFINITION 2.32 (**Primitive matrix**) Let $A \geq 0$ be an irreducible matrix. A is *primitive* if its associated graph $\mathcal{G}(A)$ has index of imprimitivity equal to 1.
In his original work, Frobenius defined primitive matrices to be the ones that don't have other eigenvalues of absolute value equal to $\rho(A)$ except ρ itself.

THEOREM 2.33 (**Schur**) Let $S \subseteq \mathbb{N}$, $S \neq \varnothing$, such that S is closed under addition. Let d be the gcd of all elements of S. Then there exists $N \in \mathbb{N}$ such that $td \in S\ \forall\, t \geq N$.

Proof. Without loss of generality we can assume that $d = 1$: we can create a new set with $\gcd = 1$ by diving all elements of S by d.
DA COMPLETARE □

The minimum N that satisfies this condition is called the *Schur-Frobenius index* of S.

PROPOSITION 2.34 Let $\mathcal{G} = (V, E)$ be a strongly connected graph and let k be its index of imprimitivity. Let $V_1, \ldots V_k$ be the imprimitivity partition of V. Then there exists $N \in \mathbb{N}$ such that for every vertices $v_i \in V_i, v_j \in V_j$ and $\forall\, t \geq N$, there exist a walk $v_i \to v_j$ of length $L = (j - i) + tk$.

Proof. DA COMPLETARE □

COROLLARY 2.35 If $\mathcal{G}$ is an imprimitive graph, then $k = 1$ and there exists $N \in \mathbb{N}$ such that for every vertices $v, w \in V$, there exists a walk $v \to w$ of length L for every $L \geq N$.

Theorem 2.36 (**Equivalent formulation of imprimitive matrix, Frobenius 1912**) Let $A \geq 0$ be an irreducible matrix. Then A is primitive $\iff$ there exists m such that $A^m > 0$. Observe that in this case, for every $l \geq m$ we have $A^l > 0$.

Proof. DA COMPLETARE □

Teorema su $A \geq 0$ irriducible ma periodica di k, che tutti i suoi blocchi di A^k sono primitivi.

Theorem 2.37 (**Frobenius 3**) If $A \geq 0$ is an irreducible, primitive matrix, then $\rho = \rho(A)$ it's the only eigenvalue λ such that $|\lambda| = \rho$. Also, there exists the limit of $\lim_{m\to\infty} (\frac{1}{\rho})A)^m$
Combining with other Frobenius results for non-negative matrices, we obtain that for a non-negative, primitive matrix, all properties of original Perron's theorem hold.

Proof. Sugli appunti è spezzata in due parti DA COMPLETARE

□

Per le matrici imprimitive di indice k allora lo spettro è simmetrico per rotazioni nel piano complesso di $\omega = \frac{2\pi}{k}$, contate con molteplicità tranne che per l'autovalore 0. Inoltre ci sono k autovalori di modulo massimo e sono esattamente $\rho(A), \omega\rho(A), \omega^2\rho(A), \ldots \omega^{k-1}\rho(A)$.

Theorem 2.38 (**Wielandt**) Let $A \geq 0$ be a primitive matrix of size $n \times n$, then surely $A^m > 0$ where $m = n^2 - 2n + 2 = (n-1)^2 + 1$. This is also the optimal exponent for which it holds.
This results has been strengthened by J.Shan (?), as:

Theorem 2.39 (**Shan, LAA 1995**) Linear algebra applied
Let $A \geq 0$ be a matrix, such that the degree of the minimal polynomial of A is m. Then A is primitive $\iff$ $A^{(m-1)^2+1} > 0$.
There are other similar results, in which the exponents is bounded depending on the number of diagonal entries $a_{i\,i} \neq 0$.
Parte sull'adjugate-adjoint e dimostrazione di Frobenius del suo teorema.

Theorem 2.40 Let $A \in \mathbb{C}^{n\times n}$, $\phi(\lambda) = \det(\lambda\,\mathrm{I} - A)$ be its characteristic polynomial and $\phi_i(\lambda) = \det(\lambda\,\mathrm{I} - A \setminus \{i, i\})$ be the characteristic polynomial of $A \setminus \{i, i\}$, the matrix with the i^{th} row and column removed. Then

$$\phi'(\lambda) = \sum_{i=1}^{n} \phi_i(\lambda)$$

Matrici riducibili, e forma normale per esse. Grafo condensato.

Theorem 2.41 (**Gantmacher**) Let $A \geq 0$ written in Frobenius normal form. with $A_1, \ldots, A_g$ isolated components and $A_{g+1}, \ldots, A_t$ the non-isolated ones. Let $\rho = \rho(A)$ be the spectral radius of A. Then the Perron eigenvector z is strictly positive if and only if both the following conditions apply:

1. ρ is an eigenvalue of every block $A_1, \dots A_g$.
2. ρ is NOT an eigenvalue of any block $A_{g+1}, \dots, A_t$.

Achtung! Questo teorema funziona per autovalori dx ma non sx, va sistemato.

Proof. Note that $\sigma(A) = \bigcup_{i=1}^{t} \sigma(A_i)$, so $\rho(A_i) \le \rho \; \forall i$.

$\Rightarrow$ Let $z > 0$ be the Perron eigenvector of A. Then $Az = \rho z$ can be written as:

$$\begin{cases} A_i\, z_i = \rho z_i & \forall\, 1 \le i \le g \\ \sum_{k=1}^{j-1} A_{j\,k}\, z_k + A_j\, z_j = \rho z_j & \forall\, g+1 \le j \le t \end{cases} \tag{1}$$

The first g equations tell us that ρ is an eigenvalue for $A_1, \dots, A_g$.
From the other we have that $A_j\, z_j \le \rho z_j$ but surely $A_j\, z_j \ne \rho z_j$, since $A_{j\,k}$ can't be 0 for all k. We obtain that $\rho \ge \frac{[A_j\, z_j]_i}{[z_j]_i}$ for all i, or also $\rho \ge \max_i \frac{[A_j\, z_j]_i}{[z_j]_i}$. For theorem ?.14, we have that $\max_i \frac{[A_j\, z_j]_i}{[z_j]_i} \ge \rho(A_j)$.
Suppose by contradiction that $\rho(A_j) = \rho$, this means that the two above inequalities are equalities, so there exists i such that $[A_j\, z_j]_i = \rho(A_j)\,[z_j]_i$. We want to prove that z_j is an eigenvector for A_j.
Recall the proof of ?.14, we used that $\rho(B) \le \max_h \{R_h : \text{sum of } k^{\text{th}} \text{ row of } B\}$ where $B = D^{-1}AD$ and $D = \operatorname{diag}(z_1, \dots, z_n)$. If there was an index k such that $[A_j\, z_j]_k < \rho(A_j)\,[z_j]_k$, then the corresponding row sum $R_k < R_i$. There would exists a matrix $B' \ge B$ with all row sums equal to R_i, and it would be irreducible since A and thus B are irreducible. By ?.27, this would mean that $R_i = \rho(B) < \rho(B') = R_i$, which is a contradiction.
This means that the for all k, we have $[A_j\, z_j]_k = \rho(A_j)\,[z_j]_k$, i.e. z_j is an eigenvalue for A_j relative to $\rho(A_j) = \rho$. This means that $A_j\, z_j = \rho z_j$ and this is a contradiction because of equation 1.

$\Leftarrow$ Assume that ρ is an eigenvalue for $A_1, \dots, A_g$ but not for $A_{g+1}, \dots, A_t$. Then we choose the relative eigenvectors $z_1, \dots, z_g$ such that $A_i = \rho z_i$.
For the other vectors, we need to solve the equation:

$$(\rho\,\mathrm{I} - A_j)\, z_j = \sum_{k=1}^{j-1} A_{j\,k}\, z_k \qquad \forall\, g+1 \le j \le t$$

$$z_j = \frac{1}{\rho}\left(\mathrm{I} - \frac{1}{\rho}A_j\right)^{-1}\left(\sum_{k=1}^{j-1} A_{j\,k}\, z_k\right) \qquad \forall\, g+1 \le j \le t$$

Which is a positive vector, because A_j is irreducible and thus the following matrix is strictly positive:

$$\left(\mathrm{I} - \frac{1}{\rho}A_j\right)^{-1} = \mathrm{I} + \frac{1}{\rho}A_j + \left(\frac{1}{\rho}A_j\right)^2 + \left(\frac{1}{\rho}A_j\right)^3 + \dots$$

$\square$

THEOREM 2.42 Let $A \geq 0$, $\rho = \rho(A)$. The following are equivalent:

1. A has both right p and left $q^\top$ eigenvectors for ρ, with $p, q > 0$.
2. The Frobenius Normal form of A is block-diagonal, and every block is irreducible.

REMARK 2.43 Let $A \geq 0$, $\rho = \rho(A)$. If ρ is a simple eigenvalue and A has both right and left positive eigenvectors for ρ, then A is irreducible.
Ultimo teorema

3 Related Matrix Classes

3.1 Z and M matrices

DEFINITION 3.1 (**essentially positive and non-negative matrix**) Let $A \in \mathbb{R}^{n\times n}$, then:

- A is *essentially non-negative* if the off-diagonal entries are non-negative: $\forall\, i \neq j \;\; a_{ij} \geq 0$.
- A is *essentially positive* if is essentially non-negative and also irreducible.

Observe that if $A \geq 0$ implies essential non-negativity and also $A > 0$ implies essential positivity. Also, if A has all non diagonal entries strictly greater than 0, then A is essentially positive; however, we chose to include a larger set of matrices in this definition.

PROPOSITION 3.2 Let A be an essentially non-negative matrix. Then there exists a real eigenvalue λ_* of A such that for every eigenvalue $\lambda \in \sigma(A)$, $\lambda_* \geq \mathfrak{Re}(\lambda)$.

Proof. Since A is essentially non-negative, it's off-diagonal entries are non-negative, but the diagonal ones could be less than 0. Take $\alpha = \max\limits_{1\leq i\leq n} |a_{i\,i}|$, then surely $A + \alpha\,\mathrm{I}$ is a non-negative matrix. Then for Frobenius theorem we have $\rho(A + \alpha\,\mathrm{I}) = \rho$ is an eigenvalue of $A + \alpha\,\mathrm{I}$.
Every eigenvalue $\lambda_{A+\alpha\,\mathrm{I}}$ of $A+\alpha\,\mathrm{I}$ corresponds to an eigenvalue λ_A of A such that $\lambda_{A+\alpha\,\mathrm{I}} = \lambda_A+\alpha$. So using the eigenvalue ρ, we obtain $\rho = \lambda_* + \alpha$; we claim that λ_* is the eigenvalue we're searching for.
Take λ_A and $\lambda_{A+\alpha\,\mathrm{I}}$ two corresponding eigenvalues, then $\rho \geq |\lambda_{A+\alpha\,\mathrm{I}}|$, which implies:

$$\lambda_* + \alpha = \rho \geq |\lambda_{A+\alpha\,\mathrm{I}}| \geq \mathfrak{Re}(\lambda_{A+\alpha\,\mathrm{I}}) = \mathfrak{Re}(\lambda_A + \alpha) = \mathfrak{Re}(\lambda_A) + \alpha$$

From which we obtain $\lambda_* \geq \mathfrak{Re}(\lambda_A)$, as desired. $\square$

Observe that transforming A to $A+\alpha\,\mathrm{I}$ corresponds to shift the eigenvalues in the complex plane by α. λ_* is the "rightmost" eigenvalue, so shifting by α can make it the farthest point from the origin and thus the spectral radius $\rho(A+\alpha\,\mathrm{I})$, but λ_* itself doesn't have to be the spectral radius of A. As an example take $A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$, it's eigenvalues are $\lambda_A = \{0, \frac{-3+\sqrt{3}i}{2}, \frac{-3-\sqrt{3}i}{2}\}$, we have $\lambda_*(A) = 0$ but $\rho(A) = 3$.

