M. Giaquinta, G. Modica

Il teorema di punto fisso e applicazioni
Note 2004–2005

Versione preliminare

Mariano Giaquinta
Scuola Normale Superiore
Piazza dei Cavalieri, 7
I-56100 Pisa

`giaquinta@sns.it`

Giuseppe Modica
Dipartimento di Matematica
Applicata, Università di Firenze
Via S. Marta, 3
I-50139 Firenze

`modica@dma.unifi.it`

Indice

1. Il metodo dei minimi quadrati

1.a Minimi quadrati lineare

Supponiamo di avere dei dati sperimentali, ad esempio numeri $y_1, y_2, \dots, y_N$ e di voler dare un senso ai dati, naturalmente all'interno di una ipotesi sulla struttura dei dati.

Nel caso del *modello lineare* si suppone che *i dati debbano essere combinazioni lineari note di parametri arbitrari*: supponiamo cioè di avere una mappa lineare assegnata $\mathbf{A} : \mathbb{R}^k \to \mathbb{R}^N$ e, per ogni valore dei parametri $x = (x_1, x_2, \dots, x_k)$, di considerare la N-upla dei relativi *valori attesi* $\mathbf{A}^1 x, \dots \mathbf{A}^N x$ righe per colonne. Se, come è naturale, rappresentiamo i valori attesi per il parametro x come la N-upla $\mathbf{A}x$ di $\mathbb{R}^N$, l'insieme dei possibili valori attesi è

$$\Big\{ z \in \mathbb{R}^N \mid \exists x \in \mathbb{R}^k \text{tale che } \mathbf{A}x = z \Big\} = \mathrm{Im}\,(\mathbf{A}).$$

Si vuole ora selezionare il valore del parametro $x \in \mathbb{R}^k$ in modo da minimizzare una assegnata funzione *errore* o *costo* che misura lo scostamento tra i dati sperimentali e i valori attesi,

$$C(x) := F(\mathbf{A}x, y).$$

La scelta di $C(x)$ puo' essere fatta in base a considieriazioni di ordine vario, ad esempio sulla base di considerazioni statistiche sulla distribuzione degli errori.

Nel caso del metodo dei *minimi quadrati*, l'errore da minimizzare è lo *scarto quadratico medio tra i valori attesi e i valori misurati*,

$$C(x) := \sum_{i=1}^{N} |\mathbf{A}^i x - y_i|^2 = |\mathbf{A}x - y|^2 \to min,$$

o, più in generale,

$$C(x) = ||Ax - y||^2$$

dove $||\ ||$ è una norma derivante da un prodotto scalare in $\mathbb{R}^N$. Il problema dei *minimi quadrati* è dunque il seguente:

Assegnati un punto $y \in \mathbb{R}^N$, una mappa lineare $\mathbf{A} : \mathbb{R}^k \to \mathbb{R}^N$ e un prodotto scalare $\cdot \bullet \cdot$ in $\mathbb{R}^N$, trovare $x \in \mathbb{R}^k$ tale che

$$C(x) := (\mathbf{A}x - y) \bullet (\mathbf{A}x - y) = ||\mathbf{A}x - y||^2 \qquad \textit{sia minimo.}$$

In base al teorema della proiezione ortogonale, si tratta di trovare $x \in \mathbb{R}^k$ in modo che $\mathbf{A}x$ sia il piede della perpendicolare a $\mathrm{Im}\,(\mathbf{A})$ per y. Dunque

$$y - \mathbf{A}x \perp \mathrm{Im}\,(\mathbf{A})$$

e, per il teorema dell'alternativa, $y-\mathbf{A}x \in \ker \mathbf{A}^*$, i.e., x è soluzione dell'*equazione canonica*

$$\mathbf{A}^*\mathbf{A}x = \mathbf{A}^*y, \tag{1.1}$$

Se $\mathbf{A}^*\mathbf{A}$ è invertibile (il che accade se e solo se $\ker \mathbf{A} = 0$), allora vi è un'unica soluzione

$$x := (\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*y$$

e l'errore commesso vale

$$C(x) = ||\mathbf{A}x - y||^2 = ||\mathbf{A}(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*y - y||^2.$$

1.1 Esercizio. Sia $\mathbf{A} : \mathbb{R}^k \to \mathbb{R}^N$ lineare iniettiva. Allora l'operatore di proiezione ortogonale $\mathbf{P} : \mathbb{R}^N \to \mathbb{R}^N$ sull'immagine di $\mathbf{A}$ è

$$\mathbf{P} = \mathbf{A}(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*.$$

La risoluzione numerica dell'equazione canonica si opera riducendo la matrice simmetrica associata a A^*A a forma canonica, ad esempio mediante l'algoritmo SVD,

$$\mathbf{A}^*\mathbf{A} = \mathbf{R}^T\mathbf{L}\mathbf{R}, \qquad \mathbf{R}^T = \mathbf{R}^{-1},\ \mathbf{L} = \mathrm{diag}\,(\lambda_1, \dots, \lambda_n),$$

e quindi calcolando per ogni $i = 1, \dots, n$ $y^i := (\mathbf{R}\mathbf{A}^*b)^i/\lambda_i$ e quindi $x = \mathbf{R}^Ty$, $y = (y^1, y^2, \dots, y^n)$.

1.b Minimi quadrati non lineare

Analogamente, se il modello per i dati dipende in modo arbitrario ma C^1 dai parametri, i.e., se si suppone che i valori attesi siano dati da una assegnata funzione $\varphi : \mathbb{R}^k \to \mathbb{R}^N$ di classe C^1 dei parametri, il relativo problema dei minimi quadrati (questa volta non lineare) diventa quello di troavare x in modo da minimizzare la funzione

$$x \to ||\varphi(x) - y||^2.$$

Si ha

1.2 Proposizione. *Se x_0 è un minimo relativo per $||\varphi(x)-y||$ allora x_0 soddisfa il sistema di equazioni non lineari*

$$[\mathbf{D}\varphi(x_0)]^*(\varphi(x_0) - b) = 0. \tag{1.2}$$

Dimostrazione. Per il teorema di Fermat se x_0 è un tale punto di minimo,

$$0 = \frac{\partial||\varphi(x) - y||^2}{\partial x^i}(x_0) = 2\sum_{i=1}^{N}(\varphi(x_0) - y)^i\mathbf{D}\varphi^i(x_0)$$

e dunque $y - \varphi(x_0) \perp \ker D\varphi(x_0) = Tan_{x_0}\mathrm{Im}\,(\varphi)$. □

1.3 Esercizio. Supponiamo che un sistema sia descritto da una funzione lineare $y = mx + q$. Siano (x_i, y_i) n punti sperimentali in $\mathbb{R}^2$. Determinare m, q dai dati sperimentali in modo da minimizzare lo scarto quadratico medio, i.e. la funzione

$$S(m, q) := \sum_{i=1}^{n}(y_i - mx_i - q)^2.$$

2. Il teorema di punto fisso di Banach

Sia X uno spazio metrico. Una mappa $T : X \to X$ si dice una *contrazione* se esiste k, $0 \leq k < 1$, tale che $d(T(x), T(y)) \leq k\, d(x, y)$ $\forall x, y \in X$. Una contrazione è una funzione lipschitziana con costante di Lipschitz strettamente minore di 1, in particolare è una funzione continua su X. Si dice che $x \in X$ è un *punto fisso* o *punto unito* per $T : X \to X$ se $T(x) = x$.

