UNIVERSITÀ DEGLI STUDI DELLA CALABRIA
Facoltà di Scienze Matematiche, Fisiche e Naturali

Corso di Laurea in Matematica

Metodo di Newton-Kantorovich per operatori non lineari tra spazi di Banach

RELATORE
prof. Espedito De Pascale

CANDIDATA
Infusino Maria
Matr. 78251

Anno Accademico 2004-05

"Nature and Nature's laws lays hid in night

God said, Let Newton be! And all was light."

(Alexander Pope, 1727).

Indice

Introduzione

Nella formulazione in termini matematici di molti problemi derivanti dalle scienze applicate, il modello matematico conduce alla risoluzione di equazioni non lineari. Le difficoltà della ricerca degli zeri di un operatore non lineare dipendono dalla natura dell'operatore. Infatti se l'equazione è algebrica di grado superiore al quarto o è trascendente, non esistono formule che conducano ad una espressione esplicita di ciascuna radice. Inoltre per equazioni algebriche di terzo e quarto grado, pur esistendo formule risolutive per radicali, la loro non semplice struttura le rende poco usate. Di conseguenza per poter risolvere equazioni di questo tipo, l'idea fondamentale cui si ricorre è usare metodi iterativi per costruire approssimazioni successive della soluzione. I metodi di tipo iterativo sono procedure numeriche che, partendo da una approssimazione iniziale della radice, generano una sequenza di approssimazioni convergente, sotto opportune ipotesi, alla radice cercata.

Diversi sono i metodi iterativi finalizzati a risolvere equazioni non lineari: bisezione, regula falsi, metodo di Newton, metodo delle secanti, ecc.

Oggetto del presente lavoro è il metodo di Newton, in particolare ci occuperemo della sua estensione agli operatori non lineari tra spazi di Banach, dovuta a Kantorovich (1912-1986), e del teorema di convergenza da lui stesso stabilito.
L'esposizione del risultato di Kantorovich richiede, tuttavia, conoscenze di analisi funzionale lineare e non lineare che non sono previste negli insegnamenti del corso di laurea triennale. Pertanto nei primi tre capitoli abbiamo riportato le nozioni fondamentali degli spazi di Banach, degli operatori lineari e del calcolo differenziale in questi spazi.
In particolare, nel primo capitolo introduciamo gli spazi normati, richiamando alcuni risultati sulle norme, dei quali ci siamo serviti per definire la convergenza di serie e successioni in questi spazi.
Quindi abbiamo rivolto la nostra attenzione allo studio degli spazi di Banach, perché è proprio a questa struttura che Kantorovich estende il metodo di Newton. Infatti è di grande utilità operare proprio all'interno degli spazi di Banach, poiché in essi è garantita la completezza, che è una proprietà spesso indispensabile per la validità di molti risultati. Nel passaggio dalla dimensione finita alla dimensione infinita vengono perse molte delle proprietà che valgono nell'usuale spazio euclideo. Questo rende ancora più interessante il risultato di Kantorovich, che si può enunciare prescindendo dalla dimensione dello spazio.
Il secondo capitolo verte sul concetto di operatore lineare ed in particolare di operatore lineare continuo. Nella sezione 2.4 esponiamo alcuni risultati classici di invertibilità

per operatori lineari, tra i quali il lemma di Banach. Quest'ultimo è un risultato di grande rilievo, di cui si serve Kantorovich nella dimostrazione del suo teorema di convergenza. Tuttavia prima di enunciare il teorema, abbiamo ritenuto opportuno dare un breve cenno alle definizioni di derivata secondo Frèchet e secondo Gateâux. Perciò nel capitolo terzo abbiamo esteso i concetti classici della differenziabilità in $\mathbb{R}^n$ agli spazi di Banach di dimensione infinita.

Infine nel capitolo quarto giungeremo al nucleo principale di questa tesina: il metodo di Newton-Kantorovich. Inizialmente diamo alcune nozioni basilari sui metodi iterativi, che devono essenzialmente la loro origine proprio al problema della ricerca degli zeri di un operatore non lineare. Dunque nella sezione 4.3 proponiamo il metodo di Newton nel caso più elementare, cioè quello di una funzione reale, con la nota interpretazione geometrica, dovuta a Mourraille e risalente al 1768. Prima, però, di trattare da vicino l'estensione del metodo stabilita da Kantorovich e il suo teorema di convergenza del 1948, abbiamo preferito dedicare un'intera sezione al principio di maggiorazione che il matematico russo introduce nel 1951. Il principio di maggiorazione ci dà condizioni sufficienti affinché un operatore ammetta un punto fisso e perciò Kantorovch lo utilizza per perfezionare e rendere più semplice la dimostrazione del risultato di convergenza. Il teorema di Kantorovich fornisce condizioni sufficienti per l'esistenza e l'unicità di soluzioni di equazioni non lineari negli spazi di Banach e mostra che sotto alcune ipotesi la sequenza generata dal metodo di Newton e da quello di Newton modificato convergono alla soluzione esatta. L'ipotesi della limitatezza della derivata seconda dell'operatore considerato, fu sostituita nel 1955 da

Bartle con l'ipotesi della lipschitzianità della derivata prima. Diverse estensioni del teorema di Kantorovich furono usate per ottenere stime dell'errore per il metodo di Newton e per i metodi che ne derivano, come i metodi Newton-like, il metodo delle secanti e il metodo di Newton modificato. Molti sono comunque gli argomenti correlati al teorema di Kantorovich, che ancora oggi attraggono l'interesse dei matematici. Il metodo di Newton-Kantorovich, infatti, nonostante la semplicità del suo principio di base è applicabile a vari tipi di problemi come i sistemi di equazioni non lineari, i problemi di autovalori, equazioni differenziali e integrali.

Capitolo 1

Spazi normati e Spazi di Banach

1.1 Introduzione

Lo scenario entro il quale Kantorovich elabora e sviluppa la sua estensione del metodo di Newton agli operatori non lineari è la struttura degli spazi di Banach. Dalla combinazione naturale dei concetti di spazio vettoriale e spazio metrico si origina la struttura di spazio normato e quindi di spazio di Banach, cioè di spazio normato completo. La definizione astratta di spazio normato compare per la prima volta intorno agli anni venti principalmente nelle opere di Stefan Banach (1894-1964) ma anche nei lavori di Hans Hahn (1879-1934) e Norbert Wiener (1894-1964). Proprio in questo periodo, infatti, Banach getta le fondamenta di quella che oggi viene denominata analisi funzionale lineare. In questo capitolo ci limiteremo a presentarne le definizioni basilari e a illustrare le conseguenze più dirette che ne derivano. Partendo dalla definizione di norma, introduciamo le nozioni ad essa collegate che saranno poi utilizzate per definire i concetti fondamentali di convergenza di una successione e completezza di uno spazio. Quindi dopo aver caratterizzato gli spazi di Banach tramite le serie assolutamente convergenti, porremo la nostra attenzione agli spazi normati di dimensione finita. Inoltre per stabilire la validità di molti risultati, anche in dimensione infinita, si rende il più delle volte indispensabile l'ipotesi della completezza ed è da questa esigenza che si comprende la convenienza di considerare gli spazi di Banach. L'estensione agli spazi di Banach di molte proprietà, note nei casi elementari, è una delle pietre miliari della matematica del nostro secolo, giacchè è notevole il contributo che essa offre nel campo della risoluzione di vaste classi di problemi.