THEOREM 3.3 (**Varga**) Let $A \in \mathbb{R}^{n\times n}$. A is essentially positive $\iff$ $\forall\, t > 0$ we have $e^{tA} > 0$.

Proof. $\Rightarrow$ Assume that A is essentially positive. Then it's irreducible and the non diagonal entries are non-negative. Take $\alpha_0 = \max_{1 \geq i \geq n} |a_{i\,i}|$, then for $\alpha > \alpha_0$ the matrix $A + \alpha\,\mathrm{I}$ is non-negative, irreducible; its diagonal entries are all positive, so $A + \alpha\,\mathrm{I}$ is also primitive.
Since it's primitive, it follows that $(A + \alpha\,\mathrm{I})^k$ is non-negative for $k \geq 0$ and is positive for k sufficiently large. We can write $e^{t(A+\alpha\,\mathrm{I})} = \mathrm{I} + t(A+\alpha\,\mathrm{I}) + \frac{t^2}{2}(A+\alpha\,\mathrm{I})^2 + \frac{t^3}{3!}(A+\alpha\,\mathrm{I})^3 + \ldots$, which means that $e^{t(A+\alpha\,\mathrm{I})}$ is positive $\forall\, t > 0$ and $\forall\, \alpha > \alpha_0$.
Observe that $e^{tA} = e^{t(A+\alpha\,\mathrm{I})} \cdot e^{-t\alpha}$ where $e^{-t\alpha}$ is a positive real number for all t and α. This means that also e^{tA} is a positive matrix.

$\Leftarrow$ We need to prove that A is essentially positive, i.e is irreducible and with non-negative off-diagonal entries. Write $e^{tA} = \mathrm{I} + tA + \frac{t^2}{2}A^2 + \frac{t^3}{3!}A^3 + \ldots$
A is irreducibile. Suppose not, then exists a permutation matrix P such that $PAP^\top$ is block upper triangular; also $(PAP^\top)^k = PA^kP^\top$ and $Pe^{tA}P^\top = e^{tPAP^\top}$ is the sum of block upper triangular matrices (with the same zero block), so $Pe^{tA}P^\top$ is block upper triangular and cannot be a positive matrix. Hence a contradiction.
If there exist $a_{i\,j} < 0$ with $i \neq j$; for small t we have $e^{tA} = \mathrm{I} + tA + \mathcal{O}(t^2)$, so the $[e^{tA}]_{i\,j} = ta_{i\,j} + \mathcal{O}(t^2)$. This means that for t small enough $[e^{tA}]_{i\,j} < 0$, which is a contradiction. This means that A is essentially non-negative. $\square$

The definition of essentially positive is useful in the solution of linear systems $x'(t) = Ax(t)$ with initial condition $x(0) = x_0$. The solution is $x(t) = e^{tA}x_0$; if A is essentially positive and $x_0 \geq 0$, we have $x(t) \geq 0 \ \ \forall\, t > 0$. This has application, for example, in the heat equation.

Definition 3.4 (**Z-matrix**) $A \in \mathbb{R}^{n\times n}$ is called a *Z-matrix* if it's off-diagonal entries are non positive: $\forall\, i \neq j$ we have $a_{i\,j} \leq 0$.

Remark 3.5 A is a Z-matrix if and only if $-A$ is an essentially non-negative matrix.

Definition 3.6 (**monotone matrix**) $A \in \mathbb{R}^{n\times n}$ is *monotone* if A is non-singular and it's inverse A^{-1} is a non-negative matrix, i.e. $A^{-1} \geq 0$.

Theorem 3.7 Let $A \in \mathbb{R}^{n\times n}$ be a non-singular matrix. Then A is monotone if and only if the following condition holds: $\forall\, x, y \in \mathbb{R}^n$, $Ax \geq Ay$ implies $x \geq y$.

Proof. $\Rightarrow$ Assume A monotone, then $A^{-1} \geq 0$ and multiplying both sides of $Ax \geq Ay$ we get $A^{-1}Ax \geq A^{-1}Ay$, so $x \geq y$.

$\Leftarrow$ Write A^{-1} columnwise, let c_j be the j^{th} column of A^{-1}. Looking at the j^{th} columnt of the product $A \cdot A^{-1} = \mathrm{I}$ we obtain $Ac_j = e_j$, which has all entries 0 except the j^{th}, so is a non-negative vector. So we have $0 = A0 \leq Ac_j$, and if we apply the hypotesis we obtain that $0 \leq c_j$. This is true for all columns c_j, so it means that A^{-1} is a non-negative matrix. $\square$

Definition 3.8 (**M-matrix**) $A \in \mathbb{R}^{n\times n}$ is a *M-matrix* if it can be written in the following form: $A = r\,\mathrm{I} - B$, where $B \geq 0$ is a non-negative matrix and $r \geq \rho(B)$.
If we have strict inequality $r > \rho(B)$, then A is called *non-singular M-matrix.*

Remark 3.9 If A is a Z-matrix, for each $r \in \mathbb{R}$ it can be written in the form $A = rI - B$, with $a_{i\,i} = r - b_{i\,i}$. The off-diagonal entries of B are $b_{i\,j} = -a_{i,j} \geq 0$, and the diagonal entries $b_{i\,i} = r - a_{i\,i}$ are non-negative if $r \geq a_{i\,i}\ \forall\, i$. So a Z-matrix can be always written in the form $rI - B$ with $B \geq 0$, but it could be that $\rho(B) < r$, so A could not be a M-matrix. For example take $A = -\mathrm{I}$; write $A = r\,\mathrm{I} - B$, then $B = (r+1)\,\mathrm{I}$ and $\rho(B) = r + 1 > r$ for all $r \in \mathbb{R}$.

Remark 3.10 A is a (non-singular) M-matrix if and only if $A^\top$ is a (non-singular) M-matrix.

Theorem 3.11 (**ZM 1**) Let $A \in \mathbb{R}^{n\times n}$ be a Z-matrix. Then A is a non-singular M-matrix $\iff$ A is monotone.

Proof. $\Rightarrow$ Since A is a non-singular M-matrix, it can be written as $A = r\,\mathrm{I} - B$. It's inverse is $A^{-1} = (\mathrm{I} - rB)^{-1} = \frac{1}{r}(\mathrm{I} - \frac{1}{r}B)^{-1}$; since $\rho(B) < r$, we have $\rho(\frac{1}{r}B) < 1$ and so we $\left(\frac{1}{r}B\right)^k \underset{k\to\infty}{\longrightarrow} 0$, so the power series is convergent: $A^{-1} = \frac{1}{r}(\mathrm{I} - \frac{1}{r}B)^{-1} = \frac{1}{r}\left(\mathrm{I} + \left(\frac{1}{r}B\right) + \left(\frac{1}{r}B\right)^2 + \dots\right)$
Each term is non-negative so their sum is also non-negative.

$\Leftarrow$ As noted in ?.9, a Z-matrix can be written in the form $A = r\,\mathrm{I} - B$ with $B \geq 0$. Suppose by contradiction that for a such decomposition, $r \leq \rho(B)$. Applying Frobenius theorem to B we obtain that $\rho(B)$ is an eigenvalue of B, so $r - \rho(B)$ is an eigenvalue of A with positive eigenvector $p > 0$ (which is the Perron eigenvector of B).
If $r = \rho(B)$, then A has eigenvalue 0, which is impossibile since we assumed A monotone and thus non-singular.
If $r < \rho(B)$, then $Ap = (r\,\mathrm{I} - B)p = (r - \rho(B))p < 0$ because $p > 0$ as a vector and $r - \rho(B) < 0$ as scalar. Since A is monotone, $A^{-1} > 0$ and $A^{-1}Ap < A^{-1}0$, so $p < 0$ which is a contradiction. So the only remaining possibility is $r > \rho(B)$, which implies that A is a non-singular M-matrix. □

Corollary 3.12 f A is non-singular M-matrix, then for every decomposition $A = r\,\mathrm{I} - B$ with $B \geq 0$, necessarily $r > \rho(B)$.
Indeed for the $\Rightarrow$ part of the above theorem, A being a non-singular M-matrix implies that A is monotone. Applying the $\Leftarrow$ part proof we obtain that every decomposition with $B \geq 0$ has $r > \rho(B)$.

Proposition 3.13 Let $A \in \mathbb{R}^{n\times n}$ be a singular, irreducible M-matrix. Then $\mathrm{rank}(A) = n - 1$ and for every decomposition $A = r\,\mathrm{I} - B$ with $B \geq 0$, necessarily $r = \rho(B)$.

Proof. Since A is a M-matrix, there exists a decomposition $A = r_0\,\mathrm{I} - B_0$ with $r_0 \geq \rho(B_0)$. Since A is singular $0 \in \sigma(A)$; this means there exists an eigenvalue $\lambda_0 \in \sigma(B)$ such that $r_0 - \lambda_0 = 0$. Since $|\lambda_0| \leq \rho(B_0) \leq r_0$, we obtain that necessarily $\lambda_0 = \rho(B_0) = r_0$.
Take another decomposition $A = r\,\mathrm{I} - B$ with $B \geq 0$. Then $r\,\mathrm{I} - B = r_0\,\mathrm{I} - B_0$, or equivalently $B = B_0 + (r - r_0)\,\mathrm{I}$. This means that the eigenvalues of B are shifted by $r - r_0$ compared to the eigenvalues of B_0. So $\lambda = \rho(B_0) + r - r_0$ is an eigenvalue of B, and is the maximum real eigenvalue. Since $B \geq 0$, for Frobenius theorem $\rho(B)$ is an eigenvalue and it's the maximum

real eigenvalue, so necessarily $\rho(B) = \rho(B_0) + r - r_0 = r$ since $\rho(B_0) = r_0$.

We have proved that for every decomposition $A = r\,\mathrm{I} - B$ with $B \geq 0$, necessarily $\rho(B) = r$. Now we will prove that $\operatorname{rank}(A) = n-1$. $B \geq 0$ is also irreducible, thus for Frobenius theorem $\rho(B)$ is a simple eigenvalue of B, hence $0 = r - \rho(B)$ is a simple eigenvalue of A. This implies that $\operatorname{rank}(A) = n - 1$. □

By combining the two above statements, we make the following conclusion: A is a Z-matrix, then by taking any decomposition $A = r_0\,\mathrm{I} - B_0$ with $B_0 \geq 0$, we can have only one of the following possibilities:

1. $r_0 > \rho(B_0)$ and A is a non-singular M-matrix.
2. $r_0 = \rho(B_0)$ and A is a singular M-matrix.
3. $r_0 < \rho(B_0)$ and A is not a M-matrix.

For every other decomposition $A = r\,\mathrm{I} - B$, r must be greater/equal/less than $\rho(B)$ according to the same case as above for r_0 and $\rho(B_0)$.

Remark 3.14 Dice cose che A è M-matrice, allora $a_{ii} \geq 0$. Se è NSMM, allora > 0. DA COMPLETARE Perché lo fa anche dopo

3.2 A lot of equivalences

Theorem 3.15 (**ZM 2 irreducible**) Let $A \in \mathbb{R}^{n\times n}$ be a Z-matrix. Then the following are equivalent:

1. A is an irreducible non-singular M-matrix.
2. $A^{-1} > 0$.

Proof. $1 \Rightarrow 2$
Write $A = r\,\mathrm{I} - B$ with $r > \rho(B)$. Then $A^{-1} = (r\,\mathrm{I} - B)^{-1} = \frac{1}{r}\left(\mathrm{I} + \left(\frac{1}{r}B\right) + \left(\frac{1}{r}B\right)^2 + \dots\right)$
Since for irreducibility we need to check only if off-diagonal entries are 0 or $\neq 0$, it follows that also B is irreducible. If B is primitive then for sufficiently large m, we have $B^m > 0$ and thus also $A^{-1} > 0$. If B is impritive of index k, write it a Frobenius canonical form with primitive blocks $B_1 \dots B_k$. It can be seen that the above sum cannot have zero entries. DA COMPLETARE

$2 \Rightarrow 1$ By theorem ?.11, A is a non-singular M-matrix. If $A = r\,\mathrm{I} - B$ were reducible, then also B would be reducible and it could be written by a permutation P in a block triangular form: $PBP^\top = \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix}$. Thus we se that $A^{-1} = \frac{1}{r}\left(\mathrm{I} + \left(\frac{1}{r}B\right) + \left(\frac{1}{r}B\right)^2 + \dots\right)$ has to be in the same block triangular form, and cannot be a positive matrix. □

Remark 3.16 A is an irreducible non-singular M-matrix, then $-A$ is essentially positive.
We see that it's just a matter of notation choosing to use "Z- and M-matrices" opposed to "essentially positive and non-negative matrices". We'll use mostly the former approach.

