

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

SCUOLA DI SCIENZE
Corso di Laurea in Matematica

Tesi di Laurea in Geometria

**Teorema di Perron-Frobenius
e utilizzo nello studio
dell'andamento del traffico**

Relatore:
Chiar.mo Prof.
Stefano Francaviglia

Presentata da:
Giorgia Petracci

Sessione Unica
Anno Accademico 2017/2018

Indice

| | |
|--|-----------|
| Introduzione | i |
| 1 Prime definizioni e Teorema di Perron | 1 |
| 1.1 Norma di una matrice | 1 |
| 1.2 Raggio spettrale | 3 |
| 1.3 Matrici non negative | 5 |
| 1.4 Teorema di Perron | 12 |
| 2 Teorema di Perron-Frobenius | 21 |
| 2.1 Estensione del teorema di Perron a matrici non-negative | 21 |
| 2.2 Irriducibilità, grafi e teorema di Perron-Frobenius | 23 |
| 2.3 Matrici Primitive | 29 |
| 3 Modellizzazione rete viaria e previsione andamento del traffico | 33 |
| 3.1 Modellizzazione rete stradale | 34 |

Introduzione

Il Teorema di Perron-Frobenius deve il suo nome ai matematici tedeschi Oskar Perron (7 Maggio 1880 - 22 Novembre 1975) e Ferdinand Georg Frobenius (26 Ottobre 1849 - 3 Agosto 1917). Fu Perron che per primo si dedicò allo studio di matrici quadrate a entrate positive, e pubblicò nel 1907 il ‘Teorema di Perron’, in cui si dimostra che tali matrici hanno un unico autovalore positivo di modulo massimo, chiamato radice di Perron, e tutti gli altri autovalori diversi da lui hanno modulo minore. L’autovettore associato alla radice di Perron ha componenti tutte positive e ogni altro autovettore a componenti non negative della matrice considerata è un suo multiplo positivo. Già dal 1908 Frobenius si dedicò alla estensione di questi risultati a matrici quadrate a entrate non negative; il suo studio culminò nel 1912 con la pubblicazione del ‘Teorema di Perron-Frobenius’.

In questa tesi cercheremo di spiegare nel modo più chiaro possibile come si giunse a tali risultati. Nel primo capitolo daremo alcune definizioni preliminari e ne richiameremo altre già note dal percorso triennale. Indagheremo le proprietà del raggio spettrale e degli autovettori ad esso associati. Solo alla fine enunciamo il Teorema di Perron.

Nel secondo capitolo dapprima individuiamo tutte le proprietà delle matrici positive che possono essere estese a matrici non negative utilizzando il concetto di limite. Definiamo, poi, nuovi oggetti, come i grafi, che ci permettono di dimostrare importanti caratteristiche delle matrici irriducibili (già definite nel primo capitolo) e alla fine illustriamo il Teorema di Perron-Frobenius vero e proprio. Infine diamo la definizione di matrice primitiva, il che permette alle matrici non-negative di avere esattamente tutte le stesse proprietà delle matrici positive.

Benchè poco conosciuto, questo teorema ha molte applicazioni, per esempio nell’algoritmo di PageRank utilizzato da Google per la ricerca delle pagine, o nel modello di Leslie per prevedere la crescita di una popolazione, etc... Nel terzo, ed ultimo, capitolo facciamo vedere come si applica la teoria di Perron-Frobenius allo studio dei sistemi dinamici, indagando lo sviluppo del traffico all’interno di

una rete stradale con numero di macchine costante.

Capitolo 1

Prime definizioni e Teorema di Perron

In questo primo capitolo andremo a dare definizioni che ci accompagneranno per tutta la trattazione, come ad esempio quella di matrice non negativa, ed enunceremo e dimostreremo una serie di teoremi i cui risultati andranno a costituire l'enunciato del teorema di Perron, che quindi esporremo solo alla fine. Tratteremo solo con matrici quadrate di dimensione finita $n \times n$, che nella maggior parte dei casi saranno a coefficienti reali; tuttavia, per un'ulteriore chiarezza, esplicheremo sempre il campo in cui lavoriamo di volta in volta.

1.1 Norma di una matrice

Definizione 1.1.1 (Norma Matriciale). Una *norma matriciale* è una funzione $\|\cdot\| : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ che soddisfi, per ogni $A, B \in \mathbb{R}^{n \times n}$, le seguenti proprietà:

- $\|A\| \geq 0$ e $\|A\| = 0 \iff A = 0$,
- $\|\lambda A\| = |\lambda| \|A\|$ per ogni scalare λ ,
- $\|A + B\| \leq \|A\| + \|B\|$,
- $\|AB\| \leq \|A\| \|B\|$.

Definizione 1.1.2 (Norma di Frobenius). La norma di Frobenius di una matrice

$A \in \mathbb{R}^{n \times n}$ è definita come

$$\|A\|_F := \sqrt{\sum_{i,j=1}^n |a_{ij}|^2}$$

Osserviamo che $\|A\|_F^2 = \text{tr}(A^T A)$. Con $\text{tr}(A^T A)$ intendo la traccia di $A^T A$, ossia la somma di tutti gli elementi diagonali di $A^T A$.

Definizione 1.1.3 (Norma matriciale indotta). Una norma vettoriale $\|\cdot\|$ definita su \mathbb{R} induce una norma matriciale su $\mathbb{R}^{n \times n}$ così definita:

$$\|A\| := \max_{\|x\|=1} \|Ax\|$$

con $A \in \mathbb{R}^{n \times n}$

Una proprietà che segue direttamente dalla definizione di norma matriciale indotta è

$$\|Ax\| \leq \|A\| \|x\|.$$

Andiamo a vedere alcuni esempi di particolare interesse, che ci serviranno per le successive dimostrazioni. Ricordiamo che la p -norma di un vettore è

$$\|x\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$$

con $p \in [1, +\infty)$.

Si può estendere il concetto di p -norma utilizzando il limite di una funzione: la ∞ -norma è il massimo dei valori delle componenti in valore assoluto

$$\|x\|_\infty := \lim_{p \rightarrow \infty} \|x\|_p = \max_{i=1, \dots, n} |x_i|$$

Proposizione 1.1.4 (1-Norma e ∞ -Norma). *Le norme matriciali indotte dalla 1-norma e ∞ -norma vettoriale sono le seguenti:*

$$\|A\|_1 = \max_{\|x\|_1=1} \|Ax\|_1 = \max_j \sum_{i=1}^n |a_{ij}| \quad (1.1)$$

$$\|A\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty = \max_i \sum_{j=1}^n |a_{ij}| \quad (1.2)$$

Dimostrazione. Dimostriamo la (1.1)

$\forall x$ con $\|x\|_1 = 1$ vale

$$\begin{aligned} \|Ax\|_1 &= \sum_{i=1}^n |(Ax)_i| = \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \\ &\leq \sum_{i=1}^n \sum_{j=1}^n |a_{ij}| |x_j| = \sum_{j=1}^n (|x_j| \sum_{i=1}^n |a_{ij}|) \leq \\ &\leq \sum_{j=1}^n |x_j| \left(\max_{j=1, \dots, n} \sum_{i=1}^n |a_{ij}| \right) = \max_{j=1, \dots, n} \sum_{i=1}^n |a_{ij}|. \end{aligned}$$

Ora la (1.2)

$\forall x$ con $\|x\|_\infty = 1$ vale

$$\begin{aligned} \|Ax\|_\infty &= \max_{i=1,\dots,n} |(Ax)_i| = \max_{i=1,\dots,n} \left| \sum_{j=1}^n a_{ij}x_j \right| \leq \\ &\leq \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}| |x_j| \leq \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}|. \end{aligned}$$

□

1.2 Raggio spettrale

Definizione 1.2.1 (Raggio spettrale). Definiamo *raggio spettrale della matrice* A

$$\rho(A) = \max_{\lambda \in \text{Spec}(A)} |\lambda|$$

con ‘ $\text{Spec}(A)$ ’ intendiamo l’insieme di tutti gli autovalori della matrice A .

Teorema 1.2.2 (Hirs). *Sia $A \in \mathbb{R}^{n \times n}$ e sia $\|\cdot\|$ una norma matriciale, allora per ogni norma matriciale vale*

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

Dimostrazione. Siano (λ, x) un’ autocoppia di A , $x \neq 0$, sia $X = [x|0|\dots|0] \neq 0$, $X \in \mathbb{R}^{n \times n}$, allora $\lambda X = AX \Rightarrow \|\lambda X\| = \|AX\|$, osserviamo che $\|\lambda X\| = |\lambda| \|X\|$ e $\|AX\| \leq \|A\| \|X\| \Rightarrow |\lambda| \|X\| \leq \|A\| \|X\|$ quindi $|\lambda| \leq \|A\|$. Ciò vale per ogni λ , quindi anche per quello che massimizza il modulo □

Teorema 1.2.3. *Data una matrice $A \in \mathbb{R}^{n \times n}$ e dato $\epsilon > 0$, allora esiste una norma matriciale $\|\cdot\|$ per cui vale*

$$\|A\| \leq \rho(A) + \epsilon$$

Dimostrazione. Data una norma vettoriale $\|\cdot\|$ e una matrice invertibile S , la applicazione $x \rightarrow \|Sx\|$ è una norma vettoriale. Sia $\|A\|_S$ la norma matriciale indotta da questa norma vettoriale, allora

$$\|A\|_S := \max_{\|x\|_S=1} \|Ax\|_S = \max_{\|Sx\|=1} \|SAx\| = \max_{\|y\|=1} \|SAS^{-1}y\| = \|SAS^{-1}\|.$$

Nel penultimo passaggio è stata fatta la sostituzione $Sx = y$. Abbiamo quindi $\|A\|_S = \|SAS^{-1}\|$.

Sia J la matrice in forma di Jordan della matrice A data, ossia $J = R^{-1}AR$, con R e R^{-1} matrici invertibili. Sia $\epsilon > 0$, consideriamo la matrice diagonale

$D_\epsilon = \text{diag}(\epsilon, \epsilon^2, \epsilon^3, \dots, \epsilon^n)$ di dimensione $n \times n$ e la matrice $\hat{J} = D_\epsilon^{-1} J D_\epsilon$, \hat{J} è una matrice diagonale a blocchi, con ogni blocco del tipo

$$\begin{pmatrix} \lambda & \epsilon & 0 & \dots & 0 \\ 0 & \lambda & \epsilon & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \lambda & \epsilon \\ 0 & \dots & 0 & 0 & \lambda \end{pmatrix}$$

quindi se l'autovalore λ che massimizza il modulo sta in un blocco di Jordan di dimensione maggiore di 1, vale $\|\hat{J}\|_\infty = \max_{\lambda \in \text{Spec}(A)} (|\lambda| + \epsilon)$, se invece tutti gli autovalori λ che massimizzano il modulo stanno in blocchi di Jordan di dimensione 1 e se ϵ è scelto abbastanza piccolo tale che $|\lambda| + \epsilon \leq \rho(A)$ per ogni altro autovalore λ , allora $\|\hat{J}\|_\infty = \rho(A)$.

Il teorema è verificato prendendo $\|A\| = \|A\|_S$, con $S = D_\epsilon^{-1} R^{-1}$.

Inoltre $\|A\| = \rho(A)$ se tutti gli autovalori che massimizzano il modulo stanno in blocchi di dimensione 1.

□

Proposizione 1.2.4 (Convergenza a zero). *Sia $A \in \mathbb{R}^{n \times n}$, allora*

$$\lim_{k \rightarrow \infty} A^k = 0 \iff \rho(A) < 1.$$

Dimostrazione. Supponiamo che $A^k \rightarrow 0$ per $k \rightarrow \infty$ e sia $x \neq 0$ un vettore tale che $Ax = \lambda x$, allora $A^k x = \lambda^k x \rightarrow 0 \iff |\lambda| < 1$ questo vale per ogni autovalore λ , quindi anche per $\rho(A) \leq 1$.