1.2 Norme e spazi normati

Definizione 1.2.1 *Sia X uno spazio vettoriale e sia $\|\cdot\| : X \to \mathbb{R}^+$. La funzione reale $\|\cdot\|$ è una norma se valgono le seguenti proprietà:*

- $\|x\| > 0$, $\forall x \in X$ *e* $x \neq 0$ *e* $\|0\| = 0$ *positività*
- $\|\lambda x\| = |\lambda| \, \|x\|$, $\forall x \in X$ $\forall \lambda \in \mathbb{R}$ *omogeneità*
- $\|x + y\| \leq \|x\| + \|y\|$, $\forall x, y \in X$ *disuguaglianza triangolare*

Osservazione 1.2.2 *La norma è una funzione continua e lipschitziana.*
(si rimanda a [5])

Osservazione 1.2.3 *Dalle proprietà della norma segue facilmente la seguente proprietà:*

$$\left|\|x\| - \|y\|\right| \leq \|x - y\| \ , \forall x, y \in X \tag{1.1}$$

Definizione 1.2.4 *Uno spazio X su cui sia definita una norma $\|\cdot\|$ si dice spazio normato.*

Osservazione 1.2.5 *Uno spazio normato si indica con una coppia $(X, \|\cdot\|)$ costituita da uno spazio vettoriale e dalla norma definita in esso.*

Il concetto di norma è strettamente connesso a quello di prodotto scalare. Infatti se X è uno spazio vettoriale reale con prodotto scalare $x \cdot y$, allora X è uno spazio normato con la norma $\|x\| = \sqrt{x \cdot x}$; analogamente se X è uno spazio vettoriale complesso un prodotto hermitiano su esso definisce una norma. Tuttavia non sempre vale il viceversa, cioè in generale un prodotto scalare non deriva da una norma, ma affinché questo accada è necessario imporre alcune condizioni. Al riguardo si ricordi l'identità del parallelogramma (si veda [5]), che è condizione necessaria e sufficiente affinché esista un prodotto scalare indotto dalla norma.

Osserviamo che uno spazio normato $(X, \|\cdot\|)$ è uno spazio metrico con la distanza naturale $d : X \to \mathbb{R}$ definita da:

$$d(x, y) \doteq \|x - y\|, \qquad \forall x, y \in X$$

che gode di tutte le proprietà di distanza ed inoltre è invariante per traslazioni, cioè

$$d(x + z, y + z) = d(x, y)$$

Ad ogni norma si può dunque associare una distanza che dipende dalla norma fissata. Inoltre è evidente che per uno spazio normato si possono dare tutte le nozioni topologiche di aperto, chiuso e intorno.

Di particolare importanza, specialmente nell'ambito degli spazi normati di dimensione finita, risulta il concetto di norme equivalenti.

Definizione 1.2.6 *Due norme ρ_1 e ρ_2 su uno stesso spazio vettoriale X si dicono equivalenti se le rispettive distanze inducono la stessa topologia su X, cioè generano sullo spazio la stessa famiglia di aperti.*

Teorema 1.2.7 *ρ_1 e ρ_2 sono due norme equivalenti sullo spazio X se e solo se esistono due costanti m, M con $0 < m < M$ tali che:*

$$m\rho_1(x) \leq \rho_2(x) \leq M\rho_1(x) \quad , \forall x \in X \tag{1.2}$$

Dimostrazione. Siano ρ_1 e ρ_2 equivalenti, allora l'applicazione identità

$$id : (X, \rho_1\ ) \to (X, \rho_2)$$

è continua e lo è in particolare nel punto 0, ciò significa che:

$$\forall\ 0 < \varepsilon < 1\ \ \exists\ \delta > 0 : \rho_2(id(z) - id(0)) = \rho_2(z)\ \leq \varepsilon\ \ \forall\ z \in X \text{ con } \rho_1\ (z) \leq \delta \tag{1.3}$$

cioè $\lim\limits_{z\to\infty} id(z) = id(0) = 0$.

Prendiamo ora $x \in \mathbb{R}$ e $z := \frac{\delta x}{\rho_2(x)}$ allora $\rho_2(z) = \rho_2\left(\frac{\delta x}{\rho_2(x)}\right) = \frac{\delta}{\rho_2(x)}\rho_2(x) = \delta$ quindi dalla (1.3) si ha che $\rho_1\ (z) \leq \delta = \rho_2(z) \leq 1$.

E ancora $\forall x \in X \quad \rho_1\ \left(\frac{\delta x}{\rho_2(x)}\right) \leq 1 \iff \rho_1\ (x) \leq \frac{1}{\delta}\rho_2(x)$.

Scambiando i ruoli a ρ_1 e ρ_2 si ottiene analogamente che $\rho_2\ (x) \leq \frac{1}{\delta_1}\rho_1(x)$ e quindi

$$\delta\rho_1\ (x) \leq \rho_2\ (x) \leq \frac{1}{\delta_1}\rho_1(x) \quad \forall x \in X \tag{1.4}$$

infine ponendo $m = \delta$ e $M = \frac{1}{\delta_1}$ dalla (1.4) si ha la (1.2)

Viceversa sia vera la (1.2). Vogliamo verificare che ogni aperto in (X, ρ_1) è aperto anche in (X, ρ_2) e viceversa, cioè vogliamo dimostrare che ρ_1 e ρ_2 inducono su X la stessa topologia.

Considerando la palla aperta di centro x_0 e raggio r_1 in (X, ρ_1) e la palla aperta di centro x_0 e raggio r_2 in (X, ρ_2), che indicheremo con:

$$B_1(x_0, r_1) = \{x \in X : \rho_1(x - x_0) < r_1\}$$

$$B_2(x_0, r_2) = \{x \in X : \rho_2(x - x_0) < r_2\}$$

si ha che:

$$B_1(x_0, r_1) \subseteq B_2(x_0, Mr_1) \quad . \tag{1.5}$$

Infatti $\forall x \in B_1(x_0, r_1) \quad \rho_1(x - x_0) < r_1$ allora per la (1.2) :

$$\rho_2(x - x_0) \leq M\rho_1(x - x_0) < Mr_1$$

cioè $x \in B_2(x_0, Mr_1)$.

E si ha pure che:

$$B_2(x_0, r_2) \subseteq B_1\left(x_0, \frac{r_2}{m}\right) \tag{1.6}$$

infatti $\forall x \in B_2(x_0, r_2) \quad \rho_2(x - x_0) < r_2$ allora per la (1.2) :

$$m\rho_1(x - x_0) \leq \rho_2(x - x_0) < r_2\,,$$

che equivale a dire che $\rho_1(x - x_0) < \frac{r_2}{m}$, cioè $x \in B_1\left(x_0, \frac{r_2}{m}\right)$.

Dunque se A è un aperto di (X, ρ_1) allora per definizione:

$\forall \varepsilon > 0 \; \forall x_0 \in A \;\; \exists B_1(x_0, \varepsilon) \subset A$ quindi per la (1.6)

$$\forall x_0 \in A \; \exists B_2(x_0, m\varepsilon) \subset B_1(x_0, \varepsilon) \subset A$$

cioè A è aperto in $(X, \rho_2\ )$.