Definition 3.17 (**stable positive-stable matrix**) $A \in \mathbb{C}^{n\times n}$ is a *positive-stable* matrix if for any eigenvector $\lambda \in \sigma(A)$, we have $\mathfrak{Re}(\lambda) > 0$.
$A \in \mathbb{C}^{n\times n}$ is a *stable* matrix if for any eigenvector $\lambda \in \sigma(A)$, we have $\mathfrak{Re}(\lambda) < 0$.
Observe that A is stable if and only if $-A$ is positive-stable. The term stability comes from the solution of differential equations like $x'(t) = Ax$, in which the orbit always converges to 0 if A is stable.

Proposition 3.18 Let $A \in \mathbb{R}^{n\times n}$ be a non-singular M-matrix, then A is positive-stable.

Proof. Write $A = r\,\mathrm{I} - B$ with $r > \rho(B) \geq |\lambda_B| \geq \mathfrak{Re}(\lambda_B)$ for every eigenvalue $\lambda_B \in \sigma(B)$. Since every $\lambda_A \in \sigma(A)$ is related as $\lambda_A = r - \lambda_B$, we have that $\mathfrak{Re}(\lambda_A) = \mathfrak{Re}(r - \lambda_B) = r - \mathfrak{Re}(\lambda_B) > 0$, as desidered. □

Proposition 3.19 Let $A \in \mathbb{R}^{n\times n}$ be a Z-matrix with $\mathfrak{Re}(\lambda) \geq 0$ for every eigenvalue $\lambda \in \sigma(A)$. Then A is a M-matrix.

Proof. Write $A = r\,\mathrm{I} - B$ with $B \geq 0$, we need to prove that $r \geq \rho(B)$. For Frobenius theorem B has eigenvalue $\rho(B)$, and let $\lambda_A = r - \rho(B)$ be its associated eigenvalue.
Using the hypothesis we obtain $0 \leq \mathfrak{Re}(\lambda_A) = \mathfrak{Re}(r - \rho(B)) = r - \rho(B)$, which implies $r \geq \rho(B)$. □

DA COMPLETARE unire i due statement sopra in un teorema

Definition 3.20 (**Stieltjes matrix**) A symmetric non-singular M-matrix is called a *Stieltjes matrix*

Remark 3.21 The class of M-matrix is not closed under addition

Proposition 3.22 Let $A, B \in \mathbb{R}^{n\times n}$ be monotone matrices such that $A \leq B$.
Then $A^{-1} \geq B^{-1} \geq 0$.

Proof. Non dimostrato. □

Theorem 3.23 (**ZM 4**) Let $A \in \mathbb{R}^{n\times n}$ be a Z-matrix. Then the following are equivalent:

1. A is a non-singular M-matrix.
2. Every principal minor of A is positive.

Proof. $1 \Rightarrow 2$
Write $A = r\,\mathrm{I} - B$, with $B \geq 0$ and $r > \rho(B)$. We will prove two propositions:
Step 1 $\det(A) > 0$.
We can write $\det(A) = \prod \lambda_i(A)$. Since $A \in \mathbb{R}^{n\times n}$, the complex eigenvalues $\lambda_i(A)$ will come in conjugate pairs $\lambda, \overline{\lambda}$ with real part greater than 0 for proposition ?.18. This implies that $\lambda \cdot \overline{\lambda} > 0$. Similarly, the real eigenvalues of A will be already positive also for proposition ?.18. This implies that their product $\det(A)$ is a positive number.

Step 2 Every principal submatrix $\widetilde{A}$ of A is a non-singular M-matrix.
We can write $\widetilde{A} = r\,\mathrm{I} - \widetilde{B}$ where $\widetilde{B}$ is the corresponding submatrix of B. For ?.9 we have $\rho(\widetilde{B}) \leq \rho(B) < r$, so also $\widetilde{A}$ is a non-singular M-matrix.

Finally, we can see that every principal minor of A is the determinant of a principal submatrix $\widetilde{A}$. For step 2, $\widetilde{A}$ is a non-singular M-matrix, and for step 1 we have $\det(\widetilde{A}) > 0$.

$2 \Rightarrow 1$ We will prove this by induction on n.
For $n = 1$ it's trivial. For $n = 2$, A is a Z-matrix so it can be written as $A = \begin{pmatrix} a_{11} & -a_{12} \\ -a_{21} & a_{22} \end{pmatrix}$ with $a_{12}, a_{21} \geq 0$. The hypothesis is that $a_{11}, a_{22} > 0$ and $\det(A) = a_{11}a_{22} - a_{12}a_{21} > 0$. Write explicitly the inverse $A^{-1} = \frac{1}{\det A}\begin{pmatrix} a_{22} & a_{21} \\ a_{12} & a_{11} \end{pmatrix}$, we can se that $A^{-1} \geq 0$ so A is a monotone matrix. Then for theorem ?.11 A is a non-singular M-matrix.

Assume now that the thesis is true for $n-1$ and let's prove it for n. Write $A = A_n$ in block form of dimension $n-1, 1$:

$$A_n = \left(\begin{array}{c|c} A_{n-1} & -c \\ \hline -b^\top & a_{nn} \end{array}\right)$$

Since A_n is a Z-matrix, $c, b > 0$ are positive vectors and A_{n-1} is a Z-matrix too. For hypothesis all principal minors of A_n are positive, so all positive minors of A_{n-1} are positive too and by inductive hypothesis we can conclude that A_{n-1} is a non-singular M-matrix. For theorem ?.11 this implies that A_{n-1} is monotone, i.e. $A_{n-1}^{-1} \geq 0$.
Write the following block factorization of A_n:

$$A_n = \left(\begin{array}{c|c} A_{n-1} & -c \\ \hline -b^\top & a_{nn} \end{array}\right) = \left(\begin{array}{c|c} \mathrm{I}_{n-1} & 0 \\ \hline -b^\top A_{n-1} & 1 \end{array}\right) \left(\begin{array}{c|c} A_{n-1} & -c \\ \hline 0 & \sigma \end{array}\right)$$

With $\sigma = a_{nn} - b^\top A_{n-1}^{-1} c$. Since $\det(A_n), \det(A_{n-1}) > 0$ for hypothesis, from the above factorization and Binet's theorem we obtain $\det(A_n) = \det(A_{n-1}) \cdot \sigma$, and thus $\sigma > 0$. Write now the inverse of A_n in block form:

$$A_n^{-1} = \left(\begin{array}{c|c} A_{n-1} & -c \\ \hline 0 & \sigma \end{array}\right)^{-1} \left(\begin{array}{c|c} \mathrm{I}_{n-1} & 0 \\ \hline -b^\top A_{n-1} & 1 \end{array}\right)^{-1} = \left(\begin{array}{c|c} A_{n-1}^{-1} & \frac{1}{\sigma} A_{n-1}^{-1} c \\ \hline 0 & \frac{1}{\sigma} \end{array}\right) \left(\begin{array}{c|c} \mathrm{I}_{n-1} & 0 \\ \hline b^\top A_{n-1} & 1 \end{array}\right)$$

Since $A_{n-1}^{-1}, b, c, \sigma$ are all positive, this implies that $A_n^{-1} \geq 0$, thus A_n is monotone and by theorem ?.11, A_n is a non-singular M-matrix. $\square$

Theorem 3.24 (**LU factorization for M-matrices**)
Let $A \in \mathbb{R}^{n\times n}$ be a non-singular M-matrix. Then there exists matrix L, U such that: $A = LU$, L is lower triangular, U is upper triangular, both L and U have non-positive off-diagonal entries and positive diagonal entries.
This factorization is in general non unique, but it becomes so if we require that the diagonal entries of L are all 1.

Proof. As in theorem ?.23, write A_n in block triangular factorization:

$$A_n = \left(\begin{array}{c|c} A_{n-1} & -c \\ \hline -b^\top & a_{n\,n} \end{array}\right) = \left(\begin{array}{c|c} \mathrm{I}_{n-1} & 0 \\ \hline -b^\top A_{n-1} & 1 \end{array}\right) \left(\begin{array}{c|c} A_{n-1} & -c \\ \hline 0 & \sigma \end{array}\right)$$

By using induction on A_{n-1} we obtain an LU factorization with all diagonal entries of L equal to 1. $\square$

Proposition 3.25 Let $A \in \mathbb{R}^{n\times n}$ be a (lower or upper) triangular matrix, with all diagonal entries > 0 and off-diagonal ≤ 0. Then A is a non-singular M-matrix.

Proof. We will prove the proposition for A upper triangular, as the other case is similar.
Write $A = D - U$ where D is diagonal, $D > 0$, and U is strictly upper triangular (i.e it's diagonal entries are zero) and $U \geq 0$.
Note that U is nilpotent with $U^n = 0$; also call $B = D^{-1}U$ is also nilpotent and $B^n = 0$.
Since $\det(A) = \det(D) > 0$, A is non-singular and we can write the inverse:

$$A^{-1} = (D-U)^{-1} = (\,D\,(I - D^{-1}U)\,)^{-1} = (I-B)^{-1}D^{-1} = (I + B + B^2 + \ldots + B^{n-1})D^{-1}$$

Observe that $\rho(B) = 0$ and thus not only $(I-B)^{-1}$ exists, but when written as a power series it's a finite sum since B is nilpotent.
We can see that the RHS of the above formula is non-negative, so $A^{-1} \geq 0$ and by theorem ?.11 we obtain that A is a non-singular M-matrix. $\square$

Proposition 3.26 Let A be a Z-matrix and suppose A is also diagonally dominant. Then A is a non-singular M-matrix.

Proof. By Gershgorin's theorems we know that A is non-singular and all it's eigenvalues lie on the right half of the complex plane, i.e. $\mathfrak{Re}(\lambda) > 0 \; \forall \lambda \in \sigma(A)$.
By theorem ?.19 we obtain that A is a non-singular M-matrix. $\square$

Theorem 3.27 (**Schur's complement is a non-singular M-matrix**)
Let $A \in \mathbb{R}^{n\times n}$ be a non-singular M-matrix. Write it in a block form:

$$A = \left(\begin{array}{c|c} A_{1\,1} & A_{1\,2} \\ \hline A_{2\,1} & A_{2\,2} \end{array}\right)$$

With $A_{1\,1}$ and $A_{2\,2}$ square matrices. Then $A_{1\,1}, A_{2\,2}$ and the Schur's complement of A, defined as $S = A_{2\,2} - A_{2\,1}A_{1\,1}^{-1}A_{2\,1}$, are all three non-singular M-matrices.