Supponiamo ora $\rho(A) \leq 1$. Dai due teoremi appena dimostrati sappiamo che esiste una qualche norma matriciale per cui

$$\rho(A) \leq \|A\| \leq \rho(A) + \epsilon \quad \forall \epsilon > 0,$$

quindi $\|A\| < 1$. Per definizione di norma matriciale vale $\|A^k\| \leq \|A\|^k$, poiché $\|A\| < 1$, $\|A\|^k \rightarrow 0$ per $k \rightarrow \infty$, per il teorema del confronto anche $\|A^k\| \rightarrow 0$ e di conseguenza $A^k \rightarrow 0$ per $k \rightarrow \infty$, rispetto la norma $\|\cdot\|$. Facciamo vedere che vale per ogni norma: su $\mathbb{R}^{n \times n}$ tutte le norme matriciali sono equivalenti, cioè per ogni coppia di norme $\|\cdot\|$ e $\|\cdot\|'$ esistono due costanti $\alpha, \beta > 0$ tale che $\alpha\|X\| \leq \|X\|' \leq \beta\|X\| \quad \forall X \in \mathbb{R}^{n \times n}$. Allora $\alpha\|A^k\| \leq \|A^k\|' \leq \beta\|A^k\|$. Se $\|A^k\| \rightarrow 0$ i termini destro e sinistro della disuguaglianza tendono a zero, quindi, per il teorema dei due carabinieri, anche $\|A^k\|' \rightarrow 0$. Allora $A^k \rightarrow 0$ rispetto ogni norma: solo adesso posso scrivere $A^k \rightarrow 0$. □

Lemma 1.2.5. Sia $A \in \mathbb{R}^{n \times n}$, allora $\rho(A)^k = \rho(A^k)$, $\forall k = 1, \dots, n$.

Dimostrazione. Innanzitutto osserviamo che se $\lambda_1, \dots, \lambda_n$ sono gli autovalori di A , allora gli autovalori di A^k sono $\lambda_1^k, \dots, \lambda_n^k$.

Utilizziamo la definizione di $\rho(A)$:

$$\rho(A)^k := \left(\max_{\lambda \in \text{Spec}(A)} |\lambda| \right)^k = \max_{\lambda \in \text{Spec}(A)} |\lambda|^k = \max_{\lambda^k \in \text{Spec}(A^k)} |\lambda|^k =: \rho(A^k). \quad \square$$

Teorema 1.2.6. Per ogni norma matriciale $\|\cdot\|$ e per ogni matrice A vale

$$\lim_{k \rightarrow \infty} \|A^k\|^{\frac{1}{k}} = \rho(A).$$

Dimostrazione. Dal teorema di Hirsh $\rho(A) \leq \|A\|$ per ogni norma di matrice.

$\rho(A)^k = \rho(A^k) \leq \|A^k\|$ quindi $\rho(A) \leq \|A^k\|^{\frac{1}{k}}$.

Ora consideriamo la matrice $\frac{1}{\rho(A) + \epsilon} A$, il suo raggio spettrale

$$\rho\left(\frac{1}{\rho(A) + \epsilon} A\right) < 1 \quad \forall \epsilon > 0; \text{ dalla proposizione appena dimostrata}$$

$$\lim_{k \rightarrow \infty} \left(\frac{1}{\rho(A) + \epsilon} A\right)^k = 0 \Rightarrow \lim_{k \rightarrow \infty} \frac{1}{(\rho(A) + \epsilon)^k} A^k = 0 \Rightarrow \lim_{k \rightarrow \infty} \frac{1}{(\rho(A) + \epsilon)^k} \|A\|^k = 0.$$

Cioè $\exists K_\epsilon > 0, K_\epsilon \in \mathbb{Z}$ tale che $\frac{1}{(\rho(A) + \epsilon)^k} \|A\| < 1 \quad \forall k \geq K_\epsilon$.

Allora $\rho(A) \leq \|A^k\|^{\frac{1}{k}} \leq \rho(A) + \epsilon \quad \forall \epsilon > 0$, quindi

$$\lim_{k \rightarrow \infty} \|A^k\|^{\frac{1}{k}} = \rho(A).$$

Quanto appena dimostrato non dipende dalla norma, poichè abbiamo utilizzato tutte proposizioni e teoremi i cui risultati erano indipendenti dalla norma utilizzata. \square

1.3 Matrici non negative

Definizione 1.3.1. Siano $A, B \in \mathbb{R}^{n \times n}$, $A = (a_{ij}), B = (b_{ij})$

- $A > 0 \iff a_{ij} > 0 \quad \forall i, j = 1, \dots, n$
- $A \geq 0 \iff a_{ij} \geq 0 \quad \forall i, j = 1, \dots, n$
- $A \geq B \iff a_{ij} \geq b_{ij} \quad \forall i, j = 1, \dots, n$
- $A > B \iff a_{ij} > b_{ij} \quad \forall i, j = 1, \dots, n$

Le relazioni \leq e $<$ tra matrici sono definite analogamente .

Definizione 1.3.2. Sia $A \in \mathbb{R}^{n \times n}$, A è una **matrice non-negativa** se $A \geq 0$. Analogamente A è una **matrice positiva** se $A > 0$.

Definizione 1.3.3. Sia $A \in \mathbb{C}^{n \times n}$, con $|A|$ intendiamo la matrice che ha come elementi i moduli dei rispettivi elementi della matrice A di partenza. Cioè $|A| = (|a_{ij}|)_{0 \leq i, j \leq n}$.

Definizione 1.3.4 (Matrice riducibile). Sia A una matrice, $A \in \mathbb{R}^{n \times n}$, A è detta **riducibile** se vale una delle seguenti condizioni:

- se $n = 1$ e $A = 0$,
- se $n \geq 2$, allora esiste una matrice di permutazione $P \in \mathbb{R}^{n \times n}$ ed esiste un intero r , con $1 \leq r \leq n$ tale che

$$P^T A P = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}$$

Dove $X \in \mathbb{R}^{r \times r}$, $Z \in \mathbb{R}^{n-r \times n-r}$, $0 \in \mathbb{R}^{n-r \times r}$, $Y \in \mathbb{R}^{r \times n-r}$.

In questa definizione non si danno condizioni sulle entrate dei blocchi X, Y, Z , ad esempio non ci si sofferma sulla presenza o meno di eventuali entrate nulle; il nocciolo della questione è ottenere un blocco di dimensione $n - r \times r$ di zeri, permutando le righe o le colonne della matrice A di partenza.

Detto ciò, se A è positiva, sicuramente non è riducibile.

Definizione 1.3.5. Sia $A \in \mathbb{R}^{n \times n}$, A è detta **irriducibile** se non è riducibile.

Per capire meglio il concetto di matrice irriducibile diamone una caratterizzazione attraverso gli spazi invariati: sia $\mathcal{B} = \{e_1, \dots, e_n\}$ la base canonica di \mathbb{R}^n , definiamo *sottospazio-coordinato* un sottospazio $\mathcal{S} := \text{span}(e_{i_1}, \dots, e_{i_k})$ con $1 \leq i_1, \dots, i_k \leq n$, cioè $\{e_{i_1}, \dots, e_{i_k}\}$ è sottoinsieme proprio della base canonica \mathcal{B} . Per dirlo in altre parole un *sottospazio-coordinato* è un sottospazio definito su un sottoinsieme proprio di vettori della base canonica \mathcal{B} .

A è una matrice irriducibile se non ha sottospazi-coordinati invarianti all'infuori di quelli banali.

Osservazione 1.3.6. Per ogni $A, B \in \mathbb{R}^{n \times n}, \forall x, y \in \mathbb{R}^n, \forall k = 1, \dots, n$ vale

$$|A^k| \leq |A|^k, \tag{1.3}$$

$$\|A\|_\infty = \||A|\|_\infty, \tag{1.4}$$

$$0 \leq A \leq B \Rightarrow 0 \leq A^k \leq B^k, \tag{1.5}$$

$$|A| \leq |B| \Rightarrow \|A\|_\infty \leq \|B\|_\infty, \tag{1.6}$$

$$A > 0, \quad x \geq 0, \quad x \neq 0 \Rightarrow Ax > 0, \tag{1.7}$$

$$A \geq 0, \quad A \neq 0, \quad x > y > 0 \Rightarrow Ax > Ay. \tag{1.8}$$

Dimostrazione. La (1.3) deriva, si può dire direttamente, dalla disuguaglianza triangolare e dalla definizione di prodotto tra matrici. Dimostriamola per induzione:

Passo base:

per $k = 1$ vale $|A| \leq |A|$? Sì, possiamo procedere.

Passo induttivo:

supponiamo che valga per tutti i $k = 1, \dots, n$, vogliamo vedere se vale anche per $k = n + 1$.

$|A^{n+1}| = |A^n A|$. Indichiamo con \hat{a}_{ij} gli elementi di A^n .

$$\begin{aligned} |A^n A| &= \left| \left(\sum_{k=1}^n \hat{a}_{ik} a_{kj} \right)_{1 \leq i, j \leq n} \right| = \left(\left| \sum_{k=1}^n \hat{a}_{ik} a_{kj} \right| \right)_{1 \leq i, j \leq n} \leq \\ &\leq \left(\sum_{k=1}^n |\hat{a}_{ik}| |a_{kj}| \right)_{1 \leq i, j \leq n} = |A^n| |A|. \end{aligned}$$

Per ipotesi induttiva vale $|A^n| \leq |A|^n$, quindi $|A^n| |A| \leq |A|^n |A| = |A|^{n+1}$.

Continuiamo con la (1.4): dalla Proposizione 1.1.4, sappiamo che

$\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$, allora, considerando che gli elementi di $|A|$ sono del tipo

$|a_{ij}|$, con $i, j = 1, \dots, n$, $\| |A| \|_\infty = \max_i \sum_{j=1}^n |(a_{ij})| = \max_i \sum_{j=1}^n |a_{ij}|$, confrontando

gli ultimi termini delle due catene di uguaglianze deduciamo $\|A\|_\infty = \| |A| \|_\infty$.

Dimostriamo adesso la (1.5): anche qui useremo l'induzione:

Passo base:

per $k = 1$ vale $0 \leq A \leq B \Rightarrow 0 \leq A \leq B$? Sì, possiamo procedere.

Passo induttivo:

supponiamo che la nostra tesi valga per tutti i $k = 1, \dots, n$, vogliamo vedere se vale anche per $k = n + 1$.

$A^{k+1} = A^k A$, $B^{k+1} = B^k B$, per ipotesi induttiva vale $A^k \leq B^k$, cioè, indicando con \hat{a}_{ij} e \hat{b}_{ij} gli elementi di A^k e B^k , abbiamo

$$\hat{a}_{ij} \leq \hat{b}_{ij} \quad \forall i, j = 1, \dots, n \quad \Rightarrow \quad \hat{a}_{ij} - \hat{b}_{ij} \leq 0 \quad \Rightarrow \quad \sum_{k=1}^n \hat{a}_{ik} - \hat{b}_{ik} \leq 0 \quad \forall i = 1, \dots, n.$$

L'ultima disuguaglianza vale in quanto somma finita di soli termini negativi.