Viceversa se A è un aperto di $(X, \rho_2\ )$ allora per definizione:

$\forall \mu > 0\ \forall x_0 \in A\ \ \exists\ B_2\left(x_{0,}\mu\right) \subset A$ quindi per la (1.5)

$$\forall x_0 \in A\ \exists B_1\left(x_{0,}\frac{\mu}{M}\right) \subset B_2\left(x_{0,}\mu\right) \subset A$$

cioè A è aperto in (X, ρ_1) ■

1.3 Dagli spazi normati agli spazi di Banach

Nella sezione precedente si è evidenziato come uno spazio normato possa sempre essere considerato uno spazio metrico con la distanza naturale indotta dalla norma. Dunque possiamo considerare in uno spazio normato il concetto di convergenza, che deve sempre essere considerata come relativa alla distanza indotta dalla norma.

Definizione 1.3.1 *Sia $(X, \|\cdot\|)$ uno spazio normato, una successione $(x_n)_{n\in\mathbb{N}}$ a valori in X converge a $x \in X$ se:*

$$\forall \varepsilon > 0\ \ \exists \nu \in \mathbb{N} : \|x_n - x\| \leq \varepsilon\ \ \forall n \geq \nu \tag{1.7}$$

Osservazione 1.3.2 *Come nel caso delle successioni reali la (1.7) si può esprimere utilizzando la nozione di limite cioè:*

$$\lim_{x\to\infty} \|x_n - x\| = 0$$

Osservazione 1.3.3 *Questa non è soltanto un'analogia, infatti tutti i risultati che nel caso delle successioni reali sono stati dimostrati utilizzando solo il concetto di limite* *e le proprietà del modulo valgono anche nel caso di uno spazio normato in generale. (si veda [6])*

Definizione 1.3.4 *Una successione* $(x_n)_{n\in\mathbb{N}}$ *a valori in* X *spazio normato è detta successione di Cauchy se:*

$$\forall\varepsilon > 0 \quad \exists\ \nu \in \mathbb{N} : \|x_n - x_m\| \leq \varepsilon \quad \forall n, m \geq \nu \tag{1.8}$$

Proposizione 1.3.5 *In uno spazio normato ogni sottosuccessione di una successione di Cauchy è di Cauchy.*

Proposizione 1.3.6 *In uno spazio normato se una sottosuccessione di una successione di Cauchy converge, la successione converge allo stesso limite.*

Proposizione 1.3.7 *In uno spazio normato una successione convergente è di Cauchy.*

La condizione di Cauchy (1.8) non cambia se si passa da una norma ad un'altra equivalente, ma in generale non è sufficiente alla convergenza. Si rende dunque necessario distinguere gli spazi normati in cui vale l'inversa della proposizione 1.3.7, ecco perché introduciamo la nozione di completezza.

Definizione 1.3.8 *Uno spazio normato* X *si dice completo se ogni successione di Cauchy in* X *è convergente.*

Definizione 1.3.9 *Uno spazio normato* X *completo rispetto alla distanza naturale si dice spazio di Banach.*

Esempi

1. $(\mathbb{R}, |\cdot|)$ è uno spazio di Banach, infatti $|\cdot|$ è una norma su $\mathbb{R}$ e lo spazio $\mathbb{R}$ è completo poiché si dimostra che ogni successione di Cauchy in $\mathbb{R}$ è convergente.

2. $(\mathbb{Q}, |\cdot|)$ non è uno spazio di Banach, poiché è uno spazio normato ma non è completo. Si pensi ad esempio alla successione $\left(1 + \frac{1}{n}\right)_{n\in\mathbb{N}}$ che è di Cauchy ma non converge in $\mathbb{Q}$.

3. $\mathbb{R}^n$ con la norma euclidea $\|x\| = \left(\sum_{k=1}^{n} |x_k|^2\right)^{\frac{1}{2}} \quad \Big(x = (x_1, ..., x_n)\Big)$ è di Banach.

4. $\mathbb{C}^n$ con la norma euclidea $\|z\| = \left(\sum_{k=1}^{n} |z_k|^2\right)^{\frac{1}{2}} \quad \Big(z = (z_1, ..., z_n)\Big)$ è di Banach.

Riprendiamo ora il teorema 1.2.7 che caratterizza le norme equivalenti, introdotto nella sezione precedente, e osserviamone due importanti conseguenze:

- ρ_1 e ρ_2 sono due norme equivalenti sullo spazio X se e solo se (X, ρ_1) e (X, ρ_2) hanno le stesse successioni convergenti.

- ρ_1 e ρ_2 sono due norme equivalenti sullo spazio X se e solo se (X, ρ_1) e (X, ρ_2) hanno le stesse successioni di Cauchy.

Quindi possiamo concludere la seguente:

Proposizione 1.3.10 *Se ρ_1 e ρ_2 sono due norme equivalenti sullo spazio X allora (X, ρ_1) è uno spazio di Banach se e soltanto se (X, ρ_2) è di Banach.*

La completezza che caratterizza gli spazi di Banach gioca un ruolo fondamentale in molti risultati, per cui è importante determinare criteri che assicurino la completezza di uno spazio normato. Il teorema 1.3.13 individua come condizione necessaria e sufficiente affinché uno spazio sia di Banach la convergenza assoluta delle sue serie. Prima di prenderlo in considerazione osserviamo che poiché in uno spazio vettoriale X le somme finite di elementi di X sono ben definite e appartengono ad X, data una successione $(x_n)_{n\in\mathbb{N}}$ a valori in X anche la serie $\sum_{k=0}^{\infty} x_k$ è ben definita ed ha dunque senso definirne la convergenza e l'assoluta convergenza.

Definizione 1.3.11 *Sia $(X, \|\cdot\|)$ uno spazio di Banach e sia $x_k \in X$, $\forall k \in \mathbb{N}$. La serie $\sum_{k=0}^{\infty} x_k$ si dice convergente se la successione delle somme parziali $s_n := \sum_{k=0}^{n} x_k$ converge in X, cioè se*

$$\exists\ s \in X \ \textit{tale che}\ \forall \varepsilon > 0 \ \ \exists\ \nu \in \mathbb{N} : \|s_n - s\| < \varepsilon \ \ \forall n > \nu \tag{1.9}$$

$s = \sum_{k=0}^{\infty} x_k$ viene detta somma della serie.

Definizione 1.3.12 *La serie $\sum_{k=0}^{\infty} x_k$ con $x_k \in (X, \|\cdot\|)$ si dice assolutamente convergente se la serie $\sum_{k=0}^{\infty} \|x_k\|$ è convergente.*

Teorema 1.3.13 *In uno spazio normato $(X, \|\cdot\|)$ ogni serie assolutamente convergente è convergente se e soltanto se X è uno spazio di Banach.*

Dimostrazione. Supponiamo che X sia uno spazio di Banach e sia $\sum_{k=0}^{\infty} x_k$ una serie assolutamente convergente di elementi di X.

Dall'assoluta convergenza segue che $\sum_{k=0}^{\infty} \|x_k\|$ è una serie convergente in $\mathbb{R}$, ciò equivale a dire che la successione delle sue somme parziali $t_n = \sum_{k=0}^{n} \|x_k\|$ converge in $\mathbb{R}$. Per il criterio di Cauchy per le successioni reali, $(t_n)_{n\in\mathbb{N}}$ è una successione di Cauchy, cioè:

$$\forall\varepsilon > 0 \;\; \exists\, \nu \in \mathbb{N} : |t_m - t_n| < \varepsilon \;\; \forall n, m \geq \nu \tag{1.10}$$

La (1.10) è equivalente a:

$$\forall\varepsilon > 0 \;\; \exists\, \nu \in \mathbb{N} : \|x_0\| + \|x_1\| + ... + \|x_m\| - \|x_0\| - ... - \|x_n\| = \sum_{k=n+1}^{m} \|x_k\| < \varepsilon \quad , \forall n, m \geq \nu$$

Allora per la disuguaglianza triangolare si ha:

$$\forall\varepsilon > 0 \;\; \exists\, \nu \in \mathbb{N} : \left\| \sum_{k=n+1}^{m} x_k \right\| \leq \sum_{k=n+1}^{m} \|x_k\| < \varepsilon \;\; , \forall n, m \geq \nu$$

cioè la successione $(s_n)_{n\in\mathbb{N}}$ delle somme parziali della serie $\sum_{k=0}^{\infty} x_k$ è una successione di Cauchy.