Proof. Clearly A_{11} and A_{22} are Z-matrices since they have the same diagonal and off-diagonal elements of A. Since A is a non-singular M-matrix it can be written as $A = r\,\mathrm{I} - B$ with $r > \rho(B)$; then by writing $A_{11} = r\,\mathrm{I} - \widehat{B}$ with $\widehat{B}$ a principal submatrix of B, we have $\rho(\widehat{B}) \leq \rho(B) < r$, so A_{11} is a non-singular M-matrix. The proof for A_{22} is analogous.
We have then that A_{11} and A_{22} are non singular. It is possibile to factorize A as follows:

$$A = \left(\begin{array}{c|c} \mathrm{I} & 0 \\ \hline A_{21}A_{11}^{-1} & \mathrm{I} \end{array}\right) \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & S \end{array}\right)$$

These two matrices are non-singular, so also S is non-singular and invertible.
Now we can write A^{-1} as a product of the inverses of the above matrices:

$$A^{-1} = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & S \end{array}\right)^{-1} \left(\begin{array}{c|c} \mathrm{I} & 0 \\ \hline A_{21}A_{11}^{-1} & \mathrm{I} \end{array}\right)^{-1} = \left(\begin{array}{c|c} A_{11}^{-1} & -A_{11}^{-1}A_{12}S^{-1} \\ \hline 0 & S^{-1} \end{array}\right) \left(\begin{array}{c|c} \mathrm{I} & 0 \\ \hline -A_{21}A_{11}^{-1} & \mathrm{I} \end{array}\right)$$

Calculate the last block of this matrix product (we do not care about the other blocks):

$$A^{-1} = \left(\begin{array}{c|c} * & * \\ \hline * & S^{-1} \end{array}\right)$$

Since A is a non-singular M-matrix, A is monotone and $A^{-1} \geq 0$, this means that also $S^{-1} \geq 0$, so S is a monotone matrix.
Now we prove that S is also a Z-matrix. $S = A_{22} - A_{21}A_{11}^{-1}A_{21}$; Since A_{21} and A_{12} are blocks that do not contain any diagonal of A, all they entries are ≤ 0 since A is a M-matrix. $A_{11}^{-1} \geq 0$ since A_{11} is a non-singular M-matrix, so all these observations imply $A_{21}A_{11}^{-1}A_{21} \geq 0$. This means that S is formed by a Z-matrix A_{22} from which be subtract a non-negative matrix, so S is also a Z-matrix.
Since S is a Z-matrix and monotone, from theorem ?.11 we obtain that is also a non-singular M-matrix. ☐

Observe that we can further expand the factorization for A:

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right) = \left(\begin{array}{c|c} \mathrm{I} & 0 \\ \hline A_{21}A_{11}^{-1} & \mathrm{I} \end{array}\right) \left(\begin{array}{c|c} A_{11} & 0 \\ \hline 0 & S \end{array}\right) \left(\begin{array}{c|c} \mathrm{I} & A_{11}^{-1}A_{12} \\ \hline 0 & \mathrm{I} \end{array}\right)$$

If A is a symmetric, positive-definite matrix, then the Schur's complement S is also symmetric and positive-definite. Indeed, we have that $\left(A_{21}A_{11}^{-1}\right)^{\top} = A_{11}^{-1}A_{12}$ and the above is a congruence relation. This implies that $\left(\begin{array}{c|c} A_{11} & 0 \\ \hline 0 & S \end{array}\right)$ is a symmetric positive-definite matrix and so also S is.

Proposition 3.28 Se A è non-singular M-matrix, allora ogni sottomatrice principale di A è non-singular M-matrix. Inoltre se $B \geq A$ è sempre Z-matrix, allora anche B è non-singular M-matrix. La seconda cosa corrisponde ad aumentare i valori diagonali, oppure a diminuire in modulo i valori fuori diagonale che sono negativi. DA COMPLETARE

Definition 3.29 Let $A \in \mathbb{C}^{n\times n}$, a difference decomposition $A = M - N$ is called a *splitting* if M is a non-singular matrix.
If we have that $\rho(M^{-1}N) < 1$, then it's called a *convergent splitting*
DA COMPLETARE Inserire esempio con il metodo del punto fisso

Definition 3.30 Let $A \in \mathbb{R}^{n\times n}$, a splitting $A = M - N$ is called:

- *regular* if $M^{-1} \geq 0$ and $N \geq 0$.
- *weak regular* if $M^{-1} \geq 0$ and $M^{-1}N \geq 0$.

Remark 3.31 Regular splitting $\Rightarrow$ weak regular splitting

Theorem 3.32 (**Ortega, Varga**) Let $A \in \mathbb{R}^{n\times n}$ and $A = M - N$ be a weak regular splitting. Then the following are equivalent:

1. A is monotone.
2. The splitting is convergent, i.e. $\rho(M^{-1}N) < 1$

Proof. 1 $\Rightarrow$ 2
Call $T = M^{-1}N$, by hypothesis $T \geq 0$. Note that $I - T = M^{-1}M - M^{-1}N = M^{-1}A$, so $(I-T)A^{-1} = M^{-1}$ which is ≥ 0 still by hypothesis.Also $A^{-1} \geq 0$ because A is monotone.
Consider the following expression:

$$(I+T+T^2+\ldots+T^m)M^{-1} = (I+T+T^2+\ldots+T^m)(I-T)A^{-1} = (I-T^{m+1})A^{-1} = A^{-1}-T^{m+1}A^{-1}$$

Let us call $S_m = \sum_{k=0}^{m} T^k$, then from the leftmost expression we obtain that $S_m M^{-1} \geq 0$ since it's a product of non-negative matrices, and from the rightmost we obtain that $S_m M^{-1} = A^{-1} - T^{m+1}A^{-1} \leq A^{-1}$ since both $T, A^{-1} \geq 0$, so we're subtracting a positive matrix from A^{-1}. So for all $m \in \mathbb{N}$ we obtain $0 \leq S_m \leq MA^{-1}$. This means that the sequence of partial sums of non-negative matrices $S_m = \sum_{k=0}^{m} T^k$ is bounded, so it must converge to some matrix in $\mathbb{C}^n$ (because each term $T^k \geq 0$). This means that $\lim_{m\to\infty} T^m = 0$, which is possible only if $\rho(T) < 1$, which is what we wanted to prove.

2 $\Rightarrow$ 1 Assume that $\rho(M^{-1}N) = \rho(T) < 1$.
Note that $M^{-1}A = I - M^{-1}N$; since $\rho(M^{-1}N) < 1$, then 0 cannot be an eigenvalue of $I - M^{-1}N$, so $M^{-1}A$ is non-singular and thus A is invertible.
From $M^{-1}A = I - T$ we obtain that $M^{-1} = (I-T)A^{-1}$ and $M^{-1}(I-T)^{-1} = A^{-1}$. Write $(I-T)^{-1}$ as $\sum_{k=0}^{\infty} T^k$ (the limit exists since $\rho(T) < 1$), so $A^{-1} = \left(\sum_{k=0}^{\infty} T^k\right) M^{-1}$ is a product of non-negative matrices, thus $A^{-1} \geq 0$ and A is monotone. $\square$

Theorem 3.33 (**ZM 5**) Let $A \in \mathbb{R}^{n\times n}$ be a Z-matrix. The following are equivalent:

1. A is a non-singular M-matrix.

2. A has a convergent, weak splitting

Proof. Since A is a Z-matrix, then it can be written as $A = r\,\mathrm{I} - B$ with $B \geq 0$. By taking $M = r\,\mathrm{I}$ and $N = B$, we obtain a regular splitting for A, since $M^{-1} = \frac{1}{r}\,\mathrm{I} > 0$ and $N = B \geq 0$, thus it's also a weak regular splitting.

$1 \Rightarrow 2$ If A is a non-singular M-matrix, then $r > \rho(B)$ and $M^{-1}N = \frac{1}{r}B$ has spectral radius $\rho(\frac{1}{r}B) < 1$. This means that the splitting is convergent.

$2 \Rightarrow 1$ A has a convergent weak regular splitting, then by theorem ?.32 it follows that A is monotone. Combining this with the fact that A is a Z-matrix we obtain by theorem ?.11 that A is a non-singular M-matrix.

□

DA COMPLETARE Considerazioni su Jacobi e Gauss-Seidel

THEOREM 3.34 (**ZM 6**) Let $A \in \mathbb{R}^{n\times n}$ be a Z-matrix. The following are equivalent:

1. A is a non-singular M-matrix.

2. $A + I$ is non-singular. Define $G = (A + I)^{-1}(A - I)$, we have $\rho(G) < 1$.

Proof. $1 \Rightarrow 2$
Consider the function $\varphi : \mathbb{C} \setminus \{-1\} \to \mathbb{C} \setminus \{1\}$, $\varphi(z) = \frac{z-1}{z+1}$. φ is a biolomorphic map between the right half-plane $H = \{z | \mathfrak{Re}(z) > 0\}$ and the disk $D = \{z | \, |z| < 1\}$.
When viewed as matrix function, we have $\phi(A) = G$ and there is a relation between eigenvalues: $\lambda_i(G) = \dfrac{\lambda_i(A) - 1}{\lambda_i(A) + 1}$. Since A is a non-singular M-matrix, its eigenvalues $\lambda_i(A) \in H$ and so $\lambda_i(G) = \varphi(\lambda_i(A)) \in D$, i.e. $|\lambda_i(G)| < 1$.

$2 \Rightarrow 1$ DA COMPLETARE □

THEOREM 3.35 (**General conditions for non-singular M-matrices**)
Let $A \in \mathbb{R}^{n\times n}$ with $n \geq 2$. The following are equivalent:

1. A is a non-singular M-matrix.

2. $\forall$ D non-negative diagonal matrix, $A + D$ is a monotone matrix.

3. $\forall \alpha \geq 0$, $A + \alpha I$ is a monotone matrix.

4. Every principal submatrix of A is monotone

5. Every principal submatrix of A of size 1,2,n is monotone.

Proof. We immediatly see that 2 $\Rightarrow$ 3 and 4 $\Rightarrow$ 5, as the latter are special cases of the former. We already saw in theorem ?.23 that if A is a non-singular M-matrix, then every principal minor of A is positive, so 1 $\Rightarrow$ 4.

1 $\Rightarrow$ 2 Let D be a non-negative diagonal matrix and A is a non-singular M-matrix. Then $A+D$ is certainly a Z-matrix, since its off-diagonal entries are still ≤ 0. Let $d_{\max}$ be the maxium entry of D, then $d_{\max}\,\mathrm{I} - D \geq 0$. Since A is a non-singular M-matrix, we can write it as $A = r\,\mathrm{I} - B$ with $r > \rho(B)$.

$$A + D = r\,\mathrm{I} - B + D = (r + d_{\max})\,\mathrm{I} - (B + d_{\max}\,\mathrm{I} - D)$$

This gives us a Z-matrix splitting for $A + D$, since $B + d_{\max}\,\mathrm{I} - D \geq 0$. If we prove that $A + D$ is a non-singular M-matrix, then by theorem ?.11 we will obtain that $A + D$ is monotone.
We have $B + d_{\max}\,\mathrm{I} - D \leq B + d_{\max}\,\mathrm{I}$, then by theorem 9 and using that $B \geq 0$ we obtain $\rho(B + d_{\max}\,\mathrm{I} - D) \leq \rho(B + d_{\max}\,\mathrm{I}) = \rho(B) + d_{\max}$. This is exactly what we needed to prove that $A + D$ is a non-singular M-matrix.

3 $\Rightarrow$ 1 By taking $\alpha = 0$, we obtain that A is a monotone matrix. We want to prove that A is also a Z-matrix. Assume by contradiction that there exists $a_{i\,j} > 0$ with $i \neq j$.
Take $\beta > 0$ small enough such that $\rho(\beta A) < 1$. Then $I + \beta A$ is invertible, its inverse can be written as a power series and $(\mathrm{I} + \beta A)^{-1} = I - \beta A + \mathcal{O}(\beta^2)$, then the i, j^{th} component of $(\mathrm{I} + \beta A)^{-1}$ is approximately $-\beta a_{i\,j} < 0$. This means that $(\mathrm{I} + \beta A)$ is not monotone, and neither is $\frac{1}{\beta}(\mathrm{I} + \beta A) = \frac{1}{\beta}\,\mathrm{I} + A$, which is a contradiction.