Analogamente dall'ipotesi $0 \leq A \leq B$ deriva $a_{ij} \leq b_{ij}$ e $a_{ij}, b_{ij} \geq 0$

$\forall i, j = 1, \dots, n$. Grazie al fatto che $a_{ij} \geq 0$ vale $\sum_{k=1}^n \hat{a}_{ij} a_{kj} - \hat{b}_{ij} a_{kj} \leq 0$; quindi

$\sum_{k=1}^n \hat{a}_{ij} a_{kj} \leq \sum_{k=1}^n \hat{b}_{ij} a_{kj} \leq \sum_{k=1}^n \hat{b}_{ij} b_{kj} \quad \forall i, j = 1, \dots, n.$ Abbiamo scritto l'ultimo passaggio grazie al fatto che $a_{ij} \leq b_{ij} \quad \forall i, j = 1, \dots, n.$ Il primo e l'ultimo termine della catena di disuguaglianze altro non sono, rispettivamente, che gli elementi ij delle matrici A^{k+1} e B^{k+1} , quindi $0 \leq A^{k+1} \leq B^{k+1}.$

La (1.6):

$|A| \leq |B|$ significa che $|a_{ij}| \leq |b_{ij}| \quad \forall i, j = 1, \dots, n,$ quindi

$$|a_{ij}| - |b_{ij}| \leq 0 \Rightarrow \sum_{j=1}^n |a_{ij}| - |b_{ij}| \leq 0 \quad \forall i = 1, \dots, n,$$

in quanto somma finita di numeri negativi. $\sum_{j=1}^n |a_{ij}| - |b_{ij}| \leq 0 \Rightarrow \sum_{j=1}^n |a_{ij}| \leq \sum_{j=1}^n |b_{ij}|$ questo vale $\forall i = 1, \dots, n,$ quindi anche per quella i che massimizza

la somma, cioè $\max_i \sum_{j=1}^n |a_{ij}| \leq \max_i \sum_{j=1}^n |b_{ij}|,$ cioè $\|A\|_\infty \leq \|B\|_\infty.$

La (1.7):

Scriviamo esplicitamente il vettore $Ax = \left(\sum_{j=1}^n a_{ij} x_j \right)_{1 \leq i \leq n}.$ Per ipotesi ogni $a_{ij} \neq 0,$ quindi la i -esima componente del vettore Ax per essere uguale a zero deve avere $x_j = 0 \quad \forall j = 1, \dots, n,$ cioè $x = 0;$ ma contraddice l'ipotesi $x \neq 0.$

Per finire la (1.8):

Scriviamo esplicitamente Ax e Ay e andiamo a valutare la loro differenza, cioè la differenza delle componenti i -esime.

$$Ax - Ay = \left(\sum_{j=1}^n a_{ij} x_j \right)_{1 \leq i \leq n} - \left(\sum_{j=1}^n a_{ij} y_j \right)_{1 \leq i \leq n} \iff$$

$$(Ax)_i - (Ay)_i = \sum_{j=1}^n a_{ij} x_j - \sum_{j=1}^n a_{ij} y_j = \sum_{j=1}^n a_{ij} (x_j - y_j) > 0 \quad \forall i = 1, \dots, n.$$

Quindi $Ax > Ay.$

Possiamo dire $\sum_{j=1}^n a_{ij} (x_j - y_j) > 0$ perchè per ipotesi $x_j - y_j > 0 \quad \forall j = 1, \dots, n$ e gli a_{ij} sono non tutti nulli. \square

Le affermazioni (1.3), (1.4) e (1.6) valgono anche per $A, B \in \mathbb{C}^{n \times n}.$ Rifare le dimostrazioni nel caso complesso non serve perchè sono uguali a quelle nel caso reale.

Teorema 1.3.7. Siano $A \in \mathbb{C}^{n \times n}$, $B \in \mathbb{R}^{n \times n}$, se $|A| \leq B$, allora

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

Dimostrazione. Tutto ciò che utilizzeremo in questa dimostrazione deriva dall'osservazione appena dimostrata.

Per ipotesi, $|A| \leq B$, allora $|A|^k \leq B^k$ (dalla (1.5) dell'osservazione precedente).

$$\begin{aligned} \|A^k\|_\infty &= \||A^k|\|_\infty \leq \||A|^k\|_\infty \leq \|B^k\|_\infty \\ &\Rightarrow \|A^k\|_\infty^{\frac{1}{k}} \leq \||A|^k\|_\infty^{\frac{1}{k}} \leq \|B^k\|_\infty^{\frac{1}{k}} \\ &\Rightarrow \lim_{k \rightarrow \infty} \|A^k\|_\infty^{\frac{1}{k}} \leq \lim_{k \rightarrow \infty} \||A|^k\|_\infty^{\frac{1}{k}} \leq \lim_{k \rightarrow \infty} \|B^k\|_\infty^{\frac{1}{k}} \\ &\Rightarrow \rho(A) \leq \rho(|A|) \leq \rho(B). \end{aligned}$$

L'ultimo passaggio deriva direttamente dal Teorema 1.2.6. □

Corollario 1.3.8. Siano $A, B \in \mathbb{R}^{n \times n}$, se $0 \leq A \leq B$, allora $\rho(A) \leq \rho(B)$.

Lemma 1.3.9. Sia $A \in \mathbb{R}^{n \times n}$, $A \geq 0$. Se le somme degli elementi di ciascuna riga di A sono costanti, allora $\rho(A) = \|A\|_\infty$. Se le somme degli elementi di ciascuna colonna di A sono costanti, allora $\rho(A) = \|A\|_1$.

Dimostrazione. Dal Teorema 1.2.2 $\rho(A) \leq \|A\|$ per qualsiasi norma matriciale $\|\cdot\|$. Dalla Proposizione 1.1.4 sappiamo che $\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$, ma in questo

caso $a_{ij} = |a_{ij}|$ e $\sum_{j=1}^n a_{ij} = c \in \mathbb{R} \quad \forall i = 1, \dots, n$; quindi $\|A\|_\infty = c$.

Inoltre il vettore $x = (1, 1, 1, \dots, 1)^T$ è un autovettore di A con autovalore c , allora $c \leq \rho(A) = \|A\|_\infty = c \Rightarrow \rho(A) = c = \|A\|_\infty$. La dimostrazione nel caso in cui le somme delle colonne di A siano costanti è analoga a quella appena vista, utilizzando A^T a posto di A . □

Teorema 1.3.10. Sia $A \in \mathbb{R}^{n \times n}$, e $A \geq 0$, allora

$$\min_{i=1, \dots, n} \sum_{j=1}^n a_{ij} \leq \rho(A) \leq \max_{i=1, \dots, n} \sum_{j=1}^n a_{ij}; \quad (1.9)$$

$$\min_{j=1, \dots, n} \sum_{i=1}^n a_{ij} \leq \rho(A) \leq \max_{j=1, \dots, n} \sum_{i=1}^n a_{ij}. \quad (1.10)$$

Dimostrazione. Sia $\alpha = \min_{i=1, \dots, n} \sum_{j=1}^n a_{ij}$ e costruiamo una nuova matrice B tale

che $A \geq B \geq 0$ e $\sum_{j=1}^n b_{ij} = \alpha \quad \forall i = 1, \dots, n$. Per il Lemma 1.3.9 $\rho(B) = \alpha$ e per

il Corollario 1.3.8 $\rho(B) \leq \rho(A)$, quindi $\min_{i=1, \dots, n} \sum_{j=1}^n a_{ij} \leq \rho(A)$.

Per dimostrare l'altra disuguaglianza si procede in maniera simile: consideriamo $\gamma = \max_{i=1, \dots, n} \sum_{j=1}^n a_{ij}$, e costruiamo una nuova matrice C tale che $0 \leq A \leq C$ e $\sum_{j=1}^n c_{ij} = \gamma \quad \forall i = 1, \dots, n$. Per il Lemma 1.3.9 $\rho(C) = \gamma$ e per il Corollario 1.3.8 $\rho(A) \leq \rho(C)$.

Per dimostrare la seconda catena di disuguaglianze si procede analogamente a quanto fatto sopra, sostituendo A con A^T e lavorando con le somme degli elementi delle righe di A^T \square

Vediamo nel prossimo corollario un'importante conseguenza di questo teorema.

Corollario 1.3.11. *Sia $A \in \mathbb{R}^{n \times n}$. Se $A \geq 0$ e $\sum_{j=1}^n a_{ij} > 0 \quad \forall i = 1, \dots, n$, allora $\rho(A) > 0$. In particolare, se $A > 0 \Rightarrow \rho(A) > 0$; oppure se $A \geq 0$ e irriducibile, allora $\rho(A) > 0$.*

Vogliamo generalizzare il risultato del teorema 1.3.10: date $A, D \in \mathbb{R}^{n \times n}$, D invertibile, le matrici $D^{-1}AD$ e A hanno gli stessi autovalori, di conseguenza $\rho(D^{-1}AD) = \rho(A)$. Se $D = \text{diag}(x_1, x_2, \dots, x_n)$, se $A \geq 0$, e se $x_i > 0 \quad \forall i = 1, \dots, n$, allora $D^{-1}AD \geq 0$. Applicando il teorema 1.3.10 alla matrice $D^{-1}AD = (x_i^{-1}a_{ij}x_j)_{1 \leq i, j \leq n}$, otteniamo un risultato generale, illustrato nel seguente teorema

Teorema 1.3.12. *Sia $A \in \mathbb{R}^{n \times n}$, $A \geq 0$, allora per ogni vettore positivo $x \in \mathbb{R}^n$ abbiamo*

$$\min_{i=1, \dots, n} \frac{1}{x_i} \sum_{j=1}^n a_{ij} x_j \leq \rho(A) \leq \max_{i=1, \dots, n} \frac{1}{x_i} \sum_{j=1}^n a_{ij} x_j; \quad (1.11)$$

$$\min_{j=1, \dots, n} x_j \sum_{i=1}^n \frac{a_{ij}}{x_i} \leq \rho(A) \leq \max_{j=1, \dots, n} x_j \sum_{i=1}^n \frac{a_{ij}}{x_i}. \quad (1.12)$$

Corollario 1.3.13. *Sia $A \in \mathbb{R}^{n \times n}$, sia $x \in \mathbb{R}^n$, supponiamo $A \geq 0$ e $x > 0$. Se esistono $\alpha, \beta \geq 0$ tali che $\alpha x \leq Ax \leq \beta x$, allora $\alpha \leq \rho(A) \leq \beta$. Se $\alpha x < Ax$, allora $\alpha < \rho(A)$, se $Ax < \beta x$, allora $\rho(A) < \beta$.*

Dimostrazione. Siano

$$Ax = \left(\sum_{j=1}^n a_{ij}x_j \right)_{1 \leq i \leq n} \in \mathbb{R}^n \quad \text{e} \quad \alpha x = (\alpha x_i)_{1 \leq i \leq n} \in \mathbb{R}^n,$$

l'espressione $\alpha x \leq Ax$ equivale a $\alpha x_i \leq \sum_{j=1}^n a_{ij}x_j \quad \forall i = 1, \dots, n$. Allora

$$\alpha \leq x_i^{-1} \sum_{j=1}^n a_{ij}x_j \quad \forall i = 1, \dots, n \Rightarrow \alpha \leq \min_{i=1, \dots, n} x_i^{-1} \sum_{j=1}^n a_{ij}x_j.$$

Applicando il Teorema 1.3.12 otteniamo proprio $\alpha \leq \rho(A)$. Se $\alpha x < Ax$, allora esiste sicuramente un $\alpha' > \alpha$ tale che $\alpha' x \leq Ax$, perciò $\rho(A) \geq \alpha' > \alpha$, quindi $\rho(A) > \alpha$. L'altra disuguaglianza si verifica in modo analogo, prendendo

$$\beta \geq \max_{i=1, \dots, n} x_i^{-1} \sum_{j=1}^n a_{ij}x_j. \quad \square$$

Corollario 1.3.14. *Sia $A \in \mathbb{R}^{n \times n}$, sia $A \geq 0$. Se A ha un autovettore positivo, allora il corrispondente autovalore è $\rho(A)$; cioè se $Ax = \lambda x$, $x > 0$ e $A \geq 0$, allora $\lambda = \rho(A)$.*

Dimostrazione. Se $x > 0$ e $Ax = \lambda x$, poichè $Ax \geq 0$, quando $A \geq 0$, allora $\lambda x \geq 0$, quindi $\lambda \geq 0$. Inoltre $\lambda x \leq Ax \leq \lambda x$, dal corollario precedente abbiamo $\lambda \leq \rho(A) \leq \lambda$. \square

Corollario 1.3.15 (Formula di Collatz-Wieland). *Sia $A \in \mathbb{R}^{n \times n}$, $A \geq 0$. Se A ha un autovettore positivo allora il raggio spettrale $\rho(A)$ soddisfa*

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

$$\text{Cioè } \rho(A) = \max_{x > 0} \min_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}.$$

Dimostrazione. Per ogni $x > 0$ vale $\rho(A) \geq \min_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij}x_j$ (Teorema 1.3.12), quindi vale anche per quella x che massimizza $\{x \text{ t.c. } x > 0\}$, cioè

$$\rho(A) \geq \max_{x > 0} \min_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij}x_j. \quad \text{Se scelgo } x \text{ autovettore positivo di } A, \text{ il suo}$$

autovalore associato è proprio $\rho(A)$ (Corollario 1.3.14), allora vale l'uguaglianza

$$\rho(A) = \max_{x > 0} \min_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij}x_j. \quad \square$$

1.4 Teorema di Perron

In questa sezione concentriamo la nostra attenzione solo su matrici positive, nel capitolo successivo estenderemo anche a matrici non negative i risultati qui descritti.