Essendo per ipotesi X di Banach, la successione $(s_n)_{n\in\mathbb{N}}$ è convergente e quindi la serie $\sum_{k=0}^{\infty} x_k$ converge.

Viceversa supponiamo che ogni serie assolutamente convergente converga in X e dimostriamo che X è di Banach.

Consideriamo una successione $(x_n)_{n\in\mathbb{N}}$ di Cauchy in X, poiché vale la proposizione 1.3.5 si può estrarre una sottosuccessione $(x_{n_k})_{k\in\mathbb{N}}$ tale che $\left\|x_{n_k} - x_{n_{k-1}}\right\| < 2^{-k}$ da cui $\sum_{k=1}^{\infty} \left\|x_{n_k} - x_{n_{k-1}}\right\| < \sum_{k=1}^{\infty} 2^{-k} < +\infty$.

Questo significa che la serie $\sum_{k=1}^{\infty} \left(x_{n_k} - x_{n_{k-1}}\right)$ è assolutamente convergente, quindi convergente in X per ipotesi. Supponiamo sia y la somma di questa serie, allora:

$$\lim_{p\to\infty} \left\| \sum_{k=1}^{p} \left(x_{n_k} - x_{n_{k-1}}\right) - y \right\| = 0 \iff \lim_{p\to\infty} \left\| x_{n_1} + x_{n_2} + .. + x_{n_p} - x_{n_0} - x_{n_1} - .. - x_{n_{p-1}} - y \right\| = 0$$

$$\Leftrightarrow \lim_{p\to\infty} \left\| x_{n_p} - x_{n_0} - y \right\| = 0$$

ponendo $x := y + x_{n_0}$ abbiamo che la sottosuccessione $\left(x_{n_p}\right)_{p\in\mathbb{N}}$ converge a x, allora per la proposizione 1.3.6 anche $(x_n)_{n\in\mathbb{N}}$ è convergente e converge a x. Dunque X è completo, cioè è uno spazio di Banach. ■

1.4 Spazi normati di dimensione finita

A partire da $\mathbb{R}^n$, utilizzando i risultati esposti sino ad ora, illustriamo alcune proprietà comuni a tutti gli spazi normati di dimensone finita:

Teorema 1.4.1 *In $\mathbb{R}^n$ tutte le norme sono equivalenti.*

Dimostrazione. Sia $\|\cdot\|$ la norma euclidea e ρ una qualunque altra norma in $\mathbb{R}^n$. Siano $\{e_1, .., e_n\}$ la base canonica di $\mathbb{R}^n$ e siano x, $y\in \mathbb{R}^n$. Allora $x = \sum_{i=1}^{n} x^i e_i$ e $y = \sum_{i=1}^{n} y^i e_i$. Per la proprietà (1.1) delle norme si ha:

$$\|\rho(x) - \rho(y)\| \leq \rho(x-y) = \rho\left(\sum_{i=1}^{n} x^i e_i - \sum_{i=1}^{n} y^i e_i\right) = \rho\left(\sum_{i=1}^{n}\left(x^i - y^i\right)e_i\right) \tag{1.11}$$

Utilizzando la disuguaglianza triangolare, la proprietà di omogeneità della norma e la disuguaglianza di Cauchy-Swartz otteniamo dalla (1.11):

$$\begin{aligned}
|\rho(x) - \rho(y)| &\leq \rho\left(\sum_{i=1}^{n}\left(x^i - y^i\right)e_i\right) \leq \sum_{i=1}^{n} \rho\left(\left(x^i - y^i\right)e_i\right) \\
&= \sum_{i=1}^{n}\left|x^i - y^i\right| \rho(e_i) \leq \left(\sum_{i=1}^{n}\left|x^i - y^i\right|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} (\rho(e_i))^2\right)^{\frac{1}{2}} \\
&= \|x-y\| \left(\sum_{i=1}^{n} (\rho(e_i))^2\right)^{\frac{1}{2}} = c\,\|x-y\|
\end{aligned}$$

dove c è una costante indipendente da x nè da y

Abbiamo quindi verificato che per la norma $\rho : (\mathbb{R}^n, \|\cdot\|) \to (\mathbb{R}^+, |\cdot|)$ esiste una costante c tale che:

$$\|\rho(x) - \rho(y)\| \leq c \|x - y\| \quad ,\forall x, y \in \mathbb{R}^n. \tag{1.12}$$

Consideriamo ora la sfera di raggio unitario in $\mathbb{R}^n$ che indicheremo con:

$$S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$$

essa è un compatto di $\mathbb{R}^n$.

Per la (1.12) ρ è una funzione lipschitziana, quindi continua su $\mathbb{R}^n$, dunque ρ sarà continua e mai nulla su S^{n-1}. Allora per il teorema di Weirstrass, la funzione ρ ammette un massimo M e un minimo m sulla sfera (con $m, M > 0$). Dall'omogeneità delle norme segue che:

$$m \leq \rho(x) \leq M \ , \forall x \in S^{n-1} \iff m \|x\| \leq \rho(x) \leq M \ \|x\| , \forall x \in S^{n-1}. \tag{1.13}$$

Per ogni $y \in \mathbb{R}^n$possiamo sempre trovare $\alpha \in \mathbb{R} : \alpha y = x$ con $\|x\| = 1$; dunque la (1.13) diventa:

$$m \|\alpha y\| \leq \rho(\alpha y) \leq M \ \|\alpha y\| , \ \forall y \in \mathbb{R}^n \quad \Leftrightarrow$$

$$\iff m |\alpha| \|y\| \leq |\alpha| \rho(y) \leq M \ |\alpha| \|y\| , \ \forall y \in \mathbb{R}^n \quad \Leftrightarrow$$

$$\Leftrightarrow m \|y\| \leq \rho(y) \leq M \ \|y\| , \ \forall y \in \mathbb{R}^n$$

dunque per il teorema 1.2.7, la norma euclidea e la norma ρ sono equivalenti. ■

Corollario 1.4.2 *Ogni spazio normato di dimensione finita è uno spazio di Banach.*

Dimostrazione. Essendo in $\mathbb{R}^n$ tutte le norme equivalenti ed essendo $(\mathbb{R}^n, \|\cdot\|)$ uno spazio di Banach ($\|\cdot\|$ è la norma euclidea), allora per la proposizione 1.3.10 si ha che anche $(\mathbb{R}^n, \rho)$ è uno spazio di Banach. Sia (X, ρ) un qualunque spazio normato di dimensione finita n e consideriamo il sistema di coordinate:

$$\xi : \mathbb{R}^n \to X$$
$$\xi : \left(x^1, x^2, ..., x^n\right) = x \to \xi(x) = \sum_{i=1}^{n} x^i e_i$$

La ξ è un isomorfismo tra $\mathbb{R}^n$ e X quindi dire che $(\mathbb{R}^n, \rho)$ è uno spazio di Banach equivale a dire che anche (X, ρ) lo è. ■

Osservazione 1.4.3 *Si può dunque estendere il teorema 1.4.1:*

"In uno spazio normato di dimensione finita tutte le norme sono equivalenti"

Corollario 1.4.4 *Sia X uno spazio normato. Ogni sottospazio $Y \subset X$ di dimensione finita è un chiuso.*

Dimostrazione. Sia $(X, \|\cdot\|)$ e sia $\{\omega_{1,}..., \omega_n\}$ una base dello spazio Y, che supponiamo abbia dimensione n. Definiamo su Y la norma:

$$\rho(v) = \sum_{i=1}^{n} \left|v^i\right| \quad \text{con} \quad \sum_{i=1}^{n} v^i \omega_i = v \in Y$$

(Y, ρ) è uno spazio normato di dimensione finita, dunque per l'osservazione 1.4.3, tutte le norme su Y sono equivalenti.