5 $\Rightarrow$ 1 The principal submatrices of size one are the diagonal entries $a_{i\,i}$, it's inverse is $\frac{1}{a_{i\,i}}$ and is non-negative because the submatrix is monotone; this means that $a_{i\,i} \geq 0 \;\forall\, i$.
Now let's prove that A is a Z-matrix, so take $i \neq j$ and we want to show that $a_{i\,j}, a_{j\,i} \leq 0$.
Take the size 2 principal submatrix given by $i^{\text{th}}, j^{\text{th}}$ row an column,

$$B = \begin{pmatrix} a_{i\,i} & a_{i\,j} \\ a_{j\,i} & a_{j\,j} \end{pmatrix}$$

We can explicitly write the inverse B^{-1}, which is non-negative for the hypothesis:

$$B^{-1} = \frac{1}{\det(B)} \begin{pmatrix} a_{j\,j} & -a_{i\,j} \\ -a_{j\,i} & a_{i\,i} \end{pmatrix} = \frac{1}{a_{i\,i}\,a_{j\,j} - a_{i\,j}\,a_{j\,i}} \begin{pmatrix} a_{j\,j} & -a_{i\,j} \\ -a_{j\,i} & a_{i\,i} \end{pmatrix} \geq 0$$

By looking at the diagonal entries we have that $\frac{a_{i\,i}}{\det(B)} \geq 0$ and since $a_{i\,i} \geq 0$, we obtain that $\det(B) > 0$. Now by looking at the off-diagonal entries, we have $\frac{-a_{i\,j}}{\det(B)} \geq 0$ and thus $a_{i\,j} \leq 0$.
Finally, we have that $A^{-1} \geq 0$ still from the hypothesis; combined with the fact that A is a Z-matrix, we obtain that A is a non-singular M-matrixfor theorem ?.11.

□

Ci starebbe anche fare un mega teorema ZM che raccoglie tutti i se e solo se detti in precedenza
DA COMPLETARE

We have obtained a lot of properties for non-singular M-matrices, so let's see how these results can be applied to (singular) M-matrices.

THEOREM 3.36 (**M-matrix is the limit of non-singular M-matrices**)
Let $A \in \mathbb{R}^{n\times n}$ be a Z-Matrix. The following are equivalent:

1. A is a M-matrix.
2. $A + \varepsilon\,\mathrm{I}$ is a non-singular M-matrix for every $\varepsilon > 0$.

Proof. $1 \Rightarrow 2$
A is a M-matrix, so it can be written as $A = r\,\mathrm{I} - B$ with $r \geq \rho(B)$ adn $B \geq 0$. This means that $A + \varepsilon\,\mathrm{I} = (r+\varepsilon)\,\mathrm{I} - B$ has $r + \varepsilon > r \geq \rho(B)$, and so is a non-singular M-matrix.

$2 \Rightarrow 1$
As a Z-matrix, A can be written as $A = r\,\mathrm{I} - B$ with $B \geq 0$. Then $A + \varepsilon\,\mathrm{I} = (r+\varepsilon)\,\mathrm{I} - B$ is a non-singular M-matrix, and we have showed that for every decomposition of this form with $B \geq 0$, necessarily $r + \varepsilon > \rho(B)$. By taking the limit $\varepsilon \to 0$, we obtain that $r \geq \rho(B)$. □

We have obtained that a singular M-matrix can be always viewed as limit of non-singular M-matrices. This means that a lot of theorems proved for non-singular M-matrices are valid even for singular M-matrices, changing strict inequalities $>$ with $\geq$ when needed.

THEOREM 3.37 (**Z-matrix is a M-matrix when...**)
Let $A \in \mathbb{R}^{n\times n}$ be a Z-matrix. The following are equivalent:

1. A is a M-matrix.
2. Every principal minor of A (including $\det(A)$) is ≥ 0
3. Every eigenvalue of A, $\lambda \in \sigma(A)$ satisfies $\mathfrak{Re}(\lambda) \geq 0$. (A is positive semi-stable)
4. Every non-zero eigenvalue $\lambda \in \sigma(A)$ satisfies $\mathfrak{Re}(\lambda) > 0$.
5. Every real eigenvalue of A is non-negative: $\lambda \in \mathbb{R} \Rightarrow \lambda \geq 0$.
6. For every positive diagonal matrix D, $A + D$ is non-singular.
7. $\forall\,\alpha > 0$, $A + \alpha\,\mathrm{I}$ is non-singular.
8. There exists a P permutation, L lower triangle and U upper triangle matrices such that the diagonal entries of L and U are non-negative and $PAP^\top = LU$
9. $A + I$ is non-singular. Define $G = (A+I)^{-1}(A-I)$, we have $\rho(G) \leq 1$

Proof. We will not prove it :(□

The following is an analogous of theorem ?.35 for general M-matrices.

THEOREM 3.38 (**General conditions for M-matrices**)
Let $A \in \mathbb{R}^{n\times n}$. The following are equivalent:

1. A is a M-matrix.
2. $A+D$ is monotone for every diagonal D with positive diagonal entries.
3. $A+\alpha\,\mathrm{I}$ is monotone for every $\alpha>0$

Proof. We will not prove it :($\square$

3.3 Property c

DEFINITION 3.39 Let $A\in\mathbb{C}^{n\times n}$. The *index* of A is the smallest integer $k\in\mathbb{N}$ sucht that $\operatorname{rank}(A^k)=\operatorname{rank}(A^{k+1})$. It will be denoted by $\operatorname{index}(A)$

REMARK 3.40 The index is the size of the largest Jordan block of A relative to eigenvalue 0.

REMARK 3.41 A is non-singular $\iff$ the index of A is 0.

DEFINITION 3.42 Let $A\in\mathbb{R}^{n\times n}$ be a M-matrix. We say that A has *property c* if there exists $s\in\mathbb{R}, s>0$ such that $A=s\,\mathrm{I}-B$ with $B\geq 0$ and such that $T=\dfrac{1}{s}B$ is convergent.

We recall that T is convergent if there exists the limit $\lim_{m\to\infty}T^m=B\in\mathbb{C}^{n\times n}$, and that T is zero-convergent if such limit is zero.

REMARK 3.43 If A is a non-singular M-matrix, then A has *property c*.
Indeed, write $A=r\,\mathrm{I}-B$, with $r>\rho(B)$. By taking $s=r$ we have $T=\dfrac{1}{r}B$ has $\rho(T)<1$, so T is zero-convergent.

Not every M-matrix has *property c*. Take for example $A=\begin{pmatrix}0&-1\\0&0\end{pmatrix}$. Since $A+\varepsilon\,\mathrm{I}$ is a non-singular M-matrix, then by theorem ?.36 we have that A is a M-matrix. For every $s>0$, $A=s\,\mathrm{I}-\begin{pmatrix}s&1\\0&s\end{pmatrix}$, and we can easily see that the powers of $\frac{1}{s}B=\begin{pmatrix}1&\frac{1}{s}\\0&1\end{pmatrix}$ diverge.
Cose sul caso opposto, DA COMPLETARE

THEOREM 3.44 Let $A\in\mathbb{R}^{n\times n}$ be a M-matrix. Then the following are equivalent:

1. A has *property c*.
2. $\operatorname{index}(A)\leq 1$.

Proof. $1\Rightarrow 2$ Write $A=s\,\mathrm{I}-B$ as the splitting given by *property c*, then $T=\frac{1}{s}B$ is convergent. Let J be the Jordan Normal form of T, it has a block (possibly of size 0) of eigenvalues equal to 1, no other eigenvalues of absolute size 1, and the other eigenvalues have $|\lambda|<1$.

$$T = XJX^{-1} = X \left(\begin{array}{c|c} I & 0 \\ \hline 0 & K \end{array} \right) X^{-1}$$

With $\rho(K) < 1$. Then $A = s(I - T)$:

$$A = sX(I-J)X^{-1} = s\,X \left(\begin{array}{c|c} I-I & 0 \\ \hline 0 & I-K \end{array} \right) X^{-1} = X \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & s(I-K) \end{array} \right) X^{-1}$$

Since $\rho(K) < 1$, $s(I-K)$ is non singular and thus A has eigenvalues 0 only in the upper left block, which means that $\operatorname{index}(A) \leq 1$.

2 $\Rightarrow$ 1 If $\operatorname{index}(A) = 0$, then A is a non-singular M-matrix; by the above remark A has *property c*.
If $\operatorname{index}(A) = 1$, write $A = s\,\mathrm{I} - B$ with $B \geq 0$ and $s = \rho(B)$. By Perron-Frobenius, $T = \frac{1}{s}B$ has spectral radius $\rho(T) = 1$.
By theorem ?.37 and the fact that $\mathfrak{Re}(\lambda_B) \leq s$ for every eigenvalue of B, we obtain that $0 \leq \lambda_A \leq s$ for every $\lambda_A \in \sigma(A)$.
Write A in its Jordan Normal form:

$$A = XJX^{-1} = X \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & K \end{array} \right) X^{-1}$$

With $0 < \lambda_K \leq s$ for every $\lambda_K \in \sigma(K)$. Let $T = \dfrac{1}{s}B = \mathrm{I} - \dfrac{1}{s}A$; change to the base given by X:

$$T = X \left(\mathrm{I} - \frac{1}{s}J \right) X^{-1} = X \left(\begin{array}{c|c} \mathrm{I} & 0 \\ \hline 0 & \mathrm{I} - \frac{1}{s}K \end{array} \right) X^{-1}$$

DA COMPLETARE e sistemare tutto :(

□

THEOREM 3.45 Let $A \in \mathbb{R}^{n \times n}$ be a singular, irreducible M-matrix. Then:

1. $\operatorname{rank}(A) = n - 1$
2. A has *property c*.
3. Every principal submatrix of A (escluding A itself) is a non-singular M-matrix.

Proof. DA COMPLETARE □

REMARK 3.46 Corollario che se A è singular irreducible M-matrix, esiste una fattorizzazione LU con diag(L)=(1,1,...1) e $\operatorname{diag}(U) = (u_1, u_2 \ldots u_{n-1}, 0)$. DA COMPLETARE

3.4 Generalized inverses

A matrix A if invertible if and only if is square and is non-singular, in this case $A \cdot A^{-1} = \mathrm{I}$. It is possible to generalize this notion to singular or non-square matrices, by requiring that some cancellation laws can be applied to the matrix and its *pseudoinverse*.

Definition 3.47 (**Inner and outer inverse**) Let $A \in \mathbb{C}^{n\times n}$, then $X \in \mathbb{C}^{n\times n}$ is called an

- *Inner inverse* or *{1}-inverse* if $AXA = A$.
- *Outer inverse* or *{2}-inverse* if $XAX = X$.

If X is both inner and outer inverse, it is called also a *{1,2}-inverse*.
There are many matrices that are the {1,2}-inverse of A, but we can add other conditions to have uniqueness, obtaining the:

Definition 3.48 (**Moore-Penrose pseudoinverse**) Let $A \in \mathbb{C}^{n\times n}$. Then there exists and it's unique a matrix X such that:

1. $AXA = A$.
2. $XAX = X$.
3. $(AX)^H = AX$.
4. $(XA)^H = XA$.

The first two are the inner-outer inverse conditions, while the last two require AX, XA to be Hermitian. The Moore-Penrose pseudoinverse is usually denoted by A^+. It can be defined also through the singular-value decomposition. It exists and is unique also for non-square matrices $A \in \mathbb{C}^{m\times n}$.