Lemma 1.4.1. *Sia $A \in \mathbb{R}^{n \times n}$, sia $A > 0$, $Ax = \lambda x$, $x \neq 0$, $|\lambda| = \rho(A)$. Allora*

$$A|x| = \rho(A)|x| \text{ e } |x| > 0.$$

Dimostrazione. Sia (λ, x) l'autocoppia di A con $|\lambda| = \rho(A)$, abbiamo

$$\rho(A)|x| = |\lambda||x| = |\lambda x| = |Ax| \leq |A||x|.$$

Definisco $y := A|x| - \rho(A)|x| \geq 0$.

Poichè $|x| \geq 0$ e $|x| \neq 0$, dall'Osservazione 1.3.6, abbiamo $A|x| > 0$.

Se $y=0$, allora $A|x| = \rho(A)|x|$. Sappiamo dal Corollario 1.3.11 che $\rho(A) > 0$,

allora possiamo scrivere $|x| = \frac{1}{\rho(A)}A|x| > 0$.

Se $y \neq 0$, definisco $z := A|x| > 0$, poichè $y \geq 0 \Rightarrow 0 < Ay = Az - \rho(A)z \Rightarrow$

$Az > \rho(A)z$. Dal Corollario 1.3.13 abbiamo $\rho(A) > \rho(A)$, assurdo. Allora $y = 0$

e abbiamo il risultato. \square

Teorema 1.4.2. *Sia $A \in \mathbb{R}^{n \times n}$, supponiamo $A > 0$. Allora $\rho(A) > 0$, $\rho(A)$ è un autovalore di A ed esiste un vettore positivo x tale che $Ax = \rho(A)x$.*

Dimostrazione. Sicuramente esiste un autovalore λ con $|\lambda| = \rho(A)$ e un autovettore associato $x \neq 0$. Per il Lemma 1.4.1, il vettore richiesto è proprio $|x|$. \square

Lemma 1.4.3. *Sia $A \in \mathbb{R}^{n \times n}$, $A > 0$, $Ax = \lambda x$, $x \neq 0$, $x \in \mathbb{C}^n$ e $\lambda = \rho(A)$. Allora esiste almeno un $\theta \in \mathbb{R}$ tale che $e^{-i\theta}x = |x| > 0$.*

Dimostrazione. Possiamo scrivere $|Ax| = |\lambda x| = |\lambda||x| = \rho|x|$.

Dal Lemma 1.4.1, abbiamo $A|x| = \rho(A)|x|$ e $|x| > 0$. Scrivendo queste catene di uguaglianze componente per componente e utilizzando la disuguaglianza triangolare, abbiamo $\forall i = 1, \dots, n$

$$\begin{aligned} \rho(A)|x_i| &= |\lambda||x_i| = |\lambda x_i| = \left| \sum_{j=1}^n a_{ij}x_j \right| \leq \\ &\leq \sum_{j=1}^n |a_{ij}||x_j| = \sum_{j=1}^n a_{ij}|x_j| = \rho(A)|x_i|. \end{aligned}$$

Il primo e l'ultimo termine sono uguali, allora deve valere $\left| \sum_{j=1}^n a_{ij}x_j \right| = \sum_{j=1}^n |a_{ij}||x_j|$.

Ciò è vero se e solo se gli $a_{ij}x_j$ giacciono tutti sulla stessa semiretta del piano

complesso. Ogni $a_{ij}x_j$ si può scrivere come $a_{ij}r_j e^{i\theta_j}$, con $r_j = |x_j|$ e $\theta_j = \arg$ omento di x_j ; se stanno tutti sulla stessa semiretta allora per qualche θ vale $\theta = \theta_j \quad \forall j = 1, \dots, n$.

Allora $e^{-i\theta} a_{ij}x_j = a_{ij}r_j > 0 \quad \forall j = 1, \dots, n$. Questo perchè $a_{ij}r_j > 0$ in quanto gli $a_{ij} > 0$ per ipotesi e $r_j > 0$ perchè sono le componenti del modulo dell'autovettore associato a $\rho(A)$, come dimostrato nel Lemma 1.4.1. Allora

$$e^{-i\theta}x > 0.$$

□

Teorema 1.4.4. *Sia $A \in \mathbb{R}^{n \times n}$, $A > 0$. Allora ogni altro autovalore di A diverso da $\rho(A)$ ha modulo strettamente minore di $\rho(A)$; cioè $|\lambda| < \rho(A)$ per ogni autovalore $\lambda \neq \rho(A)$.*

Possiamo anche dire che $\rho(A)$ è l'unico autovalore che massimizza il modulo.

Dimostrazione. Per definizione di raggio spettrale, $|\lambda| \leq \rho(A)$ per ogni autovalore λ di A . Supponiamo $|\lambda| = \rho(A)$ e $Ax = \lambda x$, $x \neq 0$. Dal lemma 1.4.3, $y = e^{-i\theta}x > 0$, per qualche $\theta \in \mathbb{R}$, allora $Ay = \lambda y$. Dal corollario 1.3.14 $\lambda = \rho(A)$. □

Fin qui abbiamo visto che se $A > 0$, allora $\rho(A)$ è un autovalore di A ed è l'autovalore con modulo maggiore di tutti gli altri. Inoltre sappiamo che esiste almeno un autovettore positivo associato a $\rho(A)$. Alla luce di questi risultati viene spontaneo chiedersi quanti siano gli autovettori associati a $\rho(A)$ (cioè quale è la sua molteplicità geometrica) e se $\rho(A)$ sia un autovalore semplice (cioè con molteplicità algebrica 1) o meno.

Con i prossimi teoremi andremo proprio a dimostrare che c'è un unico autovettore positivo p tale che $\|p\|_1 = 1$ associato a $\rho(A)$ e tutti gli altri autovettori non negativi associati a $\rho(A)$ sono suoi multipli positivi: la molteplicità geometrica di $\rho(A)$ è 1. Vedremo inoltre che la sua molteplicità algebrica è 1.

Teorema 1.4.5. *Sia $A \in \mathbb{R}^{n \times n}$, $A > 0$, siano x, y vettori non nulli tale che $Ax = \rho(A)x$ e $Ay = \rho(A)y$, allora esiste qualche $\alpha \in \mathbb{C}$ tale che $x = \alpha y$.*

Dimostrazione. Dal Lemma 1.4.3 sappiamo che esistono $\theta_1, \theta_2 \in \mathbb{R}$ tale che $p := e^{-i\theta_1}y > 0$, $q := e^{-i\theta_2}x > 0$. Sia $\beta := \min_{i=1, \dots, n} q_i p_i^{-1}$, e definiamo $r := q - \beta p$. Per come abbiamo definito β , r ha almeno una coordinata nulla, quindi $r \geq 0$. $Ar = Aq - \beta Ap = \rho(A)q - \beta \rho(A)p = \rho(A)(q - \beta p) = \rho(A)r$. Abbiamo $A > 0$, $r \geq 0 \Rightarrow Ar > 0$ (Osservazione 1.3.6). Allora $r = \frac{1}{\rho(A)}Ar > 0$; ma $r \geq 0$ per costruzione, perciò può essere solamente $r = 0$, quindi $q = \beta p$, da cui $x = \beta e^{i(\theta_2 - \theta_1)}y$. Per $\alpha = \beta \exp^{i(\theta_2 - \theta_1)}$ abbiamo il risultato. □

In altre parole la molteplicità geometrica della radice di Perron $\rho(A)$ è 1, cioè la dimensione dell'autospazio invariante associato a $\rho(A)$ è 1: è una retta.

Corollario 1.4.6 (Esistenza e unicità del vettore di Perron). *Sia $A \in \mathbb{R}^n$, $A > 0$. Allora esiste un unico vettore $x \in \mathbb{R}^n$, tale che $Ax = \rho(A)x$, $x > 0$ e*

$$\|x\|_1 = \sum_{j=1}^n x_j = 1.$$

Dimostrazione. Nel Teorema 1.4.2 abbiamo visto che un autovettore x positivo associato a $\rho(A)$ esiste. Normalizzando x , $\frac{x}{\|x\|_1} =: p$ otteniamo proprio il vettore richiesto p . Concentriamoci ora sull'unicità. Dal Teorema 1.4.5 se esiste un altro autovettore y associato a $\rho(A)$, questo sarà del tipo $y = \alpha x$ per qualche $\alpha \in \mathbb{C}$, e di conseguenza $\|y\|_1 = \sum_{j=1}^n |y_j| = \sum_{j=1}^n |\alpha| |x_j| \neq \sum_{j=1}^n |x_j| = 1$. \square

Questo unico vettore x tale che $\|x\|_1 = 1$ è chiamato **vettore di Perron** di A e da adesso in poi lo indicheremo con p ; $\rho(A)$ è chiamato **radice di Perron** di A . Se $A > 0$, anche A^T è una matrice positiva, e, poichè A e A^T hanno gli stessi autovalori, $\rho(A) = \rho(A^T)$, tutti i risultati che abbiamo dimostrato per A valgono anche per A^T .

Se $(\rho(A), p)$ è l'**autocoppia di Perron** di A , esisterà $(\rho(A), q)$ autocoppia di Perron di A^T , chiamiamo q **vettore sinistro di Perron**.

$$A^T q = \rho(A)q \iff (A^T q)^T = (\rho(A)q)^T \iff q^T A = q^T \rho(A), q^T > 0.$$

Corollario 1.4.7. *Sia $A \in \mathbb{R}^{n \times n}$, $A > 0$, allora non ci sono altri autovettori non negativi di A tranne il vettore di Perron p e i suoi multipli positivi. Cioè se esiste qualche vettore $x \in \mathbb{R}^n$, $x \geq 0$, $x \neq 0$ e $Ax = \lambda x$, allora $x = \alpha p$, per qualche $\alpha \in \mathbb{R}$ e $\lambda = \rho(A)$.*

Dimostrazione. Sia (λ, x) un' autocoppia di A , con x come nelle ipotesi, sia q il vettore sinistro di Perron (cioè il vettore di Perron di A^T), $q > 0$, allora $q^T x > 0$. $\rho(A)q^T = q^T A \Rightarrow \rho(A)q^T x = q^T Ax = q^T (\lambda x) = \lambda q^T x$. Ciò significa che $\rho(A) = \lambda$ e dal Teorema 1.4.5 esiste un qualche $\alpha \in \mathbb{C}$ tale che $x = \alpha p$; ma per ipotesi $x \in \mathbb{R}^n$, quindi α non può appartenere a \mathbb{C} , ma $\alpha \in \mathbb{R}$; inoltre $x \geq 0$ e $x \neq 0$, quindi $\alpha \in \mathbb{R}_{>0}$ \square

Enunciamo un ultimo lemma e un'ultima serie di teoremi per dimostrare che la molteplicità algebrica di $\rho(A)$ è 1.

Lemma 1.4.8. Sia $A \in \mathbb{R}^{n \times n}$ una matrice data, sia $\lambda \in \mathbb{C}$, siano $x, y \in \mathbb{C}^n$ tali che

1. $Ax = \lambda x$;
2. $A^T y = \lambda y$;
3. $x^T y = 1$.

Definiamo $L = xy^T$, allora:

- (a) $Lx = x$ e $Ly = y$;
- (b) $L^k = L \quad \forall k = 1, \dots, n$;
- (c) $A^k L = LA^k = \lambda^k L \quad \forall k = 1, \dots, n$;
- (d) $L(A - \lambda L) = 0$;
- (e) $(A - \lambda L)^k = A^k - \lambda^k L \quad \forall k = 1, \dots, n$;
- (f) ogni autovalore non-nullo di $A - \lambda L$ è anche un autovalore di A .

Se assumiamo anche

4. $\lambda \neq 0$;
5. λ è autovalore di A con molteplicità geometrica 1;

Abbiamo

- (g) λ non è un autovalore di $A - \lambda L$.