Essendo $\|\cdot\|$ una norma anche su Y, poiché $Y \subset X$ allora si ha:

$$\exists\, c \in \mathbb{R} : \rho(v) \le c \|v\| \ , \forall v \in Y$$

Consideriamo ora $(y_k)_{k\in\mathbb{N}} \subset Y$ con $y_k \to y \in X$ per $k \to \infty$, quindi per dimostrare che Y è un chiuso, dobbiamo dimostrare che $y \in Y$.

Per ogni y_k vale:

$$\rho(y_k) = \sum_{i=1}^{n} \left|y_k^i\right| \leq c \|y_k\| \implies \left|y_k^i\right| \leq c \|y_k\|$$

Allora avremo che:

$$\left|y_k^i - y_j^i\right| \leq c \|y_k - y_j\| = c \|y_k - y + y - y_j\| \leq c \|y_k - y\| + c \|y - y_j\| \tag{1.14}$$

poiché $(y_k)_{k\in\mathbb{N}}$ è convergente a $y \in X$ si ha:

$$\forall \varepsilon > 0 \;\; \exists \nu \in \mathbb{N} : \|y_k - y\| \leq \varepsilon \quad \forall k \geq \nu \tag{1.15}$$

Dalla (1.14) e dalla (1.15) otteniamo:

$$\forall \varepsilon > 0 \;\; \exists \nu \in \mathbb{N} : \left|y_k^i - y_j^i\right| \leq (2c)\, \varepsilon \quad \forall k, j \geq \nu$$

cioè $\left(y_k^i\right)_{k\in\mathbb{N}}$ è una successione reale di Cauchy, dunque è convergente. Sia $y_k^i \to \lambda^i \in \mathbb{R}$, allora:

$$y_k = \sum_{i=1}^{n} y_k^i \omega_i \to \sum_{i=1}^{n} \lambda^i \omega_i := q$$

e per l'unicità del limite $y = q$ ■

Una proprietà caratteristica degli spazi normati di dimensione finita è che la palla unitaria è compatta. Alla dimostrazione di questo importante risultato giunse Frigyes Riesz (1880-1956). Prima di esporre il teorema che ne porta il nome introduciamo il seguente lemma :

Lemma 1.4.5 ***(di Riesz)*** *Sia $(X, \|\cdot\|)$ uno spazio normato e sia $Y \subset X$ un sottospazio lineare chiuso. Allora:*

$$\exists \overline{x} \in X : \|\overline{x}\| = 1 \;\; e \;\; \|\overline{x} - x\| \geq \frac{1}{2} \quad \forall x \in Y.$$

Dimostrazione. Sia $x_0 \in X \backslash Y$ e sia $d := \inf\limits_{y \in Y} \{\|y - x_0\|\}$. Proviamo che $d > 0$. Infatti se fosse $d = 0$ esisterebbe una successione $(y_n)_{n \in \mathbb{N}} \subset Y$ con $y_n \to x_0$ per $n \to \infty$. Essendo Y chiuso ogni successione convergente a valori in Y, converge a un punto di Y, dunque risulterebbe $x_0 \in Y$ contro l'ipotesi iniziale.

Sia ora $y_0 \in Y$con $\|y_0 - x_0\| \leq 2d$. Consideriamo $\overline{x} = \frac{x_0 - y_0}{\|x_0 - y_0\|}$, risulta che : $\|\overline{x}\| = 1$. Se $y \in Y$, allora $y_0 + y\,\|x_0 - y_0\| =: z \in Y$ poiché Y è un sottospazio lineare di X. In definitiva abbiamo che:

$$\|y - \overline{x}\| = \left\| y - \frac{x_0 - y_0}{\|x_0 - y_0\|} \right\| = \left\| \frac{y\,\|x_0 - y_0\| - x_0 + y_0}{\|x_0 - y_0\|} \right\| = \frac{\|z - x_0\|}{\|x_0 - y_0\|} \geq \frac{d}{2d} = \frac{1}{2}$$

■

Teorema 1.4.6 ***(di Riesz)*** *La palla unitaria di uno spazio normato X è compatta se e solo se X è di dimensione finita.*

Dimostrazione. Sia X di dimensione finita con la norma $\|\cdot\|$ e sia

$$B = \{x \in X : \|x\| \leq 1\} \subset X$$

Sia $\xi : K^n \to X$ un sistema di coordinate. Essendo ξ un isomorfismo lineare, è anche un omeomorfismo. Consideriamo $\xi^{-1}(B) \subset K^n$. Per il teorema 1.4.1 $\xi^{-1}(B)$ è chiuso e limitato, dunque è un compatto di K^n e quindi $B = \xi\left(\xi^{-1}(B)\right)$ è un compatto di X, essendo ξ omeomorfismo.

Per dimostrare il viceversa, verifichiamo che se X ha dimensione infinita, allora la sfera unitaria di X non è compatta. Sia x_1 un vettore sulla sfera unitaria, cioè tale che $\|x_1\| = 1$. Per il lemma di Riesz esiste x_2 sulla sfera unitaria tale che dista almeno $\frac{1}{2}$ dal sottospazio generato da x_1, per cui si ha $\|x_1 - x_2\| \geq \frac{1}{2}$. Sempre per lo stesso lemma, esiste ancora x_3 con $\|x_3\| = 1$ che è distante almeno $\frac{1}{2}$ dal sottospazio generato da x_1e da x_2, cioè x_3 è tale che :$\|x_1 - x_3\| \geq \frac{1}{2}$ e $\|x_2 - x_3\| \geq \frac{1}{2}$. In questo modo si può costruire una successione $(x_k)_{k\in\mathbb{N}}$ di vettori sulla sfera unitaria , che distano almeno $\frac{1}{2}$ l'uno dall'altro. Dalla successione $(x_k)_{k\in\mathbb{N}}$, limitata sulla sfera unitaria, è evidente che non si riesce ad estrarre una sottosuccessione convergente, perciò la sfera non può essere un compatto. $\blacksquare$