DA COMPLETARE Serve davvero irriducibile? Forse solo per avere L una NSMM

We have seen that if A is an irreducible, singular M-matrix, then it has a factorization $A = LD\widehat{U}$, with L lower triangular, U upper triangular,D diagonal, $\mathrm{diag}(L) = \mathrm{diag}(U) = (1, 1, \dots, 1)$ and $\mathrm{diag}(D) = u_1, \dots, u_{n-1}, 0$.
Tben a {1,2}-inverse of A is given by $A^- = U^{-1}D^-L^{-1}$ where $D^- = \mathrm{diag}(u_1^{-1}, \dots, u_{n-1}^{-1}, 0)$.
Scrivere A^- con la forma a blocchi dell'induzione. DA COMPLETARE

Proposition 3.49 If A is a singular M-matric and A^+ is its Moore-Penrose inverse, then $A^+ \geq 0$ if and only if : $A = 0$ or there exists P permutation matrix such that $PAP^\top = \left(\begin{array}{c|c} M & 0 \\ \hline 0 & 0 \end{array}\right)$ where M is a non-singular M-matrix.
We see that A is reducible, and $A^+ = P^\top \left(\begin{array}{c|c} M^{-1} & 0 \\ \hline 0 & 0 \end{array}\right) P$.

Proof. We will not prove it. :(□

Remark 3.50 Not every singular M-matrix has an LU factorization. An example is $A = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$. However, this is possible if we allow to permute M:

Theorem 3.51 Let A be a (singular?) M-matrix. There exists a permutation matrix P, L lower triangular with $\mathrm{diag}(L) = (1, 1, \dots 1)$, U upper triangular and singular such that $PAP^\top = LU$.

Proof. DA COMPLETARE □

Definition 3.52 (**Core-nilpotent decomposition**) Let $A \in \mathbb{C}^{n\times n}$ and suppose that there exists $S \in \mathbb{C}^{n\times n}$, S non singular and $B \in \mathbb{C}^{n_1\times n_1}$, $N \in \mathbb{C}^{n_2\times n_2}$ square matrices with $n_1 + n_2 = n$ such that:

$$A = S \left(\begin{array}{c|c} B & 0 \\ \hline 0 & N \end{array}\right) S^{-1}$$

Then S, B, N form the so called *core-nilpotent* decomposition of A
There always exists a core-nilpotent decomposition for A: it is sufficient to take J the Jordan normal form of A, so $A = SJS^{-1}$, J is block diagonal and the block J_0 relative to eigenvalue 0 (if $0 \in \sigma(A)$) is nilpotent.
Clearly, the core-nilpotent decomposition is not unique. Nevertheless, we can give the following definition that (we will prove) does not actually depend on the decomposition we have used.

Definition 3.53 (**Drazin inverse**)
Let $A \in \mathbb{C}^{n\times n}$ and let $A = S \left(\begin{array}{c|c} B & 0 \\ \hline 0 & N \end{array}\right) S^{-1}$ be a core-nilpotent decomposition. Then the following matrix:

$$A^D = S \left(\begin{array}{c|c} B & 0 \\ \hline 0 & 0 \end{array}\right) S^{-1}$$

Is called the *Drazin (generalized) inverse* of A.
Cose sulla dimensione di N e sull'indice di A DA COMPLETARE

Proposition 3.54 The Drazin inverse of A does not depend on the chosen core-nilpotent decomposition.

Proof. Take two core-nilpotent decompositions of A:

$$A = S \left(\begin{array}{c|c} B & 0 \\ \hline 0 & N \end{array}\right) S^{-1} = T \left(\begin{array}{c|c} C & 0 \\ \hline 0 & M \end{array}\right) T^{-1}$$

With N, M nilpotent and B, C non-singular. By taking the k^{th} power, we obtain:

$$A^k = S \left(\begin{array}{c|c} B^k & 0 \\ \hline 0 & N^k \end{array}\right) S^{-1} = T \left(\begin{array}{c|c} C^k & 0 \\ \hline 0 & M^k \end{array}\right) T^{-1}$$

With k greater that the orders of nilpotency of N and M, we obtain that B^k and C^k must have the same rank, i.e. B and C have the same size.

Call $R = T^{-1}S$, then the above equality for A can be rewritten as:

$$R\left(\begin{array}{c|c} B & 0 \\ \hline 0 & N \end{array}\right) = \left(\begin{array}{c|c} C & 0 \\ \hline 0 & M \end{array}\right) R$$

$$\left(\begin{array}{c|c} R_{1\,1} & R_{1\,2} \\ \hline R_{2\,1} & R_{2\,2} \end{array}\right) \left(\begin{array}{c|c} B & 0 \\ \hline 0 & N \end{array}\right) = \left(\begin{array}{c|c} C & 0 \\ \hline 0 & M \end{array}\right) \left(\begin{array}{c|c} R_{1\,1} & R_{1\,2} \\ \hline R_{2\,1} & R_{2\,2} \end{array}\right)$$

$$\left(\begin{array}{c|c} R_{1\,1}\,B & R_{1\,2}\,N \\ \hline R_{2\,1}\,B & R_{2\,2}\,N \end{array}\right) = \left(\begin{array}{c|c} C\,R_{1\,1} & C\,R_{1\,2} \\ \hline M\,R_{2\,1} & M\,R_{2\,2} \end{array}\right)$$

By looking at the off-diagonal blocks, we obtain $R_{1\,2}\,N = C\,R_{1\,2}$ and $R_{2\,1}\,B = M\,R_{2\,1}$. We will prove that this implies that $R_{1\,2} = 0, R_{2\,1} = 0$.
Indeed, N is nilpotent of index b, then $R_{1\,2}N^b = 0$ and $0 = R_{1\,2}\,N^b = (R_{1\,2}\,N)\,N^{b-1} = C\,R_{1\,2}\,N^{b-1}$. Since C is non singular, necessarily $R_{1\,2}\,N^{b-1} = 0$. We can continue by induction to obtain that $0 = R_{1\,2}\,N = C\,R_{1\,2}$, which means that $R_{1\,2} = 0$. Similarly we obtain that $R_{2\,1} = 0$. From this, the above matrix equation simplifies to:

$$\left(\begin{array}{c|c} R_{1\,1}\,B & 0 \\ \hline 0 & R_{2\,2}\,N \end{array}\right) = \left(\begin{array}{c|c} C\,R_{1\,1} & 0 \\ \hline 0 & M\,R_{2\,2} \end{array}\right) \qquad \text{and} \qquad TS^{-1} = \left(\begin{array}{c|c} R_{1\,1} & 0 \\ \hline 0 & R_{2\,2} \end{array}\right)$$

We want to show that the Drazin inverses constructed from the two decompositions are equal:

$$S\left(\begin{array}{c|c} B^{-1} & 0 \\ \hline 0 & 0 \end{array}\right) S^{-1} \stackrel{?}{=} T\left(\begin{array}{c|c} C^{-1} & 0 \\ \hline 0 & 0 \end{array}\right) T^{-1}$$

$$T^{-1}S\left(\begin{array}{c|c} B^{-1} & 0 \\ \hline 0 & 0 \end{array}\right) \stackrel{?}{=} \left(\begin{array}{c|c} C^{-1} & 0 \\ \hline 0 & 0 \end{array}\right) T^{-1}S$$

$$\left(\begin{array}{c|c} R_{1\,1} & 0 \\ \hline 0 & R_{2\,2} \end{array}\right) \left(\begin{array}{c|c} B^{-1} & 0 \\ \hline 0 & 0 \end{array}\right) \stackrel{?}{=} \left(\begin{array}{c|c} C^{-1} & 0 \\ \hline 0 & 0 \end{array}\right) \left(\begin{array}{c|c} R_{1\,1} & 0 \\ \hline 0 & R_{2\,2} \end{array}\right)$$

The equality holds if and only if $R_{1\,1}B^{-1} = C^{-1}R_{1\,1}$, which is true since $R_{1\,1}B = CR_{1\,1}$ and B, C are invertible.

□

Definition 3.55 (**Equivalent definition of Drazin inverse**)
Let $A \in \mathbb{C}^{n\times n}$ be a matrix of index k. Then the Drazin inverse A^D can be defined as the only matrix that satisfies the following conditions:

1. $AA^D = A^DA$ (commutation)
2. $A^{k+1}A^D = A^k$ (A^D acts like A^{-1} algebraically)
3. $A^DAA^D = A^D$ (A^D is an outer-inverse)

Remark 3.56 In general, A^D is not an inner-inverse, i.e. $AA^DA \neq A$.

In the following statement, we see that the inner-inverses can be used for solving systems of linear equations, both if A is invertible or not.

PROPOSITION 3.57 Let $A \in \mathbb{C}^{n\times n}$ and $b \in \text{span}(A)$ a vector, then a solution of $Ax = b$ can be obtained with $x = A^- b$, where A^- is an inner-inverse of A.

Proof. Since $b \in \text{span}(A)$, then there exists $c \in \mathbb{C}^n$ such that $Ac = b$.
Take $x = A^- b = A^- Ac$, then $Ax = A(A^- Ac) = (AA^- A)c = Ac = b$ since by definition of inner-inverse we have $AA^- A = A$. $\square$

DEFINITION 3.58 (**Group-generalized inverse**) Let $A \in \mathbb{C}^{n\times n}$ with $\text{index}(A) \leq 1$. Then the Drazin inverse is also written as $A^{\#}$ and is called the *group-generalized* inverse of A.
Gruppo generalizzato, dirlo meglio DA COMPLETARE

PROPOSITION 3.59 Let $A \in \mathbb{C}^{n\times n}$, then $AA^D A = A$ if and only if $\text{index}(A) \leq 1$

Proof. If $\text{index}(A) = 0$, then A is invertible and $A^D = A^{-1}$, so the equality holds. Assume now that $k \geq 1$, and take a core-nilpotent decomposition of A:

$$A = S \left(\begin{array}{c|c} B & 0 \\ \hline 0 & N \end{array}\right) S^{-1} \qquad A^D = S \left(\begin{array}{c|c} B^{-1} & 0 \\ \hline 0 & 0 \end{array}\right) S^{-1}$$

Calculate $AA^D A$:

$$\begin{aligned} AA^D A &= S \left(\begin{array}{c|c} B & 0 \\ \hline 0 & N \end{array}\right) \left(\begin{array}{c|c} B^{-1} & 0 \\ \hline 0 & 0 \end{array}\right) \left(\begin{array}{c|c} B & 0 \\ \hline 0 & N \end{array}\right) S^{-1} \\ &= S \left(\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right) \left(\begin{array}{c|c} B & 0 \\ \hline 0 & N \end{array}\right) S^{-1} = S \left(\begin{array}{c|c} B & 0 \\ \hline 0 & 0 \end{array}\right) S^{-1} \end{aligned}$$

Which is equal to A if and only if $N = 0$, i.e. only in the index of N (and so the index of A) is equal to 1.

$\square$

THEOREM 3.60 Let $A \in \mathbb{C}^{n\times n}$. The following are equivalent:

1. $A \in G$, where G is a multiplicative group of matrices.
2. $\text{index}(A) \leq 1$.

Proof. First of all, note that if A is non-singular then $\text{index}(A) = 0$ and $A \in GL(\mathbb{C}, n)$, which is a multiplicative group of matrices. From now on, assume that A is singular.

$1 \Rightarrow 2$ If $A \in G$, there exists the group inverse A^g. Let I_G be the identity of G; it is possible that $\mathrm{I}_G \neq I$, and actually I can't be in G if A is singular. In fact, we only require that $\mathrm{I}_G \cdot B = B$ for every $B \in G$, but not for every $B \in \mathbb{C}^{n\times n}$.

However, we must necessarily have that $A^g A = AA^g = \mathrm{I}_G$ from the group laws. This implies that $A^{m+1}A^g = A^m \; \forall\, m \geq 1$ and $A^g A A^g = A^g$.
These relations satisfy the definition ?.55, so $A^g = A^D$ must be the Drazin inverse of A.
From $A^g A = \mathrm{I}_G$ we obtain that $AA^g A = A$, so by proposition ?.59 we obtain that $\mathrm{index}(A) \leq 1$.