Infine, assumiamo che

6. $|\lambda| = \rho(A) > 0$;
7. λ è l'unico autovalore di A con modulo $\rho(A)$; e ordinando dal più piccolo al più grande i moduli degli autovalori, $|\lambda_1| \leq |\lambda_2| \leq \dots \leq |\lambda_{n-1}| < |\lambda_n| = |\lambda| = \rho(A)$, allora

- (h) $\rho(A - \lambda L) \leq |\lambda_{n-1}| < \rho(A)$;
- (i) $\left(\frac{A}{\lambda}\right)^k = L + \left(\frac{A}{\lambda} - L\right)^k \rightarrow 0$ per $k \rightarrow \infty$.

Dimostrazione. Iniziamo con la (a)

$$Lx = xy^T x = \underbrace{(x^T y x^T)}_{=1}^T = (x^T)^T = x;$$

$$y^T L = y^T xy^T = (y \underbrace{x^T y}_{=1})^T = y^T.$$

La (b) è quasi ovvia:

$$L^k = (xy^T)^k = x \underbrace{y^T x}_{=1} \underbrace{y^T x}_{=1} \cdots \underbrace{y^T x}_{=1} y^T = xy^T = L.$$

per $k-1$ volte

Dimostriamo ora la (c)

$$\begin{aligned} A^k L &= A^k xy^T = \lambda^k xy^T = \lambda^k L, \\ LA^k &= xy^T A^k = ((A^T)^k yx^T)^T = (\lambda^k yx^T) = \lambda^k xy^T. \end{aligned}$$

Per dimostrare la (d) utilizzeremo le proprietà (b) e (c) appena viste.

$$L(A - \lambda L) = LA - \lambda L^2 = AL - \lambda L = \lambda L - \lambda L = 0,$$

Dimostriamo la (e) per induzione, avvalendoci ancora una volta di (b) e (c).

passo base:

Per $k = 1$ vale $A - \lambda L = A - \lambda$? Sì, possiamo procedere.

passo induttivo:

Considero $k \geq 2$, supponiamo che la nostra tesi valga $\forall k = 1, \dots, n$, dimostriamo anche per $k = n + 1$.

$$\begin{aligned} (A - \lambda L)^{n+1} &= (A - \lambda L)^n (A - \lambda L) \stackrel{p.i.}{=} (A^n - \lambda^n L)(A - \lambda L) = \\ &= A^{n+1} - \lambda A^n L - \lambda^n A L + \lambda^{n+1} L^2 = A^{n+1} - \lambda \lambda^n L - \lambda^n \lambda L + \lambda^{n+1} L. \end{aligned}$$

(f) Sia μ un autovalore di $A - \lambda L$, cioè per qualche $z \in \mathbb{C}^n$, $z \neq 0$, si ha $(A - \lambda L)z = \mu z$. Moltiplicando a destra e sinistra per L otteniamo

$$\begin{aligned} L(A - \lambda L)z &= 0z = \mu Lz. \\ \Rightarrow Lz = 0 &\Rightarrow (A - \lambda L)z = Az - \lambda Lz = Az - 0 = \mu z, \end{aligned}$$

quindi μ è anche autovalore di A .

Utilizzo per dimostrare la (g) gli stessi μ e z del punto (f).

Per assurdo, sia $\mu = \lambda$, sia z l'autovettore di $A - \lambda L$ associato a λ ; per la (f) λ deve essere autovalore di A . Dalle condizioni (1) e (5) abbiamo $z = \alpha x$, per un qualche $\alpha \neq 0$. Abbiamo quindi

$$\mu z = \lambda z = (A - \lambda L)z = (A - \lambda L)\alpha x = \alpha \lambda x - \alpha \lambda x = 0.$$

È un assurdo, poichè $\lambda, x, \alpha \neq 0$, quindi λ non è autovalore di $A - \lambda L$.

(h) Ogni autovalore di $A - \lambda L$ è anche autovalore di A (proprietà (f)), quindi esisterà un certo λ_j , autovalore di A , per cui $\rho(A - \lambda L) = |\lambda_j|$, oppure

$\rho(A - \lambda L) = 0$. Abbiamo considerato gli autovalori di A crescenti in modulo, $|\lambda_n| = |\lambda| = \rho(A)$ e dalla (g) sappiamo che λ non è autovalore di $A - \lambda L$, quindi $\rho(A - \lambda L) \leq |\lambda_{n-1}| \leq \rho(A)$.

Concludiamo con la (i), che è immediata considerando la (h) e la (e):
 $(A - \lambda L)^k = A^k - \lambda^k L$, dividendo entrambi i membri per λ^k otteniamo

$$\left(\frac{A}{\lambda} - L\right)^k = \left(\frac{A}{\lambda}\right)^k - L \rightarrow 0 \quad \text{per } k \rightarrow \infty.$$

□

Se applichiamo questo lemma a x vettore di Perron di A , $\lambda = \rho(A)$, radice di Perron di A e $y = (x^T z)^{-1} z$, con z vettore sinistro di Perron (cioè vettore di Perron di A^T), abbiamo il seguente teorema.

Teorema 1.4.9. *Sia $A \in \mathbb{R}^{n \times n}$, $A > 0$. Allora*

$$\lim_{k \rightarrow \infty} \left(\frac{A}{\rho(A)}\right)^k = L.$$

Con $L = xy^T$, $Ax = \rho(A)x$, $A^T y = \rho(A)y$, $x > 0$, $y > 0$, $x^T y = 1$.

Dimostrazione. La tesi è diretta conseguenza del punto (i) del Lemma.

Osserviamo che y è necessariamente multiplo del vettore di Perron z (Corollario 1.4.7). □

Corollario 1.4.10. *Sia $A \in \mathbb{R}^{n \times n}$, $A > 0$, allora $L = \lim_{k \rightarrow \infty} \left(\frac{A}{\rho(A)}\right)^k$ è una matrice positiva di rango 1.*

Dimostrazione. Il rango di una matrice, per definizione, è la dimensione dell'immagine dell'applicazione associata alla matrice. Osserviamo che $Lv = xy^T v = (v^T \underbrace{yx^T}_{=1})^T = v \quad \forall v \in \mathbb{R}^n$. Quindi la dimensione dell'immagine dell'applicazione associata alla matrice L è 1, quindi $rg(L) = 1$.

Partendo dall'ipotesi $x^T y = 1$ posso dire $y^T x = 1$ per la simmetria del prodotto scalare. Infine L è positiva perchè è stata definita con prodotto di due vettori positivi. □

Teorema 1.4.11. *Sia $A \in \mathbb{R}^{n \times n}$, $A > 0$, allora $\rho(A)$ è un autvalore di molteplicità algebrica 1; cioè $\rho(A)$ è una radice semplice del polinomio caratteristico di (A) .*

Per $k \rightarrow \infty$ i primi m elementi diagonali restano 1, gli altri $n - m$ diventano tutti 0.

$$L = P \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & * \\ & & & 0 & \\ 0 & & & & \ddots \\ & & & & & 0 \end{pmatrix} P^{-1}$$

Quindi L ha rango m , ma dal Teorema 1.4.9 sappiamo che L ha rango 1, allora $m = 1$ e la molteplicità algebrica di $\rho(A)$ è 1. \square

Siamo quindi Pronti per enunciare il teorema di Perron. Questo teorema si riferisce solo a matrici positive, vedremo più in là il caso generale per matrici non negative.

Teorema 1.4.12 (Teorema di Perron). *Sia $A \in \mathbb{R}^{n \times n}$, $A > 0$, allora:*

- (a) $\rho(A) > 0$;
- (b) $\rho(A)$ è un autovalore di A ;
- (c) la molteplicità algebrica di $\rho(A)$ è 1 (cioè $\rho(A)$ è un autovalore semplice);
- (d) esiste un vettore $x > 0$ tale che $Ax = \rho(A)x$;
- (e) l'unico vettore $p \in \mathbb{R}^n$, $p > 0$ tale che $Ap = \rho(A)p$ e $\|p\|_1 = 1$ è chiamato vettore di Perron di A .

Non ci sono altri autovettori non negativi di A ad eccezione dei multipli positivi di p , cioè se $x \geq 0$, $x \neq 0$ e $Ax = \lambda x \Rightarrow \lambda = \rho(A)$ e $x = \alpha p$, con $\alpha \in \mathbb{R}$;

- (f) vale la formula di Collatz-Wieland:

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

- (g) $\rho(A)$ è l'autovalore di modulo massimo di A , cioè per ogni altro autovalore $\lambda \neq \rho(A)$ vale $|\lambda| < \rho(A)$;

- (h) Dati $x, y > 0$ tali che $Ax = \rho(A)x$, $A^T y = \rho(A)y$, $x^T y = 1$, definita $L := xy^T \in \mathbb{R}^{n \times n}$, allora

$$\lim_{k \rightarrow \infty} \left(\frac{A}{\rho(A)} \right)^k = L.$$

Capitolo 2

Teorema di Perron-Frobenius

2.1 Estensione del teorema di Perron a matrici non-negative

In problemi pratici capita molto spesso di incontrare matrici non negative, anzichè matrici positive; per questo è necessario estendere la teoria trattata nella sezione precedente al caso in cui la matrice abbia almeno un'entrata nulla.

Purtoppo ci sono quantità come il rango e la dimensione che non sono funzioni continue e non possono essere generalizzate tramite passaggio al limite; invece altri risultati del teorema di Perron sì: illustriamoli nel seguente teorema.

Teorema 2.1.1. *Sia $A \in \mathbb{R}^{n \times n}$, $A \geq 0$. Allora:*

- $\rho(A)$ è un autovalore di A ;
- esiste un vettore non negativo $x \geq 0$, $x \neq 0$ tale che $Ax = \rho(A)x$.

Dimostrazione. Per ogni $\epsilon > 0$ definiamo $A(\epsilon) = \left(a_{ij} + \frac{1}{\epsilon} \right)_{1 \leq i, j \leq n}$; allora

$A(\epsilon) > 0$. Sia $x(\epsilon)$ il vettore di Perron di $A(\epsilon)$, allora $x(\epsilon) > 0$ e $\sum_{i=1}^n x(\epsilon)_i = 1$.

Poichè l'insieme dei vettori $\{x(\epsilon), \epsilon > 0\}$ è contenuto nel compatto

$\{x \in \mathbb{R}^n, \|x\|_1 \leq 1\}$, posso estrarre una sottosuccessione convergente $\frac{1}{\epsilon_1}, \frac{1}{\epsilon_2}, \dots, \frac{1}{\epsilon_n}$

tale che $\lim_{k \rightarrow \infty} \frac{1}{\epsilon_k} = 0$. Allora esiste $\lim_{k \rightarrow \infty} x(\epsilon_k) = x$. Osserviamo che

$$x(\epsilon_k) > 0 \quad \forall k = 1, \dots, n, \quad \text{allora } x = \lim_{k \rightarrow \infty} x(\epsilon_k) \geq 0.$$

Non può essere $x = 0$ perchè $\sum_{i=1}^n x_i = \lim_{k \rightarrow \infty} \sum_{i=1}^n x(\epsilon_k)_i = 1$.

Dal Corollario 1.3.8 $\rho(A(\epsilon_k)) \geq \rho(A(\epsilon_{k+1})) \geq \dots \geq \rho(A) \quad \forall k = 1, \dots, n$, ho una

successione monotona decrescente di numeri reali $(\rho(A(\epsilon_k)))_{1 \leq k \leq n}$.

Allora $\rho := \lim_{k \rightarrow \infty} \rho(A + \frac{1}{\epsilon_k})$ esiste e $\rho \geq \rho(A)$; ma poichè

$$Ax = \lim_{k \rightarrow \infty} A(\epsilon_k)x(\epsilon_k) = \lim_{k \rightarrow \infty} \rho(A(\epsilon_k))x(\epsilon_k) = \quad (2.1)$$

$$= \lim_{k \rightarrow \infty} \rho(A(\epsilon_k)) \lim_{k \rightarrow \infty} x(\epsilon_k) = \rho x. \quad (2.2)$$

Siccome $x \neq 0$, ρ è autovalore di A , allora $\rho \geq \rho(A) \Rightarrow \rho = \rho(A)$. \square

Anche la formula di *Collatz-Wieland* resta valida, per dimostrarla abbiamo bisogno di una estensione a vettori non negativi del Corollario 1.3.13, che ci viene fornita dal prossimo teorema.