Capitolo 2

Operatori lineari

2.1 Introduzione

Dopo aver introdotto la definizione d'applicazione lineare tra due spazi vettoriali e aver richiamato alcuni concetti alla base dell'algebra lineare, passeremo a considerare le applicazioni lineari tra spazi normati, dette operatori lineari. In generale non tutti gli operatori lineari sono continui, perciò alla luce delle nozioni introdotte sugli spazi normati e di Banach evidenzieremo le condizioni che garantiscono la continuità di un operatore lineare. Quindi dopo aver definito la norma di un operatore lineare ci soffermeremo sullo spazio degli operatori lineari continui, introducendo anche la nozione di convergenza puntuale e uniforme. Infine riporteremo alcuni risultati sull'invertibilità degli operatori lineari limitati.(cfr [13])

2.2 Linearità e continuità

Definizione 2.2.1 *Siano X e Y due spazi vettoriali su $\mathbb{R}$ e sia $\varphi : X \to Y$. L'applicazione φ è lineare se:*

$$\varphi(x+y) = \varphi(x) + \varphi(y) \quad \forall x, y \in X$$

$$\varphi(\lambda x) = \lambda\varphi(x) \quad \forall x \in X \quad \forall \lambda \in \mathbb{R}$$

Osservazione 2.2.2 *Per ogni applicazione lineare φ si ha $\varphi(0) = 0$*

Osservazione 2.2.3 *Ogni applicazione lineare è univocamente individuata dai valori che assume su una base.*

Inoltre ricordiamo che se consideriamo due spazi vettoriali X e Y di dimensione finita, fissate le basi dei due spazi, ad ogni applicazione lineare $\varphi : X \to Y$ si può associare una matrice che la rappresenta. Per le nozioni di immagine e nucleo di una applicazione lineare si rimanda a [5] o ad un qualunque testo di algebra lineare. Riportiamo solo la formula del rango di cui ci serviremo per dimostrare alcuni risultati:

Proposizione 2.2.4 *Se $\varphi : X \to Y$ è un'applicazione lineare e se X ha dimensione finita allora:*

$$Rank\ \varphi := \dim \operatorname{Im} \varphi = \dim X - \dim Ker\ \varphi \tag{2.1}$$

Osservazione 2.2.5 *Se X ha dimensione finita e $\varphi : X \to Y$ è un'applicazione lineare, anche $\varphi(X)$ ha dimensione finita.*

Poiché per gli operatori lineari la continuità in ogni punto equivale alla continuità nell'origine, (cfr [6]), sembrerebbe allora quasi immediato che ogni operatore lineare è continuo. Effettivamente però questa implicazione vale solo per gli spazi normati di dimensione finita. Prima di dimostrarlo, introduciamo però un teorema che fornisce una caratterizzazione degli operatori lineari continui tra spazi normati.

Teorema 2.2.6 *Un operatore lineare $L : X \to Y$ tra spazi normati è continuo $\forall x \in X$ se e solo se vale una delle seguenti condizioni:*

1. *L è continuo in 0*

2. *L è limitato sulla palla unitaria di X, cioè $\exists k > 0 \; \|L(x)\|_Y \leq k \quad \forall x \in X : \|x\|_X \leq 1$*

3. *$\exists k > 0 : \|L(x)\|_Y \leq k \|x\|_X \quad \forall x \in X$*

Dimostrazione. Indichiamo con (0) l'asserzione: "L è continuo $\forall x \in X$ ".

(0) $\Longrightarrow$ (1) è banale.

(1) $\Longrightarrow$ (2)

Sia L continuo in 0, cioè:

$$\forall \varepsilon > 0 \; \exists \; \delta > 0 : \|x\|_X \leq \delta \Longrightarrow \|L(x)\|_Y \leq \varepsilon \leq 1 \tag{2.2}$$

Ora se ci mettiamo sulla palla unitaria, cioè se $\forall x \in X$ consideriamo $z = \lambda x$ in modo che $\|z\|_X \leq 1$:

$$\|z\|_X = |\lambda| \|x\|_X \leq 1 \text{ cioè } \|x\|_X \leq \frac{1}{|\lambda|}$$

allora per la (2.2) e per la linearità di L si ha:

$$\exists \; \delta > 0 : \|x\|_X \leq \delta \Longrightarrow \|L(\lambda x)\|_Y = \|\lambda L(x)\|_Y = |\lambda| \|L(x)\|_Y \leq |\lambda| \tag{2.3}$$

quindi ponendo $\delta = \frac{1}{|\lambda|}$ abbiamo che:

$$\|z\|_X \leq 1 \Longrightarrow \|L(z)\|_Y \leq \frac{1}{\delta}$$

che è la (2).

$(2) \Longrightarrow (3)$

Per ogni $x \in X$, $x \neq 0$ definiamo il vettore unitario:

$$z = \frac{x}{\|x\|_X} \Longrightarrow \|z\|_X \leq 1.$$

Dalla (2) segue che:

$$\exists\ k > 0 : \|z\|_X \leq 1 \Longrightarrow \|L(z)\|_Y \leq k \|x\|_X\ .$$

Dalla linearità di L risulta che:

$$\|L(z)\|_Y = \left\| L\left(\frac{x}{\|x\|_X}\right)\right\|_Y = \frac{1}{\|x\|_X} \|L(x)\|_Y \leq k$$

cioè

$$\exists k > 0 : \|L(x)\|_Y \leq k \|x\|_X \quad \forall x \in X$$

che è la (3).

$(3) \Longrightarrow (0)$

Applichiamo la (3) a $z = x - x_0$ con x, $x_0 \in X$ allora si ha:

$$\exists k > 0 : \|L(z)\|_Y \leq k \|z\|_X$$

$$\exists k > 0 : \|L(x - x_0)\|_Y = \|L(x) - L(x_0)\|_Y \leq k \|x - x_0\|_X$$

dunque L è lipschtziana e dunque continua su tutto X ■

Proposizione 2.2.7 *Se X e Y sono spazi normati con X di dimensione finita, allora ogni operatore lineare $L : X \to Y$ è continuo.*

Dimostrazione. Essendo X di dimensione finita e L lineare per la formula del rango (2.1) si ha che anche $L(X)$ ha dimensione finita, infatti:

$$\dim\ L(X) = \dim X - \dim KerL$$

Siano:

- $n = \dim X$ e $m = \dim L(X)$
- $\{v_1, ..., v_n\}$ una base di X
- $\{\omega_1, ..., \omega_n\}$ una base di $L(X)$
- $v \in X$ cioè $v = \sum_{i=1}^{n} x^i v_i \quad x = \left(x^1, ..., x^n\right)$
- $\omega \in X$ cioè $\omega = \sum_{i=1}^{m} y^i \omega_i \quad y = \left(y^1, ..., y^m\right)$
- $\|v\|_X := \sum_{i=1}^{n} \left|x^i\right| \ \forall v \in X$
- $\|\omega\|_{L(X)} := \sum_{i=1}^{m} \left|y^i\right| \ \forall \omega \in L(X)$

La dimostrazione di quello che segue è indipendente dalla scelta delle norme, poichè essendo X e $L(X)$ di dimensione finita, in essi tutte le norme sono equivalenti tra loro. Sia A la matrice associata a L, allora:

$$\omega = L(v) \quad \Longleftrightarrow \quad y = Ax \text{ con } \ y^i = \sum_{j=1}^{n} a_{ij}\ x^j \ \text{ per } 1 \leq i \leq m$$

$$L(v) = \omega = \sum_{i=1}^{m} y^i \omega_i = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}\ x^j \omega_i$$