2.1 Teorema (del punto fisso di Banach). *Sia X uno spazio metrico completo e $T : X \to X$ una contrazione, i.e., esiste k, $0 \leq k < 1$, tale che*

$$d(Tx, Ty) \leq k\, d(x, y) \qquad \forall\ x, y \in X.$$

Allora T ha un unico punto fisso. Inoltre per ogni $x_0 \in X$, la successione $\{x_n\}$ definita per $n = 1, 2 \dots$ da $x_{n+1} := T(x_n)$ converge geometricamente a x e valgono le stime

$$\begin{aligned} d(x_n, x) &\leq \frac{k^n}{1-k} d(x_1, x_0), \\ d(x_{n+1}, x) &\leq \frac{k}{1-k} d(x_{n+1}, x_n), \\ d(x_{n+1}, x) &\leq k d(x_n, x). \end{aligned}$$

Dimostrazione. (i) (Unicità) Se x e y sono due punti fissi, da $d(x, y) = d(Tx, Ty) \leq kd(x, y)$, $0 \leq k < 1$ segue $d(x, y) = 0$, i.e., $x = y$.

(ii) (Esistenza) Sia $x_0 \in X$ e per $n \geq 1$ sia $x_{n+1} := T(x_n)$. Si ha

$$d(x_{n+1}, x_n) \leq kd(x_n, x_{n-1}) \leq k^n d(x_1, x_0) = k^n d(T(x_0), x_0),$$

dunque per la disuguaglianza triangolare, per $p > n$

$$d(x_p, x_n) \leq \sum_{j=n}^{p-1} d(x_{j+1}, x_j) \leq \sum_{j=n}^{p-1} k^j d(x_1, x_0) \leq \frac{k^n}{1-k} d(x_1, x_0) \to 0$$

per $n, p \to \infty$. Quindi $\{x_n\}$ è di Cauchy e converge. Il limite è necessariamente un punto fisso, come si vede passando al limite per $n \to \infty$ nella relazione $x_{n+1} = T(x_n)$. Lasciamo al lettore la facile dimostrazione delle stime di convergenza. $\square$

Osserviamo che la prima stima nella tesi del Teorema 2.1 permette di stimare il numero delle iterazioni sufficiente a raggiungere una fissata approssimazione. La seconda stima invece permette di valutare l'accuratezza di x_{n+1} come valore approssimato di x in termini di $d(x_{n+1}, x_n)$.

2.2 Esercizio. Sia $\phi : X \to X$ una applicazione anche non lineare da uno spazio di Banach in sé e $y \in X$. Vogliamo risolvere l'equazione

$$\phi(x) = y. \tag{2.1}$$

(Si pensi ad esempio ad $X = \mathbb{R}^n$. In questo caso la (2.1) è un sistema di n equazioni ad n incognite, in generale nonlineare).

Scrivendo l'equazione come $x = x - \phi(x) + y$ e ponendo $g(x) := x - \phi(x) + y$ la (2.1) è equivalente a trovare i punti fissi di g, i.e., gli x tali che

$$x = g(x).$$

Pertanto, se g è una contrazione su X, segue dal teorema di punto unito che $\phi(x) = y$ ha una ed una soluzione $\overline{x}$. Inoltre ogni successione $\{x_n\}$ definita da

$$\begin{cases} x_0 \in X, \\ x_{n+1} = x_n - \phi(x_n) + y \end{cases} \tag{2.2}$$

converge con velocità almeno esponenziale ad $\overline{x}$. Un caso particolare è quello in cui $X = \mathbb{R}^n$ e $\phi(x) = \mathbf{L}x$ è lineare. In questo caso $g(x) = (\,\mathrm{Id} - \mathbf{L})x + y$ e g è una contrazione se e solo se $\mathbf{L}$ è sufficientemente vicina all'identità, $||\,\mathrm{Id} - \mathbf{L}|| < 1$. Infatti

$$\sup_{x_1,x_2\in\mathbb{R}^n} \frac{|g(x_2) - g(x_1)}{|x_2 - x_1|} = \sup_{x_1,x_2\in\mathbb{R}^n} \frac{|(\,\mathrm{Id} - \mathbf{L})(x_2 - x_1)|}{|x_2 - x_1|} = ||\,\mathrm{Id} - \mathbf{L}||.$$

Inoltre (2.2) si riscrive come

$$x_{n+1} = (\,\mathrm{Id} - \mathbf{L})^{n+1} x_0 = \sum_{k=0}^{n} (\,\mathrm{Id} - \mathbf{L})^k y$$

e quindi per $n \to \infty$

$$\overline{x} = \sum_{k=0}^{\infty} (\,\mathrm{Id} - \mathbf{L})^k y$$

2.3 Esercizio. Una piccola variante del risultato precedente è la seguente. Sia $\phi : X \to X$ una applicazione da uno spazio di Banach X in sé e si voglia risolvere per un dato $y \in X$ l'equazione

$$\phi(x) = y.$$

Sia M un operatore invertibile noto $M : X \to X$. Riscrivendo l'equazione $\phi(x) = y$ come $Mx = Mx - \phi(x) + y$, è facile convincersi che x è solusione di $\phi(x) = y$ se e solo se x è un punto fisso di

$$x = T(x), \qquad T(x) := x - M^{-1}\phi(x) + M^{-1}y$$

Dal teorema delle contrazioni si conclude quindi che se T è una contrazione su X allora $\phi(x) = y$ ha un'unica soluzione $\overline{x} \in X$ e la successione

$$\begin{cases} x_0 \in X, \\ x_{n+1} = x_n + M^{-1}\phi(x_n) + M^{-1}y \end{cases}$$

converge ad $\overline{x}$ con velocità esponenziale.

2.4 Esercizio. Provare il seguente

Teorema. *Sia* $\mathbf{A} \in M_{n,n}(\mathbb{C})$*. La serie a valori matrici* $\sum_{n=0}^{\infty} \mathbf{A}^n \frac{1}{z^n}$ *converge in ogni* z *con* $|z| > ||\mathbf{A}||$*, la matrice* $\left(\,\mathrm{Id} - \frac{\mathbf{A}}{z}\right)$ *è invertibile e*

$$\sum_{n=0}^{\infty} \mathbf{A}^n \frac{1}{z^n} = \left(Id - \frac{\mathbf{A}}{z}\right)^{-1}.$$

2.5 Esercizio. Sia X di Banach e $T : X \to X$ lipschitziana. Provare che se μ è abbastanza grande in valore assoluto allora, per ogni $y \in X$, l'equazione

$$Tx + \mu x = y$$

ha unica soluzione.

3. Invertibilità di mappe $\mathbb{R}^n \longrightarrow \mathbb{R}^n$

Sia $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ una applicazione. Se f è lineare, $f(x) = \mathbf{A}x$, $A \in M_{n,n}(\mathbb{R})$, l'invertibilità di f, i.e. la risolubilità in x di $\mathbf{A}x = y$ per ogni $y \in \mathbb{R}^n$, è equivalente all'invertibilità della matrice $\mathbf{A}$ e questa alla condizione $\det \mathbf{A} \neq 0$.