Teorema 2.1.2. *Sia $A \in \mathbb{R}^{n \times n}$, $A \geq 0$, $x \in \mathbb{R}^n$, $x \geq 0$, $x \neq 0$.*

Se $Ax \geq \alpha x$ per qualche $\alpha \in \mathbb{R}$, allora $\rho(A) \geq \alpha$.

Dimostrazione. Sia $A = (a_{ij})_{1 \leq i, j \leq n}$, sia $\epsilon > 0$, andiamo a definire $A(\epsilon) := (a_{ij} + \epsilon)_{1 \leq i, j \leq n}$, allora $A(\epsilon) > 0$ e vale il Teorema di Perron.

Per ipotesi $Ax - \alpha x \geq 0 \Rightarrow A(\epsilon)x - \alpha x > Ax - \alpha x \geq 0$. Considero $q(\epsilon) > 0$ il vettore sinistro di Perron di $A(\epsilon)$, allora

$$q(\epsilon)^T (A(\epsilon)x - \alpha x) = \rho(A(\epsilon))q(\epsilon)^T x - \alpha q(\epsilon)^T x = (\rho(A(\epsilon)) - \alpha)q(\epsilon)^T x \geq 0.$$

Poichè $q(\epsilon)^T x > 0$ deve essere $(\rho(A(\epsilon)) - \alpha) \geq 0 \quad \forall \epsilon > 0$. Osserviamo che $\lim_{\epsilon \rightarrow 0} \rho(A(\epsilon)) = \rho(A)$, allora concludiamo $\rho(A) \geq \alpha$. \square

Corollario 2.1.3 (Formula di Collatz-Wieland). *Sia $A \in \mathbb{R}^{n \times n}$, $A \geq 0$, allora*

$$\rho(A) = \max_{x \geq 0} \min_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij} x_j \quad \text{con } x_i \neq 0.$$

Dimostrazione. Se $A \geq 0$, $x \geq 0$, $x \neq 0$, scegliamo $\alpha = \min_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij} x_j$,

allora $Ax \geq \alpha x$, da cui $\rho(A) \geq \alpha$. Questa disuguaglianza vale per ogni $x \geq 0$,

quindi anche per quella che massimizza $\{x \text{ t.c. } x \geq 0\}$; cioè $\rho(A) \geq \max_{x \geq 0} \min_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij} x_j$.

Se scelgo la $x \geq 0$ del Teorema 2.1.1, che è autovettore di A associato a $\rho(A)$, vale l'uguaglianza. \square

2.2 Irriducibilità, grafi e teorema di Perron-Frobenius

Definizione 2.2.1 (Grafo). Un **grafo** $G = (N, E)$ è composto da un insieme finito di nodi (o vertici) $N = \{N_1, \dots, N_m\}$ e da un insieme di archi

$E = \{E_1, \dots, E_n\} \subset N \times N$, con $E_i = (N_i, N_j)$ che connettono coppie di nodi.

Due nodi connessi da un arco sono detti **adiacenti**.

Un **cammino** in un grafo G è una successione di nodi tale che ogni nodo N_i è adiacente al nodo N_{i+1} . Un cammino in cui il primo e l'ultimo nodo coincidono è detto **ciclo**.

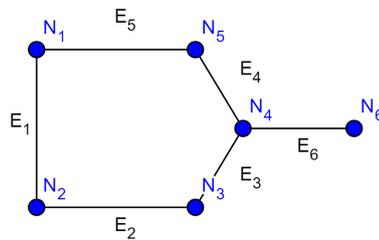


Figura 2.1: Esempio di grafo

Definizione 2.2.2 (Connettività e orientazione). Un grafo G è connesso se per ogni coppia di nodi (N_i, N_j) esiste un cammino che li collega.

Un grafo è **orientato** se a ogni arco è assegnato un verso: significa che l'insieme E è formato da coppie ordinate di nodi. L'arco (N_i, N_j) è un arco da N_i a N_j , si chiama N_j la testa e N_i la coda dell'arco.

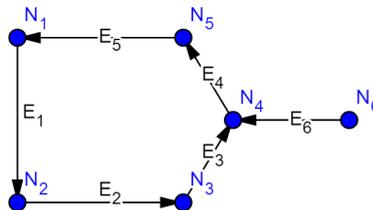


Figura 2.2: Esempio di grafo orientato.

L'essere connesso di un grafo orientato è indipendente dall'orientazione assegnata agli archi.

Definizione 2.2.3. Un grafo G è **fortemente connesso** se per ogni coppia di nodi (N_i, N_j) c'è una sequenza di archi orientata da N_i a N_j .

Definizione 2.2.4 (Grafo associato ad A). Sia $A \in \mathbb{R}^{n \times n}$, $A = (a_{ij})_{1 \leq i, j \leq n}$ indichiamo con $G(A)$ il grafo associato alla matrice A . $G(A)$ è definito come il grafo orientato avente n nodi $\{N_1, \dots, N_n\}$ e avente un arco orientato da N_i a N_j se e solo se l'elemento $a_{ij} \neq 0$.

Esempio 2.2.5. Data la matrice $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ il suo grafo associato è

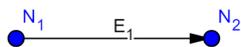


Figura 2.3: Esempio di grafo associato alla matrice A

Proposizione 2.2.6. Sia $A \in \mathbb{R}^{n \times n}$ e supponiamo $A \geq 0$. Allora A è irriducibile se e solo se $(I + A)^{n-1} > 0$.

Dimostrazione. Dimostriamo che A è riducibile $\iff (I + A)^{n-1}$ ha almeno un'entrata nulla.

Sia A riducibile, allora

$$A = P\hat{A}P^T = P \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix} P^T, \text{ con } P \text{ matrice di permutazione.}$$

Se $A \geq 0$, allora anche $\hat{A} \geq 0$, poichè P e P^T hanno solo l'effetto di scambiare righe e colonne.

$$\hat{A}^2 = \begin{pmatrix} X^2 & XY + YZ \\ 0 & Z^2 \end{pmatrix}, \dots, \hat{A}^{(n-1)} = \begin{pmatrix} X^{(n-1)} & R \\ 0 & Z^{(n-1)} \end{pmatrix},$$

con $R \in \mathbb{R}^{r \times n-r}$, $R \geq 0$.

Osserviamo che ogni potenza di \hat{A} ha un blocco di 0 di dimensione $n - r \times r$ in basso a sinistra. Allora

$$\begin{aligned} (I + A)^{n-1} &= (I + P\hat{A}P^T)^{n-1} = (P(I + \hat{A})P^T)^{n-1} = \\ &= P(I + \hat{A})^{n-1}P^T = P \left(\sum_{k=1}^n \binom{n}{k} A^k I^{n-1-k} \right) P^T = \\ &= P((n-1)A + \binom{n-1}{2} A^2 + \dots + \binom{n-1}{n-1} A^{n-1} + I)P^T. \end{aligned}$$

Tutte le matrici della somma hanno un blocco $n - r \times r$ di 0 in basso a sinistra, quindi $(I+A)^{n-1}$ è riducibile, quindi non può avere tutte entrate diverse da zero.

Supponiamo adesso che l'entrata (p, q) -esima di $(I + A)^{n-1}$ sia nulla, con $p \neq q$. Allora nel grafo $G(A)$ associato alla matrice A non c'è l'arco orientato dal punto N_p al punto N_q . Definiamo l'insieme dei nodi

$$V_1 := \{N_i | N_i = N_q, \text{ oppure esiste un cammino in } G(A) \text{ da } N_i \text{ a } N_q\}.$$

Sia V_2 l'insieme di tutti i nodi che non stanno in V_1 . Allora $V_1 \cup V_2 = \{N_1, \dots, N_n\}$, $N_q \in V_1 \neq \emptyset \Rightarrow V_2 \neq \{N_1, \dots, N_n\}$. Se c'è qualche cammino dal nodo N_i di V_2 al nodo N_j di V_1 , allora c'è un cammino $N_i \rightarrow N_j \rightarrow N_q$, allora $N_i \in V_1$, quindi non ci sono cammini dai nodi di V_2 a quelli di V_1 . Rietichettando i nodi abbiamo $V_1 = \{\hat{N}_1, \dots, \hat{N}_r\}$ e $V_2 = \{\hat{N}_{r+1}, \dots, \hat{N}_n\}$. Perciò

$$\hat{A} = P^T A P = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix},$$

con $0 \in \mathbb{R}^{n-r \times r}$. Quindi A è riducibile. \square

In questa dimostrazione abbiamo implicitamente fatto vedere che la matrice A è irriducibile se e solo se il suo grafo associato $G(A)$ è fortemente connesso.

Lemma 2.2.7. *Sia $A \in \mathbb{R}^{n \times n}$ e siano $\lambda_1, \dots, \lambda_n$ i suoi autovalori contati con molteplicità. Allora $1 + \lambda_1, \dots, 1 + \lambda_n$ sono autovalori della matrice $I + A$ e $\rho(I + A) \leq 1 + \rho(A)$; se $A \geq 0$, allora $\rho(I + A) = 1 + \rho(A)$.*

Dimostrazione. Sia λ un autovalore di A con molteplicità k , allora λ è una radice di molteplicità k del polinomio caratteristico di A : $\det(tI - A) = 0$. Osserviamo che $\det(tI - A) = \det((t+1)I - (A+I))$ quindi $\lambda+1$ è una radice di molteplicità k di $\det(rI - (A+I)) = 0$, quindi $1 + \lambda_1, \dots, 1 + \lambda_n$ sono autovalori di $I + A$. Vediamo ora chi è il raggio spettrale $\rho(I + A) = \max_{1 \leq i \leq n} |1 + \lambda_i| \leq \max_{1 \leq i \leq n} 1 + |\lambda_i| = 1 + \rho(A)$. Se $A \geq 0$, sappiamo che $\rho(A)$ è autovalore di A dal Teorema 2.1.1, allora $1 + \rho(A)$ è autovalore di $I + A$, quindi $\rho(I + A) = 1 + \rho(A)$. \square

Lemma 2.2.8. *Sia $A \in \mathbb{R}^{n \times n}$, $A \geq 0$ e $A^k > 0$ per qualche $k \geq 1$. Allora $\rho(A)$ è autovalore di molteplicità algebrica 1 di A .*

Dimostrazione. Se $\lambda_1, \dots, \lambda_n$ sono autovalori di A , allora $\lambda_1^k, \dots, \lambda_n^k$ sono autovalori di A^k . Dal Teorema 2.1.1 $\rho(A)$ è autovalore di A . Se $\rho(A)$ avesse molteplicità algebrica $m \geq 1$, allora anche $\rho(A)^k = \rho(A^k)$ avrebbe molteplicità algebrica $m > 1$, ma dal Teorema di Perron sappiamo che non è possibile. \square

Teorema 2.2.9 (Perron-Frobenius). Sia $A \in \mathbb{R}^{n \times n}$, A irriducibile e $A \geq 0$.

Allora:

(a) $\rho(A) > 0$;

(b) $\rho(A)$ è un autovalore di A ;

(c) la molteplicità algebrica di $\rho(A)$ è 1;

(d) esiste un vettore $x > 0$ tale che $Ax = \rho(A)x$;

(e) l'unico vettore $p \in \mathbb{R}^n$, $p > 0$ tale che $Ap = \rho(A)p$ e $\|p\|_1 = 1$ è chiamato vettore di Perron di A .

Non ci sono altri autovettori non negativi di A ad eccezione dei multipli positivi di p , cioè se $x \geq 0$, $x \neq 0$ e $Ax = \lambda x \Rightarrow \lambda = \rho(A)$ e $x = \alpha p$, con $\alpha \in \mathbb{R}$;

(f) vale la formula di Collatz- Wieland:

$$\rho(A) = \max_{x \in \mathcal{N}} \min_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij} x_j, \text{ con } \mathcal{N} = \{x | x \geq 0, \text{ con } x \neq 0\};$$

Dimostrazione. La (a) deriva direttamente dal Corollario 1.3.11; la (b) è stata dimostrata nel Teorema 2.1.1.