Passando alle norme:

$$\begin{aligned}
\|L(v)\|_{L(X)} &= \sum_{i=1}^{m} \left|y^i\right| = \sum_{i=1}^{m} \left|\sum_{j=1}^{n} a_{ij}\ x^j\right| \leq \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}| \ \ \left|x^j\right| \\
&= \sum_{j=1}^{n} \left|x^j\right| \sum_{i=1}^{m} |a_{ij}| = x^1 \left(a_{11} + a_{21} + ...a_{m1}\right) + ... + x^n \left(a_{1n} + a_{2n} + ...a_{mn}\right) \\
&\leq M \sum_{j=1}^{n} \left|x^j\right| = M \|v\|_X \quad \text{con} \quad M = \max_{1\leq\ j\leq n} \sum_{i=1}^{m} |a_{ij}|
\end{aligned}$$

dunque

$$\exists\ M : \|L(v)\|_{L(X)} \leq M \|v\|_X \quad \forall v \in X$$

cioè per il teorema 2.2.6 L è continua. ∎

Per un operatore lineare:

$$L : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$$

possiamo definire la norma:

$$\|L\| := \sup_{\|x\|_X \leq 1} \|L(x)\|_Y \tag{2.4}$$

Dalla (2) del teorema 2.2.6, si ha infatti che L è limitata sulla palla unitaria cioè:

$$\exists k > 0 \ \ \|L(x)\|_Y \leq k \ \ \forall x \in X : \|x\|_X \leq 1$$

che equivale a dire che $\|L\| < \infty$

quindi:

$$\|L\| = \sup_{\|x\|_X \leq 1} \|L(x)\|_Y = \inf_{k \in \mathbb{R}^+} \left\{\|L(x)\|_Y \leq k \|x\|_X, \forall x \in X\right\}$$

Osservazione 2.2.8 *Un operatore lineare è continuo se e solo se è limitato.*

Osservazione 2.2.9 *Inoltre se L è lineare vale:*

$$\|L(x)\|_Y \leq \|L\| \|x\|_X , \forall x \in X \ . \tag{2.5}$$

2.3 Lo spazio degli operatori lineari continui

Notazione 2.3.1 *Siano X e Y due spazi normati indicheremo con:*

$L(X,Y)$ *lo spazio degli operatori lineari continui* $\varphi : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$

$End(X)$ *lo spazio degli operatori lineari continui* $\varphi : (X, \|\cdot\|_X) \to (X, \|\cdot\|_X)$

Lo spazio $L(X,Y)$ è uno spazio vettoriale rispetto alle operazioni di somma e prodotto per uno scalare, in particolare è un sottospazio dello spazio degli operatori lineari ed è anche uno spazio normato munito della norma (2.4).

Teorema 2.3.2 *Se Y è uno spazio di Banach allora lo è anche $L(X,Y)$ con la norma definita da (2.4).*

Dimostrazione. Sia $(L_n)_{n\in\mathbb{N}} \subset L(X,Y)$ una successione di Cauchy in $L(X,Y)$ cioè si ha:

$$\forall \varepsilon > 0 \ \ \exists \, \nu \in \mathbb{N} : \|L_n - L_m\| \leq \varepsilon \ \ \forall n, m \geq \nu$$

da cui per la (2.5) abbiamo che:

$$\forall \varepsilon > 0 \ \ \exists \, \nu \in \mathbb{N} : \|L_n(x) - L_m(x)\|_Y \leq \varepsilon \|x\|_X \ \ \ \forall x \in X \ \ \forall n, m \geq \nu \tag{2.6}$$

Quindi $(L_n(x))_{n\in\mathbb{N}}$ è una successione di Cauchy in Y ed essendo Y di Banach, allora $(L_n(x))_{n\in\mathbb{N}}$ è convergente in Y, cioè:

$$L_n(x) \to L(x) \, , \text{per } n \to \infty \ \forall x \in X \ .$$

Ora osserviamo che:

$L : X \to Y$ è lineare. Infatti essendo L_n lineare $\forall x, y \in X$ abbiamo che:

$$L_n (x + y) = L_n (x) + L_n (y)$$

$$L_n (\lambda x) = \lambda L_n (x)$$

passando al limite per $n \to \infty$

$$L (x + y) = L (x) + L (y)$$

$$L (\lambda x) = \lambda L (x)$$

$L : X \to Y$ è continuo, infatti dalla (2.6) per $m \to \infty$ si ha che:

$$\forall \varepsilon > 0 \;\; \exists \, \nu \in \mathbb{N} : \|L_n(x) - L(x)\|_Y \leq \varepsilon \, \|x\|_X \quad \forall x \in X \;\; \forall n \geq \nu \tag{2.7}$$

dunque $L_n - L$ è continuo, segue che:

$$L = L_n - (L_n - L) \text{ è continua.}$$

In definitiva $L \in L(X, Y)$, inoltre dalla (2.7) si ha che:

$$\forall \varepsilon > 0 \;\; \exists \, \nu \in \mathbb{N} : \|L_n(x) - L(x)\|_Y \leq \varepsilon \quad \forall n \geq \nu \;\; \forall x \in X : \|x\|_X \;\; \leq 1$$

che è equivalente a:

$$\forall \varepsilon > 0 \;\; \exists \, \nu \in \mathbb{N} : \|L_n - L\|_Y \leq \varepsilon \quad \forall n \geq \nu$$

cioè la successione $(L_n)_{n \in \mathbb{N}}$ converge a $L \in L(X, Y)$ dunque $L(X, Y)$ è di Banach. ■

Siano ora $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ due spazi di Banach e sia $(L_n)_{n\in\mathbb{N}} \subset L(X,Y)$, distinguiamo il concetto di convergenza puntuale da quelllo della convergenza uniforme.

Definizione 2.3.3 $(L_n)_{n\in\mathbb{N}} \subset L(X,Y)$ *converge puntualmente a* $L \in L(X,Y)$ *se:*

$$\forall x \in X \quad \forall \varepsilon > 0 \quad \exists\ \nu \in \mathbb{N} : \|L_n(x) - L(x)\|_Y \leq \varepsilon \quad \forall n \geq \nu. \tag{2.8}$$

Definizione 2.3.4 $(L_n)_{n\in\mathbb{N}} \subset L(X,Y)$ *converge uniformemente a* $L \in L(X,Y)$ *se:*

$$\forall \varepsilon > 0 \quad \exists\ \nu \in \mathbb{N} : \sup_{\|x\|_X \leq 1} \|L_n(x) - L(x)\|_Y \leq \varepsilon \quad \forall n \geq \nu, \forall x \in X. \tag{2.9}$$

Osservazione 2.3.5 *Nella (2.8)* ν *dipende sia da* ε *che da* x*, invece nella (2.9)* ν *dipende solo da* ε*.*

Proposizione 2.3.6 *La convergenza uniforme implica sempre quella puntuale.*

Proposizione 2.3.7 *Se* X *ha dimensione finita allora la convergenza puntuale implica la convergenza uniforme.*

Osservazione 2.3.8 *Se* X *ha dimensione infinita, in generale la convergenza puntuale non implica quella uniforme.*

Consideriamo ora l'operazione di composizione di operatori lineari continui. Siano $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, $(Z, \|\cdot\|_Z)$ tre spazi normati e siano $f \in L(X,Y)$ e $g \in L(Y,Z)$; l'operatore $g \circ f : X \to Z$ è ancora un operatore lineare continuo, cioè $g \circ f \in L(X,Z)$ e si ha la seguente proprietà:

$$\forall x \in X \quad \|g \circ f(x)\|_Z \leq \|g\| \, \|f(y)\|_Y \leq \|g\| \, \|f\| \, \|x\|_X \quad \text{quindi} \quad \|g \circ f\| \leq \|g\| \, \|f\| \ . \tag{2.10}$$

Lo spazio $End(X)$, analogamente a $L(X,Y)$ è uno spazio vettoriale ed è uno spazio normato con la norma usuale (2.4) e vale:

Proposizione 2.3.9 *Se X è uno spazio di Banach allora lo è anche $End(X)$.*

Sia X di Banach, il prodotto di composizione $\circ$ definisce sullo spazio di Banach $End(X)$ una struttura di algebra. Giacché vale la (2.10) si dice che $End(X)$ è un'algebra di Banach, non commutativa rispetto alla composizione di operatori.