Per funzioni $f : \mathbb{R} \to \mathbb{R}$ derivabili, abbiamo visto con i metodi del calcolo che la condizione $f' > 0$ (o $f' < 0$) è sufficiente a garantire la monotonia e quindi la invertibilità di f e la derivabilità della funzione inversa. Inoltre per funzioni con derivata continua l'essere $f'(x_0) \neq 0$ in un punto x_0, assicura l'esistenza di un intervallo $I(x_0, r)$ su cui $f'(x)$ ha lo stesso segno di $f'(x_0)$, da cui la stretta monotonia, la continuità, la derivabilità di $(f_{|I})^{-1}$ e la formula

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))} \qquad \forall y \in f(I(x_0, r)).$$

Per funzioni $f : \mathbb{R}^n \to \mathbb{R}^n$ di classe C^1, la condizione $f'(x_0) \neq 0$ in una variabile, viene sostituita dalla condizione $\det \mathbf{D}f(x_0) \neq 0$ di non degenerazione della matrice jacobiana, una similitudine piuttosto naturale se la si interpreta geometricamente come una condizione di non degenerazione dell'applicazione lineare tangente $h \to \mathbf{D}f(x_0)(h)$, o, se si vuole, di non degenerazione dello sviluppo di Taylor al primo ordine di f in x_0,

$$f(x) = f(x_0) + \nabla f(x_0) \bullet (x - x_0) + o(|x - x_0|) \qquad \text{per } x \to x_0.$$

Sia $f = (f^1, f^2, \dots, f^n)$ una applicazione da un aperto $\Omega \subset \mathbb{R}^n$ in $\mathbb{R}^n$, $n \geq 1$, di classe C^1. Ricordiamo che $f_{|U}$ denota la restrizione di f ad $U \subset \Omega$ e che con $x_0 + U$ si intende il sottoinsieme $\{x \in \mathbb{R}^n \mid x - x_0 \in U\}$, ad esempio $x_0 + B(0, r) = B(x_0, r)$.

3.1 Definizione. *Una funzione $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ si dice* localmente invertibile *se per ogni $x \in \Omega$ esiste un intorno U di x tale che $f_{|U}$ è iniettiva.*

3.2 Teorema (di invertibilità locale). *Sia $f : \Omega \to \mathbb{R}^n$ una applicazione di classe C^k, $k \geq 1$, definita su un aperto $\Omega \subset \mathbb{R}^n$ e sia $x_0 \in \Omega$. Se $\det \mathbf{D}f(x_0) \neq 0$, allora esiste un intorno aperto U di x_0 tale che*

(i) *f ristretta ad U è iniettiva,*
(ii) *$V := f(U)$ è aperto, f è aperta e $(f_{|U})^{-1}$ è continua,*
(iii) *$(f_{|U})^{-1} : V \to U$ è di classe C^k ed inoltre $\forall\, y \in V$*

$$\mathbf{D}(f_{|U})^{-1}(y) = \Big[\mathbf{D}f(x)\Big]^{-1} \qquad x := (f_{|U})^{-1}(y). \tag{3.1}$$

Pertanto, se Ω è aperto in $\mathbb{R}^n$, $f : \Omega \to \mathbb{R}^n$ è di classe $C^k(\Omega)$ e $\det \mathbf{D}f(x) \neq 0$ $\forall x \in \Omega$, allora f è localmente invertibile, aperta e con inversa di classe C^k.

Dimostrazione. Non è restrittivo supporre che $x_0 = 0$ e che $f(x_0) = 0$.

Passo 1. Per comodità siano

$$\mathbf{M} := \mathbf{D}f(0)^{-1} \qquad M := ||\mathbf{M}|| \qquad \text{e} \qquad F(x) := f(x) - \mathbf{D}f(0)x.$$

Riscriviamo l'equazione $f(x) = y$ come

$$x = x + \mathbf{M}(-f(x) + y).$$

Evidentemente $f(x) = y$ se e solo se x è un punto unito $x = T_y x$ per la mappa

$$T_y(x) := x - \mathbf{M}f(x) + \mathbf{M}y = -\mathbf{M}F(x) + \mathbf{M}y.$$

Essendo $f \in C^1(\Omega, \mathbb{R}^n)$ per ipotesi, esiste $r > 0$ tale che

$$\sup_{||z|| \leq r} ||\mathbf{D}f(z) - \mathbf{D}f(0)|| < \frac{1}{4\,M}$$

e dalla formula del valor medio,

$$|F(x) - F(z)| \leq \frac{1}{4\,M}|x - z| \qquad \forall x, z \in \overline{B(0, r)}. \tag{3.2}$$

Sia $X := \overline{B(0, r)} \subset \mathbb{R}^n$. Proviamo che T_y è una contrazione su X per ogni $y \in B(0, \frac{r}{2M})$. Si ha infatti per ogni $x, z \in \overline{B(0, r)}$

$$\begin{aligned} |T_y(x)| &= |\mathbf{M}y - \mathbf{M}F(x)| \leq |\mathbf{M}y| + |\mathbf{M}F(x) - \mathbf{D}F(0)| \\ &\leq M|y| + M\frac{1}{4M}|x - 0| = \frac{r}{2} + \frac{r}{4} = \frac{3}{4}\,r \end{aligned}$$

e

$$|T_y(x) - T_y(z)| = |\mathbf{M}F(x) - \mathbf{M}F(z)| \leq M\frac{1}{4M}|x - z| = \frac{1}{4}|x - z|.$$

Dunque per ogni $y \in B(0, r/(2M))$ la funzione $x \to T_y x$ è una contrazione su X con fattore di contrazione inferiore a $1/2$ e con immagine in $B(0, 3r/4)$. Essendo X uno spazio metrico completo, segue dal teorema di punto fisso di Banach l'esistenza di un unico $x \in B(0, 3r/4)$ tale che $x = T_y x$, equivalentemente $\phi(x) = y$.

Ponendo $U := f^{-1}(B(0, r/(2M)))$, $V := B(0, r/(2M))$, si è così provato che $f_{|U}$ è invertibile e qindi la (i).