Passiamo alla (c): per assurdo sia $\rho(A)$ autovalore multiplo di A , $1 + \rho(A)$ è autovalore di $I + A$, ed è uguale al raggio spettrale $\rho(I + A)$, allora $1 + \rho(A)$ è autovalore multiplo di $(I + A)^{n-1}$. Ma dal Lemma 2.2.8 abbiamo: $I + A \geq 0$, $(I + A)^{n-1} > 0$, allora $\rho(I + A)$ ha molteplicità algebrica 1.

Per la (d) facciamo riferimento sempre al Teorema 2.1.1, che ci garantisce l'esistenza di un vettore non negativo $x \neq 0$ tale che $Ax = \rho(A)x$. Allora $(I + A)^{n-1}x = (1 + \rho(A))^{n-1}x$. Poichè $(I + A)^{n-1} > 0$ (Proposizione 2.2.6) e $x \geq 0$, vale $(I + A)^{n-1}x > 0$; inoltre $(1 + \rho(A))^{n-1} > 0$.

Allora $x = \underbrace{(1 + \rho(A))^{1-n}}_{>0} \underbrace{(I + A)^{n-1}x}_{>0} \Rightarrow x > 0$.

La dimostrazione della (e) è uguale a quella fatta per il teorema di Perron (in particolare Corollario 1.4.6 e Corollario 1.4.7); infine la (f) è stata dimostrata nel Corollario 2.1.3. \square

Se A è una matrice irriducibile è facile dimostrare che anche A^T lo è, allora il teorema di Perron-Frobenius vale anche per A^T . Chiamiamo *vettore sinistro di Perron* di A il vettore di Perron di A^T , che da adesso in poi indicheremo con q . Il raggio spettrale di A , $\rho(A)$, è chiamato *radice di Perron*: sono esattamente le stesse definizioni date nella conclusione del primo capitolo.

Andiamo ora a dimostrare un'estensione del Teorema 1.3.7 a matrici non negative e irriducibili. Enunciamo prima un breve lemma.

Lemma 2.2.10. *Sia $A \in \mathbb{R}^{n \times n}$, $A \geq 0$ e irriducibile, supponiamo esista un vettore $x \geq 0$, $x \neq 0$ tale che $Ax \geq \rho(A)x$. Allora $Ax = \rho(A)x$ e $x > 0$.*

Dimostrazione. Osserviamo, innanzitutto, che vale Perron-Frobenius. Supponiamo, per assurdo, $Ax > \rho(A)x$ e sia $q > 0$ il vettore sinistro di Perron di A ; abbiamo $(A - \rho(A)I)x > 0 \Rightarrow q^T(A - \rho(A)I)x > 0$, ma questo è impossibile, dato che $q^T(A - \rho(A)I)x = 0$. Allora $Ax = \rho(A)x$, cioè x deve essere un multiplo positivo del vettore di Perron p , quindi $x > 0$. \square

Teorema 2.2.11 (Wieland). *Siano $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{C}^{n \times n}$, con $A \geq 0$ e irriducibile, tali che $|B| \leq A$. Allora $\rho(B) \leq \rho(A)$.*

Se $\rho(A) = \rho(B)$, cioè se esiste qualche $\phi \in \mathbb{R}$ tale che $\lambda = e^{i\phi}\rho(A)$ sia autovalore di B , allora esistono $\theta_1, \dots, \theta_n \in \mathbb{R}$ tali che $B = e^{i\phi}DAD^{-1}$, con $D = \text{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})$.

Dimostrazione. Prima di iniziare la dimostrazione vera e propria richiamiamo alla attenzione questo fatto:

$$\text{se } A \in \mathbb{R}^{n \times n}, \quad A \geq 0 \text{ e } x \in \mathbb{R}^n, x > 0 \text{ e vale } Ax = 0 \Rightarrow A = 0. \quad (2.3)$$

Procediamo adesso con la dimostrazione.

Sappiamo che $\rho(B) \leq \rho(A)$ dal Teorema 1.3.7.

Se $\rho(B) = \rho(A)$ e se x è l'autovettore di B associato a λ tale che $|\lambda| = \rho(A)$, allora $\rho(A)|x| = |\lambda||x| = |\lambda x| = |Bx| \leq |B||x| \leq A|x|$. Dal Lemma 2.2.10 appena dimostrato deve valere $A|x| = \rho(A)|x|$, e $|x| > 0$, quindi $|B||x| = \rho(A)|x|$. Di conseguenza $A|x| = |B||x| \Rightarrow (A - |B|)|x| = 0$; ma $A - |B| \geq 0$ e $|x| > 0$, dall'equazione (2.3) abbiamo $A - |B| = 0 \Rightarrow A = |B|$.

Per ogni $\theta_k \in \mathbb{R}$ definiamo $e^{i\theta_k} := \frac{x_k}{|x_k|}$, $\forall k = 1, \dots, n$; sia $D = \text{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})$, allora $x = D|x|$.

$$BD|x| = Bx = \lambda x = e^{i\phi}\rho(A)x = e^{i\phi}\rho(A)D|x| \quad (2.4)$$

$$\Rightarrow e^{-i\phi}D^{-1}BD|x| = \rho(A)|x| = A|x|. \quad (2.5)$$

Per comodità chiamiamo $C = e^{-i\phi}D^{-1}BD$, $C \in \mathbb{C}^{n \times n}$. È immediato notare che $|C| = |B| = A$, allora possiamo riscrivere la (2.5) come $(|C| - C)|x| = 0$. Consideriamo solo la parte reale di questa equazione:

$$(|C| - \text{Re}(C))|x| = 0, \quad |x| > 0, \text{ dalla (2.3) abbiamo } |C| = \text{Re}(C) \Rightarrow \text{Im}(C) = 0.$$

Allora $C = \text{Re}(C) = |C| = A$, solo adesso posso dire $e^{-i\phi}D^{-1}BD = A$. \square

Corollario 2.2.12 (Radici h -esime di $\rho(A)$). Sia $A \in \mathbb{R}^{n \times n}$, $A \geq 0$ e irriducibile, sia $S = \{\lambda_0, \dots, \lambda_{h-1}\}$ l'insieme degli autovalori di modulo massimo $\rho(A)$, allora sono vere:

- Ogni autovalore $\lambda_k \in S$ ha molteplicità algebrica 1 $\forall k = 0, \dots, h$;
- S è l'insieme delle radici h -esime di $\rho(A)$, cioè

$$S = \{\rho(A), \rho(A)\omega, \rho(A)\omega^2, \dots, \rho(A)\omega^{h-1}\} \quad \text{con} \quad \omega = e^{\frac{2\pi i}{h}}.$$

Inoltre sia λ un autovalore qualsiasi di A , allora $e^{\frac{2\pi i k}{h}} \lambda$ è autovalore di A per ogni $k = 0, \dots, h-1$.

Dimostrazione. Ogni elemento di S , che ricordiamo essere un autovalore di modulo massimo, si può scrivere come $\lambda_k = e^{i\phi_k} \rho(A)$, con $k = 0, \dots, h-1$; cioè, ordinando i ϕ_k tali che $0 = \phi_0 < \phi_1 < \dots < \phi_{h-1}$, $S = \{\rho(A), e^{i\phi_1} \rho(A), \dots, e^{i\phi_{h-1}} \rho(A)\}$. Applichiamo il Teorema 2.2.11 con $B = A$ e $\lambda = e^{i\phi_j} \rho(A)$, questo ci assicura l'esistenza di una matrice D_k tale che

$$A = e^{i\phi_k} D_k A D_k^{-1}. \quad (2.6)$$

Le matrici $D_k A D_k^{-1}$ e A sono simili, quindi avranno gli stessi autovalori. Il teorema di Perron-Frobenius ci garantisce che $\rho(A)$ è un autovalore semplice di A , allora esso sarà anche autovalore semplice di $D_k A D_k^{-1}$. L'identità (2.6) ci assicura che l'insieme degli autovalori di A è portato in sé stesso se ruotato nel piano complesso di un angolo $\phi_k \quad \forall k = 0, \dots, h-1$. Sia $e^{i\phi_k} \rho(A)$ la rotazione di $\rho(A)$ di angolo ϕ_k , allora questa resta un autovalore semplice di A . Abbiamo visto che tutti gli elementi di S sono della forma $e^{i\phi_k} \rho(A)$, con $k = 0, \dots, h-1$, quindi ogni elemento di S è autovalore semplice di A .

Consideriamo adesso un altro autovalore $e^{i\phi_l} \rho(A) \in S$ e applichiamo ancora una volta il Teorema 2.2.11, otteniamo $A = e^{i\phi_l} D_l A D_l^{-1}$, per qualche D_l . Combinando questa scrittura con la (2.6) otteniamo

$$A = e^{i\phi_k} D_k A D_k^{-1} = A = e^{i\phi_k} D_k (e^{i\phi_l} D_l A D_l^{-1}) D_k^{-1} = e^{i(\phi_k + \phi_l)} (D_k D_l) A (D_k D_l)^{-1}.$$

Allora anche $e^{i(\phi_k + \phi_l)} \rho(A)$ è autovalore di modulo massimo, ciò implica che $S = \{\rho(A), e^{i\phi_1} \rho(A), \dots, e^{i\phi_{h-1}} \rho(A)\}$ è chiuso rispetto alla moltiplicazione.

Ciò significa che anche $G := 1, e^{i\phi_1}, \dots, e^{i\phi_{h-1}}$ è chiuso rispetto alla moltiplicazione; G ha inoltre le seguenti proprietà: $1 \in G$, $e^{i\phi_k} \in G \Rightarrow e^{-i\phi_k} \in G$, $e^{i\phi_k} e^{i\phi_l} = e^{i\phi_{k+l}} = e^{i\phi_{l+k}} = e^{i\phi_l} e^{i\phi_k}$. Allora G è un gruppo abeliano di ordine h ; quindi la potenza h -esima di ogni suo elemento dà l'elemento neutro $(e^{i\phi_k})^h = 1 \quad \forall k = 0, \dots, h-1$. Ciò significa che G è l'insieme delle radici h -esime dell'unità $e^{\frac{2\pi i k}{h}} \quad k = 0, \dots, h-1$, quindi S sarà l'insieme delle radici h -esime di $\rho(A)$. \square

2.3 Matrici Primitive

Nella maggior parte delle applicazioni del teorema di Perron-Frobenius serve valutare $\lim_{k \rightarrow \infty} A^k$, ma le condizioni $A \geq 0$ e A irriducibile non sono sufficienti affinché questo limite esista.

Esempio 2.3.1. La matrice $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ è non negativa e irriducibile, i suoi autovalori sono $\lambda_1 = 1$, $\lambda_2 = -1$ e hanno entrambi modulo massimo; ma $\lim_{k \rightarrow \infty} A^k$ non esiste: infatti per k pari ho la matrice Identità $I \in \mathbb{R}^{2 \times 2}$, per k dispari ritrovo A .

L'ipotesi che manca nel teorema di Perron-Frobenius affinché questo limite esista è proprio l'unicità dell'autovalore di modulo massimo $\rho(A)$, come si può vedere dalle ipotesi del Lemma 1.4.8. Dobbiamo quindi dare ulteriori restrizioni alla classe delle matrici irriducibili, a questo ha pensato per la prima volta Frobenius nel 1912, dando la definizione di matrice primitiva.

Definizione 2.3.2. Sia $A \in \mathbb{R}^{n \times n}$, $A \geq 0$ e irriducibile è detta *matrice primitiva* se l'unico autovalore di modulo massimo è $\rho(A)$.

Teorema 2.3.3. Sia $A \in \mathbb{R}^{n \times n}$, $A \geq 0$ e primitiva, allora

$$\lim_{k \rightarrow \infty} \left(\frac{A}{\rho(A)} \right)^k = L > 0,$$

con $x, y > 0$ tali che $Ax = \rho(A)x$, $A^T y = \rho(A)y$, $x^T y = 1$ e $L := xy^T \in \mathbb{R}^{n \times n}$.

Dimostrazione. La dimostrazione è uguale a quella fatta nel Teorema 1.4.9, utilizzando il Lemma 1.4.8. \square

Teorema 2.3.4. Sia $A \in \mathbb{R}^{n \times n}$, $A \geq 0$, allora A è primitiva se e solo se $A^k > 0$ per qualche $k > 0$.