2 $\Rightarrow$ 1 If $\mathrm{index}(A) = 1$, then the core-nilpotent decomposition of A is $A = S \left(\begin{array}{c|c} B & 0 \\ \hline 0 & 0 \end{array} \right) S^{-1}$ where B is non-singular and of size m.
Define the group G as follows:

$$G = \left\{ S \left(\begin{array}{c|c} X & 0 \\ \hline 0 & 0 \end{array} \right) S^{-1} \mid \; \mathrm{size}(X) = \mathrm{rank}(X) = m \right\}$$

The group inverse I_G is then:

$$\mathrm{I}_G = S \left(\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array} \right) S^{-1}$$

Basically this group is a copy of $GL(\mathbb{C}, m) \subseteq GL(\mathbb{C}, n)$, and the embedding is given by the change of basis associated to matrix S.

□

Remark 3.61 We give a more concrete example of such a group. Let $e = (1, 1, \ldots 1)^\top$ be the vector of all ones in $\mathbb{C}^n$, and $J = e\, e^\top$ be the matrix in $\mathbb{C}^{n \times n}$ of all ones, so J has rank 1 and $J^2 = nJ$.
The group G is defined as

$$G = \{\alpha J \mid \alpha \in \mathbb{C}, \alpha \neq 0\}$$

The multiplication results in $\alpha J \cdot \beta J = n\alpha\beta J$, and the group inverse $\mathrm{I}_G = \frac{1}{n} J$. The inverse of $A = \alpha J$ is $A^\# = \frac{1}{\alpha n^2} J$.

Proposition 3.62 (**Properties of A^D**)

1. If $\lambda \in \sigma(A)$, $\lambda \neq 0$ and $x \neq 0$ eigenvector relative to λ, so $Ax = \lambda x$. Then $\frac{1}{\lambda}$ is an eigenvalue of A^D with eigenvector x.

2. Given A, there is a polynomial $p(x)$ such thath $A^D = p(A)$.

Proof. 1) Take the core-nilpotent decomposition of A, then λ must be an eigenvalue of B, and the eigenvector x has the entries relative to N equal to 0. Then x is also an eigevector for B^{-1}, and the result follows.

2) First of all, if A is an invertible matrix, then A^{-1} can be expressed as a polynomial of A.
DA COMPLETARE □

Proposition 3.63 Let $A \in \mathbb{C}^{n\times n}$, with $\operatorname{index}(A) = k \geq 0$. Then

$$A^D = \lim_{\varepsilon\to 0} \left(A^{k+1} + \varepsilon\, \mathrm{I}\right)^{-1} A^k$$

Proof. DA COMPLETARE ☐

PROPOSIZIONE SU $AA\hat{D}$ e proiettore. DA COMPLETARE

4 Stochastic Matrices and Markov Chains

Disclaimer: this chapter is very sloppily written, and assumes that the reader has already some knowledge of Markov chains. For further (or propedeutic?) reading, we recommend the book *Markov Chains* by James R. Norris.

4.1 Basic definitions

Definition 4.1 (**Markov chain**) A *Markov chain*, or also Markov process, can be thought as a system that at discrete time intervals $t_0, t_1, t_2 \dots$ can be in one of the states s_i $i \in I$. The process can go from state s_i to s_j with probability $P_{i\,j}$, this does not depend on the states prior to s_i.
We will work mainly with finite Markov chains, so we can take $I = \{1, 2, \dots, n\}$ and P can be thought as a matrix, called *transition matrix.*

Definition 4.2 (**Stochastic matrix**) A matrix $P \in \mathbb{R}^{n\times n}$ is *stochastic* if:

1. $0 \leq p_{i\,j} \leq 1$ for every i, j. This clearly implies $P \geq 0$.
2. $\sum_{j=1}^{n} p_{i\,j} = 1$, the row sums of P are equal to 1.

Observe that a transition matrix is a stochastic matrix, since the sum of all probabilities of possible paths from s_i must be 1.

Definition 4.3 (**Doubly stochastic matrix**) P is a *doubly stochastic* matrix if both P and $P^\top$ are stochastic, i.e. also the column sums of P are equal to 1.
If P is a permutation matrix, then it is doubly stochastic. Also, if U is unitary, which means that $UU^H = U^H U = \mathrm{I}$, then P given by $p_{i\,j} = |u_{i\,j}|^2$ is doubly stochastic.
If $A \geq 0$ and has a Perron eigenvector $x > 0$, $Ax = \rho x$, then $\frac{1}{\rho}A$ is diagonally similar do a row-stochastic matrix P. Take $D = \operatorname{diag}(p)$, and $P = D^{-1}\frac{1}{\rho}AD$.

Definition 4.4 (**Probability vector**) $\pi = (\pi_1, \pi_2, \dots, \pi_n) \in \mathbb{R}^n$ is a *probability vector* (also called a probability distribution) if $\pi \geq 0$ and $\sum_{i=1}^{n} \pi_i = 1$.

Proposition 4.5 Let π be a probability vector, and P, Q be row-stochastic matrices. Then πP is a probability vector and PQ is a row-stochastic matrix.

Proof. Left as an exercise. □

Proposition 4.6 Let P be a stochastic matrix. Then $\rho(P) = 1$ and $\mathbf{1} = (1, 1, \ldots 1)^\top$ is a right Perron eigenvector for P.

Proof. The definition of a stochastic matrix is that the row sums are equal to one, i.e. $P\mathbf{1} = \mathbf{1}$, so $\rho(P) \geq 1$. Also, by theorem ?.10, we have that $\rho(P)$ is less or equal than the maximum row sum, so actually $\rho(P) = 1$ and $\mathbf{1}$ is the right Perron eigenvector for P. □

Definition 4.7 (**Stationary distribution**) A probabilty vector π is *stationary* for a stochastic matrix P if $\pi P = \pi$.
Note that $\pi P^k = \pi$ for every $k > 0$. Also note that π is a left eigenvector for P, with eigenvalue 1; there always exists such a vector by Perron-Frobenius theorem, since $\rho(P) = 1$.
If P is not an irreducible matrix, then there could be multiple linearly independent stationary vectors for P; but if P is irreducible, then by Perron-Frobenius theorem, it has a unique left eigenvector (which is positive and can be normalized to a probability vector) relative to eigenvalue 1.

Definition 4.8 (**Steady state distribution**) Let P be a stochastic matrix, and let $\{\pi^{(k)}\}_{k=0}^{\infty}$ be a sequence of probability vectors such that:

$$\pi^{(k+1)} = \pi^{(k)} P$$

A row vector π is called a *steady state distribution* (or also a *limiting distribution*) for P if for every choice of $\pi^{(0)}$, the sequence $\pi^{(k)}$ has a limit and

$$\pi = \lim_{k\to\infty} \pi^{(k)}$$

Observe that a steady state vector is necessarily a stationary vector, since $\pi = \lim_{k\to\infty} \pi^{(k)} = \lim_{k\to\infty} \pi^{(k-1)} P = \pi P$. Altough, not every stationary vector is a steady state vector, it depends on the properties of the Markov Chain P.
For example, take $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, it has a stationary vector $\pi = (\frac{1}{2}, \frac{1}{2})$ but for every $\pi^{(0)} \neq \pi$, the sequence $\pi^{(k)}$ does not converge and it alternates between $\pi^{(0)}$ and $\pi^{(1)}$.
Achtung: Some authors do not require that π is independent of initial choice $\pi^{(0)}$ to be a steady state distribution; this blurs the distinction between steady state and stationary vectors.

Theorem 4.9
Let P be a primitive (therefore irreducible), stochastic matrix. Then there exists an unique steady state distribution for P.
Also, there exists the limit $\lim_{k\to\infty} P^k = \mathbf{1}\pi$, where $\mathbf{1} = (1, \ldots, 1)^\top$ and π is the steady state distribution (row vector).

Proof. The result follows from Perron-Frobenius theorem, where π is the left Perron eigenvector of P.
We want to remark that π is a steady state distribution: for every initial distribution $\pi^{(0)}$, we have that

$$\lim_{k\to\infty} \pi^{(k)} = \lim_{k\to\infty} \pi^{(0)} P^k = \pi^{(0)} \lim_{k\to\infty} P^k = \pi^{(0)}(\mathbf{1}\pi) = (\pi^{(0)}\mathbf{1})\pi = \pi$$

$\square$

DEFINITION 4.10 Let P be stochastic, then define the $\delta(P) = \max\{\,|\lambda|\,|\lambda \in \sigma(P), \lambda \neq 1\}$.
$\delta(P)$ is related to the spectral gap of P, which is $1 - \delta(P)$. Also it controls the asintotic convergence of P^k, or the asintotic convergence of $\pi^{(k)}$ to π.
If P is primitive, then $\delta(P) < 1$. The smaller $\delta(P)$ is, the faster the convergence.

DEFINITION 4.11 Let P be the transition matrix of a Markov chain, then the Markov chain is:

- *regular* if and only if P is a primitive matrix. ($\Rightarrow \exists!$ steady state distribution)
- *ergodic* if and only if P is an irreducible matrix. ($\Rightarrow \exists!$ stationary distribution)
- *periodic* (or *cyclic*) if and only if P is an irreducible matrix with imprimitivity index $k \geq 2$.

Achtung: this terminology is not universal, in particular in Italian *ergodico* usually means *regular*.

DEFINITION 4.12 A state s_i has *access* to state s_j, if there exists a path $s_i \to s_j$ of finite length with positive probability.
If s_i, s_j have both mutual access to each other, they are called *communicating states*.
We see that we can partition all states in communicating classes.

DEFINITION 4.13 (**Absorbing state**)
A state s_i of a Markov chain is called an *absorbing state* (or also trap, sink) if the process can not exit state s. Equivalently, if P is the transition matrix, s_i is absorbing if the i^{th} row of P is all zeros, except $P_{i\,i} = 1$.

DEFINITION 4.14 (**Absorbing Markov chain**) A Markov chain is called an *absorbing chain* if there is at least one absorbing state and if for every non absorbing state, there is a positive probability of reaching an absorbing state in finite time.
Equivalently, every process in the Markov chains ends in an absorbing state with probability 1.

DEFINITION 4.15 (**Transient state**) A state s_i of a Markov chain is called *transient* if for every initial state $s(t_0)$, the state s_i will be visited only a finite number of times with probability 1.
For an absorbing Markov chain, all the non absorbent states are transient.

4.2 Reducible stochastic matrices

Now assume that the matrix P is reducible, i.e. the associated graph is not strongly connected and there are states, called *transient*, that the system can reach some finite number of times and then never again (with probability 1).
Applying the Frobenius Normal form to a stochastic matrix, we obtain:

$$P = \left(\begin{array}{cccc|ccc} P_1 & & & & & & \\ & P_2 & & & & & \\ & & \ddots & & & & \\ & & & P_g & & & \\ \hline P_{g+1,1} & \cdots & \cdots & P_{g+1,g} & P_{g+1} & & \\ \vdots & & & \vdots & \vdots & \ddots & \\ P_{n,1} & \cdots & \cdots & P_{n,g} & P_{n,g+1} & \cdots & P_t \end{array}\right)$$

Where the diagonal entries P_i $1 \le i \le t$ are all irreducible, $P_1, \ldots P_g$ are isolated components. This means the process cannot exit block P_j if it enters a state $s_i \in P_j$.
The P_i represent the communicating classes. The classes $P_1, \ldots P_g$ are the transient classes, and $P_{g+1}, \ldots P_t$ are the *recurrent* or *ergodic* classes.

Remark 4.16
A Markov Chain is absorbent $\iff$ every ergodic class consist of only one state, i.e.

$$P_i = [1] \qquad g < i \le t$$

Theorem 4.17 Let P be a stochastic matrix. Then $\lambda = 1$ is a semisimple eigenvalue with multiplicity g, which is the number of ergodic classes. Also $A = \mathrm{I} - P$ is a singular M-matrix, with $\mathrm{index}(A) = 1$ (i.e. A has *property c*).