Passo 2. Dimostriamo ora la (ii). Basta provare che $(f_{|U})^{-1} : V \to U$ è continua. Poniamo per comodità per $y, w \in V$ $x := (f_{|U})^{-1}(y)$ e $z := (f_{|U})^{-1}(w)$. Si ha allora dalla (3.2)

$$\begin{aligned} |Df(0)(x - z)| &= |-f(x) + Df(0)x + f(z) - Df(0)z + f(x) - f(z)| \\ &\leq |F(x) - F(z)| + |f(x) - f(z)| \\ &\leq \frac{1}{4M}\,|x - z| + |f(x) - f(z)| \end{aligned}$$

da cui

$$|x - z| \leq \frac{1}{4}|x - z| + M|f(x) - f(z)|$$

i.e.,

$$|(f_{|U})^{-1}(y) - (f_{|U})^{-1}(w)| \leq \frac{4M}{3}|y - w| \tag{3.3}$$

Passo 3. Resta da provare la differenziabilità dell'inversa di $f_{|U}$, $g := (f_{|U})^{-1}$ in ogni punto $y \in V = f(U)$. Non è restrittivo supporre che $y = 0$ e che $g(0) = 0$. Posto $x = g(z)$ si ha

$$g(z) - \mathbf{M}z = x - \mathbf{M}f(x) = -\mathbf{M}F(x)$$

e d'altra parte dalla (3.3) $|f(x)| \geq \frac{3M}{4}|x|$ e quindi

$$\frac{|g(z)-\mathbf{M}z|}{|z|} = \frac{|\mathbf{M}(f(x)-\mathbf{D}f(0)x)|}{|x|}\frac{|x|}{|f(x)|} \le M\frac{o(|x|)}{|x|}\frac{|x|}{|f(x)|} \le \frac{4}{3}M^2\frac{o(|x|)}{|x|}.$$

Quando $z \to 0$ anche $x = g(z) \to 0$ per la continuità di g. Il secondo membro tende quindi a zero per $z \to 0$. g è quindi differenziabile con $\mathbf{D}g(0) = \mathbf{M} = \mathbf{D}f(0)^{-1}$.

Infine se f è di classe C^k, $k \ge 2$, segue immediatamente dalla (3.1) che $(f_{|U})^{-1}$ è di classe $C^k(V)$. □

3.3 Osservazione. Con riferimento alla dimostrazione del Teorema 3.2, ricordiamo che il punto fisso di una contrazione si ottiene con un processo di approssimazione. La dimostrazione del teorema si basa quindi su un algoritmo. Con le notazioni del Teorema 3.2, per ogni $y \in B(0, r/(2M))$, si è infatti provato che le successioni definite da

$$\begin{cases} x_0 \in \overline{B(0,r)}, \\ x_{k+1} = x_k - \mathbf{M}f(x_k) + \mathbf{M}y, \end{cases}$$

convergono tutte con velocità almento esponenziale all'unica soluzione x di $f(x) = y$ in $\overline{B(0,r)}$.

4. Sistemi di equazioni differenziali ordinarie lineari

In questo capitolo discutiamo l'esistenza e la risolubilità del problema di Cauchy associato ad un istema di N equazioni lineari omogenee, i.e., della esistenza e unicità di una funzione $X : [0,T] \to \mathbb{C}^N$ di classe C^1 soluzione del sistema

$$\begin{cases} X(0) = X_0, \\ X'(t) = \mathbf{A}(t)X(t) + F(t), \qquad \forall t \in [0,T] \end{cases} \tag{4.1}$$

pensando come assegnati il *dato iniziale* $X_0 \in \mathbb{C}^N$, la funzione a valori matrici $t \to \mathbf{A}(t)$ e il termine noto $F : [0,T] \to \mathbb{R}^N$. *Supporreme sempre che* $\mathbf{A}(t)$ *e* $F(t)$ *siano funzioni continue.*

4.a Esistenza e unicità

Indichiamo con $||\mathbf{A}|| := \sup_{|x|=1} |\mathbf{A}x|$ la norma di $\mathbf{A} \in M_{n,n}$, im modo che

$$|\mathbf{A}x| \le ||\mathbf{A}||\,|x| \qquad \forall x,$$

e con

$$M := \sup_{t\in[0,T]} ||\mathbf{A}(t)||.$$

Si ha

4.1 Teorema. *Il problema di Cauchy* (4.1) *ha una e una sola soluzione. La si puo' ottenere come limite* uniforme *della successione* $\{\alpha_n(t)\}$ *di funzioni definite iterativamente da*

$$\begin{cases} \alpha_0(t) = X_0, \\ \alpha_{n+1}(t) := X_0 + \int_0^t (\mathbf{A}(s)\alpha_n(s) + F(s))\,ds. \end{cases} \tag{4.2}$$

4.2 Lemma. *Sono fatti equivalenti*

(i) $X(t)$ *è di classe* $C^1([0,T])$ *ed è soluzione di* (4.1),
(ii) $X(t)$ *è continua e*

$$X(t) = X_0 + \int_0^t (\mathbf{A}(s)X(s) + F(s))\,ds. \tag{4.3}$$

Dimostrazione. Applicare il teorema fondamentale del calcolo. □

La (4.3) caratterizza la soluzione $X(t)$ di (4.1) come *punto fisso* per la mappa

$$T : X(t) \longrightarrow X_0 + \int_0^t (\mathbf{A}(s)X(s) + F(s))\, ds. \tag{4.4}$$

Proviamo il Teorema 4.1 tesi applicando opportunamente il teorema di punto fisso.

Sia $\gamma > 0$ un numero che sceglieremo fra breve. Consideriamo lo spazio $\mathcal{S}$ delle funzioni continue su $[0, T]$ a valori in $\mathbb{R}^N$. La quantità

$$||X||_\gamma := \sup_{t \in [0,T]} \Big(|X(t)| e^{-\gamma t} \Big)$$

ha evidentemente le proprietà di una norma (è finita su ogni funzione continua su $[0, T]$, è positivamente omogenea, vale la diseguaglianza triangolare e vale zero se e solo se $X(t) = 0$) e dunque $(\mathcal{S}, ||\ ||_\gamma)$ è uno spazio normato. Poi

$$e^{-\gamma T} ||X||_{\infty,[0,T]} \le ||X||_\gamma \le ||X||_{\infty,[0,T]};$$

dunque le successioni convergenti e le successioni di Cauchy per la norma $||\ ||_\gamma$ sono rispettivamente le successioni convergenti e le successioni di Cauchy per la *norma uniforme* $||\ ||_{\infty,[0,T]}$. Poiche' $C^0([0, T], \mathbb{R}^n)$ con la norma uniforme è uno spazio di Banach, si conclude che $(\mathcal{S}, ||\ ||_\gamma)$ *è uno spazio di Banach.*

Proviamo ora

4.3 Proposizione. *La mappa $T : \mathcal{S} \to \mathcal{S}$ definita in* (4.4) *è una contrazione per la norma $||\ ||_\gamma$ se*

$$\gamma > M := \sup_{t \in [0,T]} ||\mathbf{A}(t)||.$$

Dimostrazione. Infatti $\forall X, Y \in \mathcal{S}$ e $t \in [0, T]$,

$$\begin{aligned}
|TX(t) - TY(t)| &= \left| \int_0^t \mathbf{A}(s)(X(s) - Y(s))\, ds \right| = \left| \int_0^t \mathbf{A}(s)(X(s) - Y(s)) e^{-\gamma s} e^{\gamma s}\, ds \right| \\
&\le \int_0^t ||\mathbf{A}(s)|| \Big(|X(s) - Y(s)| e^{=\gamma s} \Big) e^{\gamma s}\, ds \\
&\le M ||X - Y||_\gamma \int_0^t e^{\gamma s}\, ds = \frac{M}{\gamma} ||X - Y||_\gamma e^{\gamma t},
\end{aligned}$$

da cui passando al sup,

$$||TX - TY||_\gamma \le \frac{M}{\gamma} ||X - Y||_\gamma.$$

□

Dimostrazione del Teorema 4.1. Scegliamo $\gamma = 2M$. Per la Proposizione 4.3, $T : \mathcal{S} \to \mathcal{S}$ è una contrazione. Duqneu T ha un unico punto fisso, i.e., esiste un'unica funzione continua $X(t)$ soluzione di (4.3). Il Lemma 4.2 assicura allora che $X(t)$ trovata è la soluzione cercata del problema, e l'algoritmo di calcolo fornito dal teorema di punto fisso si traduce in questo caso nello schema di approssimazione (4.2). □