Dimostrazione. Innanzitutto dimostriamo che se $A \in \mathbb{R}^{n \times n}$ e λ autovalore di A , allora λ^k è autovalore di A^k e la molteplicità algebrica di λ è uguale alla molteplicità algebrica di λ^k .

Se λ autovalore di A , allora λ^k autovalore di A^k è ovvio, concentriamoci sulla molteplicità. Supponiamo che la molteplicità algebrica di λ sia m e sia A in forma di Jordan. Ci saranno diversi blocchi di Jordan relativi a λ la somma delle cui dimensioni è m . Supponiamo, senza perdere di generalità, ci sia un

unico blocco $m \times m$ relativo all'autovalore λ del tipo

$$J(\lambda) = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}$$

Se A è una matrice diagonale a blocchi, anche A^k lo è e i suoi blocchi sono le potenze k -esime dei blocchi di A , andiamo quindi a vedere chi è $J(\lambda)^k$. Facendo i conti e considerando che $\binom{k}{j} = 0 \quad \forall j > k$ otteniamo un blocco $m \times m$

$$J(\lambda)^k = \begin{pmatrix} \lambda^k & \binom{k}{1} \lambda^{k-1} & \binom{k}{2} \lambda^{k-2} & \dots & \binom{k}{m-1} \lambda^{k-m+1} \\ & \lambda^k & \binom{k}{1} \lambda^{k-1} & \ddots & \vdots \\ & & \ddots & \ddots & \binom{k}{2} \lambda^{k-2} \\ & & & \lambda^k & \binom{k}{1} \lambda^{k-1} \\ & & & & \lambda^k \end{pmatrix}.$$

Questo è l'unico blocco nella matrice A^k relativo all'autovalore λ^k : non ce ne possono essere altri. Dal momento che $J(\lambda)^k$ è un blocco $m \times m$, anche λ^k ha molteplicità m .

Ora siamo pronti per iniziare la dimostrazione vera e propria.

Supponiamo $A^k > 0$ per qualche k . Per prima cosa studiamo la riducibilità di A : se per assurdo A fosse riducibile, $A = P \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix} P^T$, con P matrice di permutazione. Allora $A^k = P \begin{pmatrix} X^k & R \\ 0 & Z^k \end{pmatrix} P^T$, dove R generico blocco; cioè A^k ha entrate nulle: assurdo. Quindi A è irriducibile.

Supponiamo che A abbia h autovalori $\lambda_1, \dots, \lambda_h$ di modulo massimo, cioè $\rho(A) = |\lambda_1| = \dots = |\lambda_h| > |\lambda_{h+1}| > \dots > |\lambda_n|$, allora anche $\lambda_1^k, \dots, \lambda_h^k$ sono tutti autovalori di modulo massimo di A^k . Il teorema di Perron assicura che A^k ha solo un autovalore di modulo massimo, allora $\rho(A^k) = \rho(A)^k = \lambda_1^k = \dots = \lambda_h^k$. Questo significa che la molteplicità algebrica di $\rho(A)$ è uguale a quella di $\rho(A^k)$, che è uguale a h . Dal teorema di Perron-Frobenius $h = 1$; quindi A è primitiva.

Viceversa: supponiamo $A \geq 0$ e primitiva, allora $\lim_{k \rightarrow \infty} \left(\frac{A}{\rho(A)} \right)^k = L > 0$ dal Teorema 2.3.3, ciò significa che esiste $K_\epsilon > 0$ tale che $\forall k > K_\epsilon \quad A^k > 0$. \square

Capitolo 3

Modellizzazione rete viaria e previsione andamento del traffico

Conoscere l'andamento del traffico è fondamentale per una città sia per migliorare la viabilità, sia per evitare picchi di inquinamento e di emissioni di CO_2 , di cui il traffico è la causa principale. Ci sono diversi modelli matematici sviluppati con tale scopo: quello che illustreremo di seguito è un esempio basato sulla teoria dei grafi e sul teorema di Perron-Frobenius applicato a matrici stocastiche. A tal proposito andiamo a vedere cos'è una matrice stocastica.

Definizione 3.0.5 (Matrice stocastica). Una *matrice stocastica* (per righe) $M \in \mathbb{R}^{n \times n}$ è una matrice quadrata non negativa, in cui la somma degli elementi di ciascuna riga è 1. Diremo che una matrice è stocastica per colonne se la somma degli elementi di ciascuna colonna è 1. Una matrice è *doppiamente stocastica* se è stocastica contemporaneamente per righe e per colonne.

Osserviamo che il raggio spettrale di una matrice stocastica è sempre uguale a 1 (si veda Teorema 1.3.10)

Attenzione: se abbiamo una matrice stocastica non possiamo applicare direttamente il teorema di Perron-Frobenius, dobbiamo prima assicurarci che sia irriducibile!

3.1 Modellizzazione rete stradale

Iniziamo schematizzando una ipotetica rete stradale: riportiamo, senza perdita di generalità, uno dei casi più semplici, costituito da tre incroci collegati da tre strade a doppio senso di circolazione. Questo modello può essere ampliato senza alcun problema nel caso dovessimo studiare una rete viaria più complessa. Nel grafo così ottenuto, chiamato *grafo-primario* i nodi rappresentano gli incroci e gli archi le strade che li collegano. Supponiamo che il numero di macchine totali in circolazione resti costante nel tempo: per tante macchine che parcheggiano tante si immettono in strada.

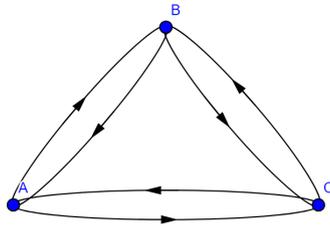


Figura 3.1: Rete stradale con tre incroci, collegati gli uni agli altri da strade a doppio senso di marcia.

Con un grafo-primario si ha immediatamente l'idea della struttura della rete stradale, ma non si riescono a includere importanti dati quantitativi quali limite di velocità, lunghezza della strada, numero di corsie, probabilità di svolta ad un incrocio etc... Trasformando il grafo-primario nel *grafo-duale* riusciamo a tenere conto di queste cose. Nel grafo-duale i nodi rappresentano le strade (es. il nodo AB rappresenta la strada che nel grafo-primario collegava l'incrocio A all'incrocio B) e gli archi le uscite di ogni incrocio.

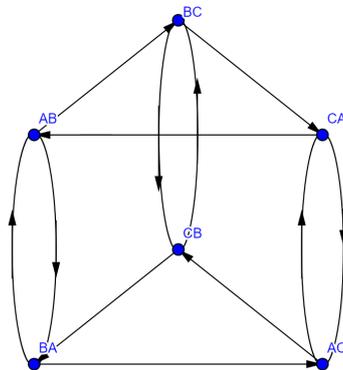


Figura 3.2: Grafo-duale del grafo 3.1.

A questo punto per procedere con la nostra trattazione è necessario dare alcune definizioni:

Definizione 3.1.1. Un **grafo pesato** è un grafo $G = (N, E)$ in cui a ciascun arco $E_i = (N_i, N_j)$ è associato un numero p_{ij} , detto peso dell'arco.

La **matrice di adiacenza** di un grafo orientato pesato è la matrice A che ha come elementi

$$a_{ij} = \begin{cases} 0 & \text{se non c'è nessun arco orientato da } N_i \text{ a } N_j \\ p_{ij} & \text{se c'è un arco orientato da } N_i \text{ a } N_j \text{ di peso } p_{ij} \end{cases}$$

Trascuriamo il tempo di percorrenza di ciascuna strada, assumendo che sia uguale per tutte. Supponiamo di conoscere la probabilità di svolta ad ogni incrocio: da un punto di vista pratico ciò può essere fatto ponendo delle videocamere ad ogni incrocio e contando il numero di veicoli che proseguono dritti o tornano indietro.

Attribuiamo a ogni arco del grafo-duale la probabilità che esso venga scelto dall'automobilista proveniente dalla strada (nodo) da cui l'arco ha origine. Per capire meglio guardiamo la figura 3.1

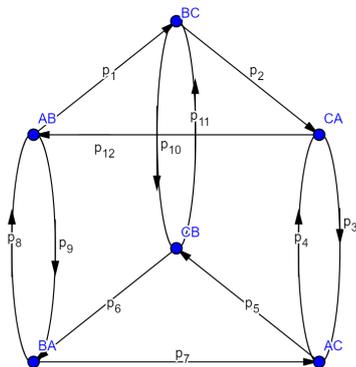


Figura 3.3: Grafo-duale G con indicate le probabilità di svolta ad ogni incrocio.

Così facendo abbiamo ottenuto un grafo orientato e pesato, andiamo a costruire la sua matrice di adiacenza A , rietichettando i nodi AB, AC, BA, BC, CA, CB, rispettivamente, con 1, 2, 3, 4, 5, 6.

$$A = \begin{pmatrix} 0 & 0 & p_9 & p_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_4 & p_5 \\ p_8 & p_7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_2 & p_{10} \\ p_{12} & p_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & p_6 & p_{11} & 0 & 0 \end{pmatrix}.$$

Osserviamo che il grafo G di figura 3.3 non è altro che il grafo associato ad A . Ma G è fortemente connesso, allora A è irriducibile. Se, infatti, G non fosse fortemente connesso, sarebbe impossibile andare da una strada all'altra all'interno della stessa città; ma questo non può mai accadere.

Nel grafo G la somma dei pesi degli archi uscenti da ogni nodo è 1, quindi la matrice delle probabilità A è stocastica per righe, perciò avrà raggio spettrale $\rho(A) = 1$. Adesso che sappiamo che la nostra matrice A è irriducibile, possiamo applicare il teorema di Perron-Frobenius.

Inoltre, facendo i conti, si vede che $A^3 > 0$, quindi, per il Teorema 2.3.4, A è primitiva: ha solo un autovettore di modulo massimo $\rho(A)$, e, soprattutto,

$$\lim_{n \rightarrow \infty} \left(\frac{A}{\rho(A)} \right)^n = \lim_{n \rightarrow \infty} A^n \text{ converge.}$$

Sia $v(0) \in \mathbb{R}^n$ un vettore che rappresenta la distribuzione del traffico al tempo $t = 0$: cioè $v(0) = (v_1(0), v_2(0), v_3(0), v_4(0), v_5(0), v_6(0))$ dove ogni $v_i(0)$ rappresenta il numero di veicoli che si trovano al tempo $t = 0$ nella strada i -esima (che corrisponde nel grafo al nodo i -esimo) diviso il numero di veicoli totali che circolano nella rete stradale. Consideriamo il tempo discreto. La distribuzione dei veicoli nella rete stradale dopo un tempo $t = n$ è data da $A^n v(0) =: v(n)$.

Per conoscere l'evoluzione del traffico a lungo termine valuto il limite di $v(n)$ per $n \rightarrow \infty$. Dal Teorema 2.3.3, tenendo presente che $\rho(A) = 1$, abbiamo

$$\lim_{n \rightarrow \infty} A^n v(0) = Lv(0),$$

con $L := xy^T \in \mathbb{R}^{n \times n}$, dove $x, y > 0$ tali che $Ax = x$, $A^T y = y$, $x^T y = 1$.

Allora $Lv(0) = xy^T v(0) = \alpha x$ per un certo $\alpha \in \mathbb{R}$.

Questo significa che dopo un lungo periodo di tempo la distribuzione del traffico si stabilizza, cioè esiste un $n_\epsilon > 0$ tale che per ogni $n > n_\epsilon$ il numero di macchine che si trovano nella strada i al tempo n diviso il numero di macchine totali in circolazione, è dato dalla componente i -esima del vettore αx , con x vettore di Perron di A o suo multiplo positivo.

Bibliografia

- [1] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, 1990.
- [2] Carl D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, 2000.
- [3] Emanuele Crisostomi, Steve Kirkland, Robert Shorten, *A Google-like Model of Road Network Dynamics and its Application to Regulation and Control*, 2010.
- [4] L. Pallottino, *Sistemi Robotici Distribuiti*, 4 Marzo 2015.