Proof.

$$P = \left(\begin{array}{cccc|ccc} P_1 & & & & & & \\ & P_2 & & & & & \\ & & \ddots & & & & \\ & & & P_g & & & \\ \hline P_{g+1,1} & \cdots & \cdots & P_{g+1,g} & P_{g+1} & & \\ \vdots & & & \vdots & \vdots & \ddots & \\ P_{n,1} & \cdots & \cdots & P_{n,g} & P_{n,g+1} & \cdots & P_t \end{array}\right) = \left(\begin{array}{cccc|c} P_1 & & & & \\ & P_2 & & & \\ & & \ddots & & \\ & & & P_g & \\ \hline R_1 & R_2 & \cdots & R_g & C \end{array}\right)$$

Write P in the Frobenius normal form, where each diagonal block is irreducible, the first blocks $P_1, \ldots, P_g$ are isolated (the Markov process can not exit the block), and $P_{g+1}, \ldots, P_t$ are transient.
By looking at the row sums of $P_1, \ldots, P_g$, we obtain that $\rho(P_1) = \ldots = \rho(P_g) = 1$. We will prove that $\rho(P_{g+1}), \ldots \rho(P_t) < 1$: take P_i, then its rows sums can't be all equal to one (since P_i has to connect to the absorbing classes), so there exists $\widetilde{P}_i \ge P_i$ with row sums equal to 1. P_i and therefore $\widetilde{P}_i$ are irreducible, so by theorem 27 we have that $\rho(P_i) < \rho(\widetilde{P}_i) = 1$.

Another way to see that the spectral radius of transient classes $P_{g+1}, \dots P_t$ is less than 1, consider them all as one block C. Every transient state has a non zero probability to have a path of length k to a recurrent state, for sufficiently large k. Then C^k has every row sum less than 1, so $1 > \rho(C^k) = \rho(C)^k$, so $\rho(C) < 1$.

So the multiplicity of $\lambda = 1$ is given by that of $P_1, \dots P_g$, and since they are irreducible $\lambda = 1$ is semisimple, so it has index 1 and $A = I - P$ has *property c.*

□

Using this theorem, we can write P separating the eigenvalues equal to 1:

$$P = X \left(\begin{array}{c|c} \mathrm{I} & 0 \\ \hline 0 & K \end{array} \right) X^{-1}$$

With $1 \in \sigma(K)$. So define

$$A = \mathrm{I} - P = X \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & \mathrm{I} - K \end{array} \right) X^{-1}$$

This is a core-nilpotent decomposition for A, so its Drazin inverse is:

$$A^{\#} = X \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & (\mathrm{I} - K)^{-1} \end{array} \right) X^{-1}$$

Then we can define the following matrix

$$L = \mathrm{I} - (\mathrm{I} - P)(\mathrm{I} - P)^{\#} = \mathrm{I} - AA^{\#} = X \left(\begin{array}{c|c} \mathrm{I} & 0 \\ \hline 0 & 0 \end{array} \right) X^{-1}$$

Note that $1 \in \sigma(K)$ does not imply that $\rho(K) < 1$, since the ergodic classes P_i $1 \leq i \leq g$ could be periodic.

Theorem 4.18 Let P be a stochastic matrix, then:

1. P is a convergent matrix $\iff$ $\gamma(P) < 1$ $\iff$ $\rho(K) < 1$ $\iff$ all the ergodic classes $P_1, \dots, P_g$ are primitive.

2. If the Markov chain is regular, then P is convergent.

3. If the Markov chain is absorbing, then P is convergent.

Proposition 4.19 Let P be a stochastic matrix and let $S_1, \dots S_g$ be the ergodic classes corresponding to $P_1, \dots P_g$. Then to each class P_i there corresponds an unique stationary distribution π_i, such that $\pi_i(j) = 0$ if $s_j \notin S_i$ and $\pi_i(j) > 0$ otherwise.
Every stationary distribution π for P, is a linear combination of π_i:

$$\pi = \sum_{i=1}^{g} a_i \pi_i \qquad \text{with } \; a_i \geq 0 \; \sum_{i=1}^{h} a_i = 1$$

Proof. The first part follows from Perron-Frobenius theorem applied to each primitive matrix P_i.
The eigenspace relative to eigenvalue 1 has dimension g, and $\pi_1, \dots \pi_g$ are g linearly indepent eigenvectors. This implies that every other stationary distribution has to be in their span. $\square$

PROPOSITION 4.20 Let P be a stochastic matrix, and A, L defined as above. Then L can be calculated as:

$$L = \lim_{k\to\infty} \frac{1}{k} \left(\mathrm{I} + P + P^2 + \ldots + P^{k-1}\right) \qquad \text{Cesaro-summability}$$

Also for each $\alpha \in (0,1)$:

$$L = \lim_{k\to\infty} \left((1-\alpha)\,\mathrm{I} + \alpha P\right)^k \qquad \text{Euler-summability}$$

Also, if P is convergent, then:

$$L = \lim_{k\to\infty} P^k$$

The first two formulas are used in the case P is periodic (or one of the ergodic classes P_i in the normal form is periodic): we have to cancel periodicity either by averaging, or by transforming P_i in $T_\alpha = (1-\alpha)\,\mathrm{I} + \alpha P_i$, which is a primitive matrix. Also, $\delta(T_\alpha) < 1$, because the spectrum T_α can be obtained from the spectrum of P_i by an homothety in 1 and ratio α.
In the case that all ergodic classes are primitive, then the stronger result from the third formula holds.

PROPOSITION 4.21 Let $A = \mathrm{I} - P$, $L = \mathrm{I} - AA^{\#}$. If s_i is a transient state, then the i^{th} row of L is zero, i.e. $Le_i = 0$.

Proof. Si fa con la forma triangolare a blocchi. DA COMPLETARE $\square$

PROPOSITION 4.22 Le s_i, s_j be transient states. Then $[A^{\#}]_{ij}$ is the expected values of the number of times the process with initial state s_i visits state s_j.

4.3 Absorbing Markov chain

PROPOSITION 4.23 (**Normal form for absorbing Markov chains**)
If the transition matrix P has size n, and there are t transient states and $r = n - t$ absorbing states, by putting all the absorbing states at the end we obtain a normal form for P:

$$P = \left(\begin{array}{c|c} Q & R \\ \hline 0 & \mathrm{I}_r \end{array}\right)$$

Then we can calculate the powers of P:

$$P^k = \left(\begin{array}{c|c} Q^k & R_k \\ \hline 0 & \mathrm{I}_r \end{array}\right)$$

Where $R_k = (\mathrm{I} + Q + Q^2 + \dots Q^{k-1})R$, and $R_{k+1} \geq R_k$.
Also, $\rho(Q) < 1$ and there exists the limit

$$\lim_{k \to \infty} P^k = \left(\begin{array}{c|c} 0 & (\mathrm{I} - Q)^{-1} R \\ \hline 0 & \mathrm{I}_r \end{array}\right)$$

Proof. The proof that $\rho(Q) < 1$ is a special case of ?.17.
Then $I - Q$ is a non-singular M-matrix, so $\mathrm{I} + Q + Q^2 + \dots = (I - Q)^{-1} \geq 0$.

□

The matrix $(\mathrm{I} - Q)^{-1}$ is called the *fundamental matrix* of the Markov chain. $I - Q$ is a non-singular M-matrixsince $\rho(Q) < 1$, so $(I - Q)^{-1} \geq 0$.

PROPOSITION 4.24 If P is an absorbing Markov chain in the above normal form, then the stationary distributions are of the form $\pi = (0|\pi_2)$, where π_2 is a probability vector of size r.

Proof. Let $\pi = (\pi_1|\pi_2)$ be a stationary vector, so:

$$\pi P = (\pi_1|\pi_2) \left(\begin{array}{c|c} Q & R \\ \hline 0 & \mathrm{I}_r \end{array}\right) = (\pi_1 Q \,|\, \pi_1 R + \pi_2)$$

Since $\rho(Q) < 1$, then $\pi_1 = \pi_1 Q$ is possible only if $\pi_1 = 0$. Then every choice of π_2 satisfies this equation.

□

PROPOSITION 4.25 Let $N = (\mathrm{I} - Q)^{-1}$ be the fundamental matrix of an absorbing Markov chain. Then

1. $N_{i\,j}$ è il numero di volte che passa dal transiente s_i al transiente s_j.
2. $[N\mathbf{1}]_i$, the sum of i^{th} row of N is the expected number of steps needed to get to an absorbent state starting from s_i.
3. $NR_{i\,j}$ represents the probability that the system will end in state s_j by passing in s_i as the last transient state. DA COMPLETARE

Questa proposizione è la stessa cosa di $A^{\#}$ detta in generale, e può essere inglobata li.

4.4 Ergodic Markov chain

PROPOSITION 4.26 If the Markov chain is ergodic, then there exists a unique stationary distribution $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$, which is positive $\pi > 0$.
Also $m_i = \dfrac{1}{\pi_i}$ is the expected time of return to state s_i.

DEFINITION 4.27 (**Mean first passage time**) Let $m_{i\,j}$ be the expect number of steps of a path from s_i to s_j.
The the matrix $M = m_{i\,j}$ is the mean first passage matrix.
Observe that this means $m_{i\,i} = \dfrac{1}{\pi_i}$.

THEOREM 4.28 (**Meyer**) Let P be the transition matrix of an ergodic Markov chain and let $\pi > 0$ be it's stationary distribution. Let $J = \mathbf{1} \cdot \mathbf{1}^\top$ be the matrix of all 1s, and $A = \mathrm{I} - P$.
Let $D = \mathrm{diag}(\frac{1}{\pi_1}, \frac{1}{\pi_2}, \ldots, \frac{1}{\pi_n})$, then:

1. $D^{-1} = \mathrm{diag}(\mathrm{I} - AA^{\#})$
2. $M = (\mathrm{I} - A^{\#} + J\mathrm{diag}(A^{\#}))D$

As an example, let's calculate the diagonal entries of M:

$$m_{i\,i} = \frac{1}{\pi_i} - \frac{a^{\#}_{i\,i}}{(I - AA^{\#})_{i\,i}} + \frac{a^{\#}_{i\,i}}{(I - AA^{\#})_{i\,i}} = \frac{1}{\pi_i}$$

So the theorem is true for the diagonal entries.

THEOREM 4.29 (**Meyer**) Let P be the transition matrix of an ergodic Markov chain, π be the (unique) stationary vector and $A, A^{\#}, M$ defined as above. Then M satisfies:

$$[\pi M]_i = 1 + \frac{a^{\#}_{i\,i}}{\pi_i} \qquad 1 \leq i \leq n$$

This represents the expected time to arrive in state s_i starting from a random state with probability given by π.

THEOREM 4.30 (**Golub - Meyer, 1986**) Questo è un teorema su piccole perturbazioni della matrice stocastica.

THEOREM 4.31 (**Normal form for stochastic convergent matrices**)

5 Graphs, walks and centrality measures

Page rank. Descrizione dell'algoritmo
Pagerank, Degree, Closeness, Betweenness, Eigenvector.
Katz centrality, Hub and authority,

subgrpah centrality (does not distinguish hubs and authorities)
total communicability (hub and authority)
Articolo di Benzi e Klymko sull limite di Katz e subgraph a degree e eigenvector.
Laplaciano di un grafo.
Laplaciano normalizzato.
Misure di connettività

6 Economia, Leontief model

Carino.
Poi aggiungere le considerazioni finali.

7 todo

DA COMPLETARE Osservazioni sparse da inserire:

- Corollario di Lappo-danilevsky delle sottomatrici
- Se A è NSMM, se ho $B \geq A$ allora anche B è NSMM.
- Fare un mega teoremone TFAE per le matrici NSMM
- Se A è NSMM, B è NSMM e AB è Z-matrice, allora AB è NSMM.
- Esercizio proiettore di AA^D.
- matrice substocastica

$A\mathbf{1} = v = \mathbf{1}^\top B$
Libro Berman - Plemons, applicazioni su M-matrici e condizioni nec e suff per Z-matrici.