4.4 Osservazione. Supponiamo per un momento che $\mathbf{A}$ sia costante. In questo caso le approssimazioni della soluzioni sono

$$\alpha_n(t) = \Big(\sum_{k=0}^{n} \frac{\mathbf{A}^k t^k}{k!}\Big) X_0$$

e dunque è ragionevole indicare la soluzione $X(t)$ del problema (4.1) con

$$X(t) = \Big(\sum_{n=0}^{\infty} \frac{\mathbf{A}^n\, t^t}{n!}\Big) X_0 =: e^{t\mathbf{A}} X_0.$$

4.b La matrice di transizione

Per ogni $(t_0, x_0) \in I \times \mathbb{R}^n$, il sistema di equazioni differenziale in forma normale

$$x' = \mathbf{A}(t)x + F(t), \qquad \mathbf{A} \in C^0(I, M_{n,n}(\mathbb{R})),\ F \in C^0(I, \mathbb{R}^n),$$

ha una ed una sola soluzione $x(t) = x(t; t_0, x_0)$ definita su tutto I, di classe C^1 e tale che $x(t_0) = x_0$, cfr. Teorema 4.1. Nel caso omogeneo,

$$x' = \mathbf{A}(t)x, \qquad \mathbf{A} \in C^0(I, M_{n,n}(\mathbb{R})),$$

l'applicazione $x_0 \to x(t; t_0, x_0)$ da $\mathbb{R}^n$ in sé è chiaramente lineare; resta quindi associata una matrice $\mathbf{W}(t, t_0)$ tale che

$$x(t; t_0, x_0) = \mathbf{W}(t, t_0)x_0$$

La matrice $\mathbf{W}(t, s)$ si chiama *matrice di transizione* (dallo stato al tempo s allo stato al tempo t) e ogni soluzione $x(t)$ di $x' = \mathbf{A}x$ verifica

$$x(t) = \mathbf{W}(t, s)x(s) \qquad \forall t, s \in I.$$

Per definizione le colonne W_1, $W_2, \dots,$ W_n della matrice di transizione $\mathbf{W}(t, s)$, sono rispettivamente le soluzioni dei problemi di Cauchy

$$\begin{cases} W_j'(t) = \mathbf{A}(t)W_j(t), \\ W_j(s) = e_j \end{cases} \qquad j = 1, \dots, n,$$

dove $\{e_1, e_2, \dots, e_n\}$ è la base canonica di $\mathbb{R}^n$. Sono quindi una base dello spazio delle soluzioni dell'equazione omogenea $x' = \mathbf{A}(t)x$. Viceversa se $\mathbf{Z}(t) : I \to M_{n,n}$ ha come colonne n soluzioni di $x' = \mathbf{A}(t)x$, allora

$$\mathbf{Z}(t) = \mathbf{W}(t, s)\mathbf{Z}(s)$$

e, se $\det \mathbf{Z}(t) \neq 0\ \forall t$, $\mathbf{Z}(t)$ si chiama una una *matrice fondamentale* del sistema e

$$\mathbf{W}(t, s) = \mathbf{Z}(t)\mathbf{Z}(s)^{-1}.$$

4.5 Proposizione. *Sia $\mathbf{W}(t, s)$ la matrice di transizione associata ad $\mathbf{A}(t)$, allora*

(i) $\mathbf{W}(t,t) = \mathrm{Id}$ *per ogni* $t \in I$,
(ii) $\mathbf{W}(t,s)\mathbf{W}(s,r) = \mathbf{W}(t,r)$,
(iii) $\mathbf{W}(t,s)^{-1} = \mathbf{W}(s,t)$,
(iv) *Si ha*

$$\frac{\partial \mathbf{W}(t,s)}{\partial t} = \mathbf{A}(t)\mathbf{W}(t,s), \qquad \frac{\partial \mathbf{W}(t,s)}{\partial s} = -\mathbf{W}(t,s)\mathbf{A}(s).$$

(v) *Vale l'*equazione di Liouville

$$\frac{\partial}{\partial t}\det \mathbf{W}(t,s) = \mathrm{tr}\,(\mathbf{A}(t))\det \mathbf{W}(t,s),$$

ed in particolare la formula di Abel

$$\det \mathbf{W}(t,s) = \exp\left(\int_s^t \mathrm{tr}\,\mathbf{A}(\tau)\,d\tau\right).$$

Dimostrazione. Lasciamo al lettore la verifica delle (i) (ii),…(iv) e proviamo (v). Abbiamo

$$\begin{aligned}\mathbf{W}(t+\epsilon,s) = \mathbf{W}(t+\epsilon,t)\mathbf{W}(t,s) &= \Big(\mathrm{Id} + \epsilon\frac{\partial \mathbf{W}}{\partial t}(t,t) + o(\epsilon)\Big)\mathbf{W}(t,s)\\ &= \Big(\mathrm{Id} + \epsilon\mathbf{A}(t) + o(\epsilon)\Big)\mathbf{W}(t,s) = (\mathrm{Id} + \epsilon\mathbf{A}(t))\mathbf{W}(t,s) + o(\epsilon)\end{aligned}$$

e quindi

$$\det \mathbf{W}(t+\epsilon,s) = \det(\mathrm{Id} + \epsilon\mathbf{A}(t))\det \mathbf{W}(t,s) + o(\epsilon) = \mathrm{tr}\,\mathbf{A}(t)\det \mathbf{W}(t,s) + o(\epsilon)$$

da cui (v). □

4.6 Esercizio. Provare che

Proposizione. *Una matrice* $\mathbf{Z}$ *le cui colonne sono soluzioni del sistema* $x' = \mathbf{A}(t)x$ *è una matrice fondamentale se e solo se esiste* $t_0 \in I$ *tale che* $\det \mathbf{Z}(t_0) \neq 0$.

[*Sugg.* Usare la formula di Abel.]

4.7 Esercizio. Osservando che $\mathbf{W}(t,s) = \mathbf{W}(t,t_0)\mathbf{W}(t_0,s)$, provare che $\mathbf{W}(t,s)$ è di classe C^1 (nella coppia di variabili (t,s)).

Per verifica diretta o con il metodo della *variazione delle costanti*, cioè cercando una soluzione particolare del tipo $u(t) := \mathbf{W}(t,s)c(t)$, $c(t) \in \mathbb{R}^n$, si conclude

4.8 Teorema. *La soluzione del problema di Cauchy*

$$\begin{cases} x' = \mathbf{A}(t)x + f(t),\\ x(t_0) = x_0\end{cases}$$

è data da

$$x(t) = \mathbf{W}(t,t_0)x_0 + \int_{t_0}^t \mathbf{W}(t,\tau)f(\tau)\,d\tau$$

dove $\mathbf{W}(t,s)$ *è la matrice di transizione associata a* $\mathbf{A}(t)$.

Definiamo per induzione

$$\begin{cases} \mathbf{W}_0(t,s) := \mathrm{Id}, \\ \mathbf{W}_{k+1}(t,s) := \int_s^t \mathbf{A}(\tau)\mathbf{W}_k(\tau,s)\,d\tau, \quad k \ge 0, \end{cases} \tag{4.5}$$

cioè

$$\begin{cases} \mathbf{W}_0(t,s) = \mathrm{Id}, \\ \mathbf{W}_1(t,s) = \int_s^t \mathbf{A}(\tau)\,d\tau, \\ \mathbf{W}_k(t,s) = \int_s^t \int_s^{\tau_k} \dots \int_s^{\tau_2} \mathbf{A}(\tau_k)\cdots\mathbf{A}(\tau_1)\,d\tau_1\cdots d\tau_k, \quad k \ge 2. \end{cases}$$

Applicando il teorema delle contrazioni come nel caso del teorema di esistenza, cfr. Vol. III Cap. XI Sez. 3.1, e usando la linearità, si trova

4.9 Proposizione. *Per ogni intervallo $J \subset\subset I$*

$$|\mathbf{W}_k(t,s)| \le ||\mathbf{A}||^k_{\infty,J}\frac{|t-s|^k}{k!} \qquad \forall t,s \in J,$$

dove $||\mathbf{A}||_{\infty,J} := \sup_{t\in J}||\mathbf{A}(t)||$. Quindi la serie $\sum_{k=0}^\infty \mathbf{W}_k(t,s)$ converge uniformemente sui compatti di $I \times I$ alla matrice di transizione,

$$\mathbf{W}(t,s) = \sum_{k=0}^\infty \mathbf{W}_k(t,s),$$

e vale la stima

$$||\mathbf{W}(t,s)|| \le e^{|t-s|\,||\mathbf{A}||_{\infty,J}} \qquad \forall t,s \in J \subset\subset I.$$

4.c Sistemi a coefficienti costanti

Supponiamo ora che $\mathbf{A}(t)$ commuti con

$$\mathbf{B}(t) := \int_s^t \mathbf{A}(\tau)\,d\tau.$$

Questo succede ad esempio se $\mathbf{A}(t)$ e $\mathbf{A}(\tau)$ commutano per ogni t e τ,

$$[\mathbf{A}(t), \mathbf{A}(\tau)] := \mathbf{A}(t)\mathbf{A}(\tau) - \mathbf{A}(\tau)\mathbf{A}(t) = 0,$$

in particolare, se $\mathbf{A}(t) := \mathbf{A}$ è una matrice costante. In questo caso

$$\frac{d}{dt}\mathbf{B}^k(t) = k\mathbf{B}^{k-1}(t)\mathbf{B}'(t) = k\mathbf{B}^{k-1}(t)\mathbf{A}(t).$$

Si ricava allora dalla (4.5) che, per $k \ge 0$,

$$\mathbf{W}_k(t,s) = \frac{1}{k!}\mathbf{B}^k(t)$$

e quindi

$$\mathbf{W}(t,s)=\sum_{k=0}^{\infty}\frac{1}{k!}\mathbf{B}^k(t)=\exp\Big(\int_s^t \mathbf{A}(\tau)\,d\tau\Big), \tag{4.6}$$

Nel caso in cui il sistema ha coefficienti costanti, $\mathbf{A}(t)=\mathbf{A}$, si ha allora

$$\mathbf{W}(t,s)=e^{(t-s)\mathbf{A}}$$

Notiamo che questo implica che $\mathbf{W}(t,s)=\mathbf{W}(t-s,0)$ e che per ogni matrice fondamentale $\mathbf{Z}(t)$ di $x'=\mathbf{A}x$, si ha $\mathbf{Z}(t)\mathbf{Z}(s)^{-1}=\mathbf{Z}(t-s)\mathbf{Z}(0)^{-1}$, formule queste ultime che si possono provare direttamente. Concludiamo quindi

4.10 Proposizione. *Sia $\mathbf{A}\in M_{n,n}$. La soluzione del problema di Cauchy*

$$\begin{cases} x'(t)=\mathbf{A}x(t)+f(t),\\ x(t_0)=x_0\end{cases}$$

è data da

$$x(t)=e^{(t-t_0)\mathbf{A}}x_0+\int_{t_0}^t e^{(t-s)\mathbf{A}}f(s)\,ds.$$

4.11 Esercizio. Sia $\mathbf{Z}(t)$ la matrice fondamentale con $\mathbf{Z}(0)=\mathrm{Id}$. Mostrare che $\mathbf{Z}(t)=e^{t\mathbf{A}}$ $\forall t\in\mathbb{R}$. Osservare che $\mathbf{Z}(t+s)=\mathbf{Z}(t)\mathbf{Z}(s)$ $\forall t,s\in\mathbb{R}$.

4.12 Esercizio. Provare le seguenti formule

(i) $e^{\mathbf{A}}=\sum_{k=0}^{\infty}\frac{\mathbf{A}^k}{k!}=\lim_{n\to\infty}\Big(\mathrm{Id}+\frac{1}{n}\mathbf{A}\Big)^n$,

(ii) $\Big|\Big|e^{\mathbf{A}}\Big|\Big|\le e^{||\mathbf{A}||}$,

(iii) Se $\mathbf{AB}=\mathbf{BA}$, allora $e^{\mathbf{A}}\mathbf{B}=\mathbf{B}e^{\mathbf{A}}$, $e^{\mathbf{A}+\mathbf{B}}=e^{\mathbf{A}}e^{\mathbf{B}}$, $\Big(e^{\mathbf{A}}\Big)^{-1}=e^{-\mathbf{A}}$, $\frac{d}{dt}e^{t\mathbf{A}}=\mathbf{A}e^{t\mathbf{A}}$,

(iv) Se $\mathbf{B}$ è invertibile, allora $e^{\mathbf{BAB}^{-1}}=\mathbf{B}e^{\mathbf{A}}\mathbf{B}^{-1}$,

(v) $\det e^{\mathbf{A}}=e^{\operatorname{tr}\mathbf{A}}$,

4.d Equazioni di ordine superiore

Come già sappiamo, una equazione differenziale di ordine n in forma normale lineare,

$$u^{(n)}+a_{n-1}(t)u^{(n-1)}+\dots+a_1(t)u'+a_0(t)u=f(t), \tag{4.7}$$

si riconduce ad un sistema lineare del primo ordine nella incognita $v:=(u,u',\dots,u^{(n-1)})$. Precisamente, se u è soluzione di (4.7) e si pone

$$v=(v_1,\,v_2,\dots,\,v_n):=(u,u',\dots,u^{(n-1)}), \tag{4.8}$$

allora

$$\begin{cases} v_0'=v_1,\\ v_1'=v_2,\\ \dots,\\ v_n'=-a_0(t)v_0-a_1(t)v_1+\dots-a_{n-1}(t)v_{n-1}+b(t)\end{cases}$$

o in forma vettoriale

$$v' = \mathbf{A}(t)v + f(t) \tag{4.9}$$

dove

$$\mathbf{A}(t) := \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & & & & \\ 0 & 0 & 0 & \dots & 1 \\ -a_0(t) & -a_1(t) & -a_2(t) & \dots & -a_{n-1}(t) \end{pmatrix} \qquad \text{e} \qquad f(t) := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ b(t) \end{pmatrix}, \tag{4.10}$$

e viceversa: se $v : I \to \mathbb{R}^n$ è soluzione di (4.9) con $\mathbf{A}$ e f date da (4.10), allora $u := v_0$ è soluzione di (4.7) e $v = (u, u', \dots, u^{(n-1)})$.

Possiamo quindi applicare la teoria svolta per i sistemi lineari del primo ordine e rappresentare le soluzioni dell'equazione di ordine n: *se* $\mathbf{W}(t,s)$ *è la matrice di transizione del sistema* (4.9), *allora le soluzioni dell'equazione omogenea* (4.7) *sono date da*

$$u(t) = \sum_{j=1}^{n} \mathbf{W}_{1j}(t,t_0)c_j + \int_{t_0}^{t} \mathbf{W}_{1n}(t,\tau)f(\tau)\,d\tau, \qquad c = (c_1, \dots, c_n) \in \mathbb{R}^n. \tag{4.11}$$

Si osservi che se $u_1,\, u_2, \dots,\, u_n$ sono n soluzioni dell'equazione omogenea

$$u^{(n)} + a_{n-1}(t)u^{(n-1)} + \dots + a_1(t)u' + a_0(t)u = 0 \tag{4.12}$$

associata a (4.7), le colonne della matrice $n \times n$

$$\mathbf{Z}(t) := \begin{pmatrix} u_1 & u_2 & \dots & u_n \\ u_1' & u_2' & \dots & u_n' \\ u_1'' & u_2'' & \dots & u_n'' \\ \dots & & & \\ u_1^{(n-1)} & u_2^{(n-1)} & \dots & u_n^{(n-1)} \end{pmatrix} \tag{4.13}$$

sono soluzioni del sistema $v' = \mathbf{A}(t)v$. Diciamo che $\{u_1,\, u_2, \dots,\, u_n\}$ sono una base di soluzioni di (4.12) se $\det \mathbf{Z}(t) \neq 0$. Si chiama *Wronskiano* del sistema di soluzioni $\{u_1,\, u_2, \dots,\, u_n\}$ la funzione

$$w(t) := \det \mathbf{Z}(t).$$

Si ha $w(t) \neq 0\ \forall t$ e $\mathbf{W}(t,s) = \mathbf{Z}(t)\mathbf{Z}(s)^{-1}$.

4.13 Esercizio. Siano $u_1,\, u_2, \dots,\, u_n$ n soluzioni dell'equazione omogenea (4.12) e $\mathbf{Z}(t)$ come in (4.13). Mostrare che $w'(t) = -a_{n-1}(t)w$ e quindi

$$w(t) = w(t_0)\exp\Big(- \int_{t_0}^{t} a_{n-1}(\tau)\,d\tau \Big).$$

In particolare o $w(t) \neq 0\ \forall t \in I$ oppure $w(t) = 0\ \forall t \in I$.

4.e Equazioni di ordine superiore a coefficienti costanti

Quando $\mathbf{A}(t) = \mathbf{A}$ è costante, si può calcolare una base di soluzioni dell'equazione omogenea

$$u^{(n)} + a_{n-1}u^{(n-1)} + \cdots + a_1 u' + a_0 u = 0 \tag{4.14}$$

in termini degli zeri del *polinomio caratteristico*

$$p(\lambda) := \det(\mathbf{A} - \lambda\,\mathrm{Id}) = \sum_{k=0}^{n} a_k \lambda^k.$$

Conviene però cercare soluzioni a valori complessi di (4.14). Si verifica allora che se $p(\lambda)$ ha m radici complesse distinte $\lambda_1, \lambda_2, \dots, \lambda_k$ con molteplicità $r_1, r_2, \dots, r_k$, i.e., $p(\lambda) = \prod_{k=1}^{m}(\lambda - \lambda_k)^{r_k}$, *le funzioni* $t^h e^{\lambda_k t}$, $1 \le k \le m$, $1 \le h \le r_k$, sono una base (su $\mathbb{C}$) dello spazio delle soluzioni di (4.14).

4.14 Esercizio. Provare direttamente l'affermazione precedente.

Passiamo ora all'equazione completa

$$u^{(n)} + a_{n-1}u^{(n-1)} + \cdots + a_1 u' + a_0 u = b(t) \tag{4.15}$$

Basterà trovare una soluzione particolare. Si osservi che l'ultima colonna di $\mathbf{W}(t, 0) = e^{t\mathbf{A}}$ è la soluzione di $v' = \mathbf{A}v$ con dato iniziale $v(0) = (0, 0, \dots, 0, 1)^T$ e la prima componente $k(t) := v_0(t)$ di v, i.e., $k(t) = W_{1n}(t)$ è la soluzione dell'equazione omogenea (4.14) con dati iniziali

$$u(0) = 0,\ u'(0) = 0,\ \dots,\ u^{(n-1)}(0) = 1.$$

Essendo $W(t, s) = W(t-s, 0) = e^{(t-s)\mathbf{A}}$, segue dall (4.11) la *formula di Duhamel*: la funzione

$$u(t) = \int_0^t k(t-\tau)b(\tau)\,d\tau$$

è la soluzione particolare di (4.15) con dati iniziali

$$u(0) = 0,\ u'(0) = 0,\ \dots,\ u^{(n-1)}(0) = 0.$$

4.f Il calcolo di $e^{t\mathbf{A}}$

Consideriamo il sistema a coefficienti costanti $x' = \mathbf{A}x$. Con un cambio di variabile nelle incognite $y = \mathbf{P}x$, il sistema si trasforma in $y' = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}y$, i.e., in un sistema lineare per una matrice $\mathbf{B} := \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ simile ad $\mathbf{A}$. Una volta trovate le soluzioni $y(t)$ di questo sistema, è possibile calcolare le soluzioni del sistema originale calcolando $x(t) = \mathbf{P}^{-1}y(t)$.

Ad esempio, se esiste una base $\{u_1, u_2, \dots, u_n\}$ di $\mathbb{R}^n$ di autovettori di $\mathbf{A} \in M_{n,n}(\mathbb{C})$ con autovalori relativi $\lambda_1, \lambda_2, \dots, \lambda_n$, allora si ha

$$\mathrm{diag}\,(\lambda_1, \dots, \lambda_n) = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

con $\mathbf{P} = \Big[u_1 | u_2 | \dots | u_n\Big]$. Pertanto

$$e^{t\mathbf{A}} = \exp\Big(t\mathbf{P}\mathrm{diag}\,(\lambda_1,\,\lambda_2,\ldots,\,\lambda_n)\mathbf{P}^{-1}\Big) = \mathbf{P}\exp\Big(t\mathrm{diag}\,(\lambda_1,\,\lambda_2,\ldots,\,\lambda_n)\Big)\mathbf{P}^{-1}$$

e, essendo

$$\exp\Big(t\mathrm{diag}\,(\lambda_1,\,\lambda_2,\ldots,\,\lambda_n)\Big) = \mathrm{diag}\,(e^{\lambda_1 t}, e^{\lambda_2 t},\ldots,\,e^{\lambda_n t}),$$

si conclude che

$$e^{t\mathbf{A}} = \mathbf{P}\,\mathrm{diag}\,(\lambda_1,\,\lambda_2,\ldots,\,\lambda_n)\mathbf{P}^{-1} = \Big[u_1 e^{\lambda_1 t}|u_2 e^{\lambda_2 t}|\ldots|u_n e^{\lambda_n t}\Big]\mathbf{P}^{-1}.$$

4.15 Esercizio. Si può provare lo stesso risultato osservando che, se u è un autovettore per $\mathbf{A}$ con autovalore corrispondente λ, allora $x(t) = e^{\lambda t}u$ è la soluzione del sistema $x' = \mathbf{A}x$ con dato iniziale $x(0) = u$. Perciò se $\{u_1,\, u_2,\ldots,\, u_n\}$ è una base di $\mathbb{C}^n$ di autovettori di $\mathbf{A}$ con autovalori relativi $\lambda_1,\, \lambda_2,\ldots,\, \lambda_n$ la matrice

$$\mathbf{Z}(t) = \Big[e^{\lambda_1 t}u_1|e^{\lambda_2 t}u_2|\ldots|e^{\lambda_n t}u_n\Big]$$

è una matrice fondamentale per $x' = \mathbf{A}x$ con $\mathbf{Z}(0) = \Big[u_1|u_2|\ldots|u_n\Big]$. Perciò

$$e^{t\mathbf{A}} = \mathbf{W}(t,0) = \mathbf{Z}(t)\mathbf{Z}(0)^{-1} = \mathbf{Z}(t)\mathbf{P}^{-1}.$$

In generale, è possibile trovare una cambiamento di base $\mathbf{P}$ in $\mathbb{C}^n$ in modo che $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ sia in forma canonica di Jordan.

Siano $\lambda_1,\, \lambda_2,\ldots,\, \lambda_k$ gli autovalori distinti di $\mathbf{A}$ e $m_1,\, m_2,\ldots,\, m_k$ le relative molteplicità. Per ogni k sia p_k la dimensione dell'autospazio relativo a λ_k. Allora esiste un cambiamento di base $\mathbf{P}$ tale che $\mathbf{J} := \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ abbia la forma

$$\mathbf{J} = \begin{pmatrix} \boxed{\mathbf{J}_{1,1}} & 0 & 0 & \ldots & 0 \\ 0 & \boxed{\mathbf{J}_{1,2}} & 0 & \ldots & 0 \\ \ldots & & & & \\ 0 & 0 & 0 & \ldots & \boxed{\mathbf{J}_{k,p_k}} \end{pmatrix}$$

dove $i = 1,\ldots,k$, $j = 1,\ldots,p_i$ e

$$\mathbf{J}_{i,j} = \begin{cases} \lambda_i & \text{se } \mathbf{J}_{i,j} \text{ ha dimensione 1,} \\ \begin{pmatrix} \lambda_i & 1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_i & 1 & 0 & \ldots & 0 \\ \ldots & & & & & \\ \ldots & & & & & \\ 0 & 0 & 0 & \ldots & \lambda_i & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda_i \end{pmatrix} & \text{altrimenti.} \end{cases}$$

Se $\mathbf{J}' = \mathbf{J}_{i,j} = (\lambda)$ ha dimensione 1, allora evidentemente $e^{t\mathbf{J}'} = e^{\lambda t}$. Se invece $\mathbf{J}' = \mathbf{J}_{i,j}$ è uno dei blocchi di dimensione ℓ maggiore o uguale a 2,

$$\mathbf{J}' = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \dots \\ 0 & \dots & 0 & \lambda & 1 \\ 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

allora

$$\mathbf{J}' = \lambda\,\mathrm{Id} + \mathbf{N}, \qquad \mathbf{N}_{ij} = \delta_{i+1,j}.$$

Poiché $\mathbf{N}$ e Id commutano tra loro,

$$e^{t\mathbf{J}'} = e^{t\lambda\,\mathrm{Id}}\, e^{t\mathbf{N}} = e^{\lambda t}\, e^{t\mathbf{N}}$$

e

$$(\mathbf{N}^k)_{ij} = \begin{cases} \delta_{i+k,j} & \text{se } k < \ell, \\ 0 & \text{se } k \geq \ell, \end{cases}$$

si ha

$$\exp{(t\mathbf{J}')} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{k-1}}{(k-1)!} \\ 0 & 1 & t & \dots & \frac{t^{k-2}}{(k-2)!} \\ \dots \\ 0 & 0 & \dots & 1 & t \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

Perciò

$$\exp{(t\mathbf{A})} = \mathbf{P} e^{t\mathbf{J}} \mathbf{P}^{-1}$$

con

$$e^{t\mathbf{J}} := \begin{pmatrix} \boxed{e^{t\mathbf{J}_{1,1}}} & 0 & \dots & 0 \\ 0 & \boxed{e^{t\mathbf{J}_{1,2}}} & \dots & 0 \\ \dots \\ 0 & 0 & \dots & \boxed{e^{tJ_{k,p_k}}} \end{pmatrix}.$$

Si osservi che ogni elemento della matrice $e^{t\mathbf{A}}$ ha la forma

$$\sum_{j=1}^{k} p_j(t) \exp{(\lambda_j t)}$$

dove $p_j(t)$ è un polinomio di grado al più $p_j - 1$ e $\lambda_1,\ \lambda_2, \dots,\ \lambda_n$ sono gli autovalori di $\mathbf{A}$. Segue immediatamente che per ogni $\rho > \max(\Re\lambda_1, \Re\lambda_2, \dots, \Re\lambda_n)$ esiste una costante C_ρ tale che

$$\Big|\exp\left(t\mathbf{A}\right)\Big| \le C_\rho e^{t\rho}, \qquad 0 \le t < \infty.$$

In particolare

4.16 Proposizione. *Se tutti gli autovalori di* $\mathbf{A}$ *hanno parte reale negativa, allora ogni soluzione di* $x' = \mathbf{A}x$ *tende a zero per* $x \to +\infty$*. Precisamente, se* $\sigma > 0$ *è tale che* $\max_{i=1,n}(\Re\lambda_i) < -\sigma < 0$*, allora esiste una costante* C_σ *tale che*

$$|x(t)| \le C_\sigma e^{-\sigma t}, \qquad t \ge 0.$$

4.17 Esercizio. Ritrovare il risultato alla fine della sezione precedente sulle equazioni di ordine n.

4.18 Esercizio. Mostrare che, se $x(t)$ è soluzione di $x' = \mathbf{A}x + f$, $\mathbf{A}$ ha autovalori tutti con parte reale negativa e f non cresce più che esponenzialmente, allora x non cresce più che esponenzialmente all'infinito.

5. Indice analitico