Osservazione 2.3.10 *Se $L \in End(X)$ e $L^n = \underbrace{L \circ L \circ ... \circ L}_{n\ volte}$ allora $\|L^n\| \leq \|L\|^n$.*

Consideriamo ora gli operatori lineari da uno spazio normato X in $\mathbb{R}$ e introduciamo il concetto di duale di uno spazio.

Definizione 2.3.11 *Sia X uno spazio normato da $\|\cdot\|_X$, il duale di X è lo spazio X^* degli operatori lineari e continui su X.*

X^*è munito della norma

$$\|f\|_{X^*} = \sup_{\substack{x \in X \\ \|x\|_X \leq 1}} f(x) \tag{2.11}$$

Notazione 2.3.12 *Se $f \in X^*$ scriveremo $\langle f, x \rangle$ in luogo di $f(x)$ e diremo che $\langle\ ,\ \rangle$ è il prodotto scalare nella dualità X^*, X.*

Introduciamo un noto corollario del teorema di Hahn-Banach, che è una caratterizzazione della norma su uno spazio tramite la norma sul suo spazio duale. Per la dimostrazione del corollario e per il teorema di Hahn-Banach si rimanda a [2] oppure a [13].

Corollario 2.3.13 *Sia X uno spazio vettoriale su $\mathbb{R}$. Allora $\forall x \in X$ si ha:*

$$\|x\|_X = \sup_{\substack{f \in X^* \\ \|f\|_{E^*} \leq 1}} |\langle f, x \rangle|$$

2.4 Invertibilità di operatori lineari

I due teoremi che seguono forniscono importanti informazioni sull'esistenza e la natura dell'inversa di un'applicazione lineare sotto opportune condizioni. In particolare la formula (2.12) è strettamente connessa alla teoria delle equazioni integrali, generalmente conosciuta come "espansione di Neumann".

Notazione 2.4.1 *In questo paragrafo indicheremo con* $\|\cdot\|$ *la norma su* $End(X)$*.*

Teorema 2.4.2 *Sia* $(X, \|\cdot\|_X)$ *di Banach e sia* $A \in End(X)$ *tale che la sua serie di Neumann* $\sum_{n=1}^{\infty} A^n$ *sia convergente in* $End(X)$*. Allora l'operatore* $I - A$ *è invertibile e*

$$(I - A)^{-1} = \sum_{n=1}^{\infty} A^n \ . \tag{2.12}$$

Dimostrazione. Per ipotesi la serie di Neumann di A converge in $End(X)$, dunque $\exists B \in End(X)$ tale che $B = \sum_{n=1}^{\infty} A^n$ e poichè la serie è convergente il suo termine generale è infinitesimo, cioè:

$$\lim_{n\to\infty} A^n = 0 \tag{2.13}$$

Consideriamo ora

$$(I - A)\sum_{k=1}^{n} A^k = \sum_{k=1}^{n} A^k - \sum_{k=1}^{n} A^{k+1} = I - A^{n+1}$$

passando al limite per $n \to \infty$, per la (2.13) si ottiene:

$$(I - A)\, B = I$$

e poiché

$$AB = BA = \sum_{n=1}^{\infty} A^{n+1}$$

si ha:

$$(I - A)\,B = B(I - A) = I$$

dunque la (2.12) ■

Corollario 2.4.3 *Sia* $(X, \|\cdot\|_X)$ *di Banach e sia* $A \in End\,(X)$ *tale che* $\|A\| < 1$,*allora* $I - A$ *è invertibile, vale la (2.12) e*

$$\left\|(I - A)^{-1}\right\| \leq \frac{1}{1 - \|A\|} \quad . \tag{2.14}$$

Dimostrazione. Per l'osservazione 2.3.10 si ha:

$$\|A^n\| \leq \|A\|^n$$

e quindi

$$\sum_{n=0}^{\infty} \|A^n\| \leq \sum_{n=0}^{\infty} \|A\|^n$$

Poiché $\|A\| < 1$, la sua $\sum\limits_{n=0}^{\infty} \|A\|^n$converge a $\frac{1}{1-\|A\|}$. Allora:

$$\sum_{n=0}^{\infty} \|A^n\| \leq \sum_{n=0}^{\infty} \|A\|^n = \frac{1}{1 - \|A\|} \tag{2.15}$$

cioè la serie $\sum\limits_{n=0}^{\infty} A^n$è assolutamente convergente e quindi convergente in $End(X)$.
Allora per il teorema 2.4.2 vale la (2.12) e inoltre:

$$\left\|(I - A)^{-1}\right\| = \left\|\sum_{n=1}^{\infty} A^n\right\| \leq \sum_{n=0}^{\infty} \|A^n\| \tag{2.16}$$

Dalla (2.15) si ha la tesi (2.14). ■

Lemma 2.4.4 ***(di Banach)*** *Sia* X *di Banach e siano* $A, B \in End(X)$. *Se* A *è invertibile e se*

$$\|A - B\| \leq \frac{1}{\|A^{-1}\|} \tag{2.17}$$

allora anche B è invertibile e inoltre valgono:

$$\left\|B^{-1}\right\| \leq \frac{\left\|A^{-1}\right\|}{1-\|A^{-1}\|\,\|A-B\|} \tag{2.18}$$

$$\left\|B^{-1}-A^{-1}\right\| \leq \frac{\left\|A^{-1}\right\|^{2}\|A-B\|}{1-\|A^{-1}\|\,\|A-B\|}\ . \tag{2.19}$$

Dimostrazione. Osservando che:

$$I-A^{-1}B=A^{-1}(A-B)$$

passando alle norme e tenendo conto della (2.10) e della ipotesi (2.17) si ha:

$$\left\|I-A^{-1}B\right\|=\left\|A^{-1}(A-B)\right\| \leq \left\|A^{-1}\right\|\|A-B\|<1.$$

Allora per il corollario 2.4.3, l'operatore $I-\left(I-A^{-1}B\right)=A^{-1}B$ è invertibile e di conseguenza anche B lo è. Si ha:

$$\left\|\left(A^{-1}B\right)^{-1}\right\| \leq \frac{1}{1-\|I-A^{-1}B\|}=\frac{1}{1-\|A^{-1}(A-B)\|} \leq \frac{1}{1-\|A^{-1}\|\,\|A-B\|} \tag{2.20}$$

Essendo:

$$B^{-1}=B^{-1}\left(AA^{-1}\right)=\left(B^{-1}A\right)A^{-1}=\left(A^{-1}B\right)^{-1}A^{-1}$$

passando alle norme e tenendo conto della (2.10) e della (2.20) si ha la tesi (2.18)

$$\left\|B^{-1}\right\|=\left\|\left(A^{-1}B\right)^{-1}A^{-1}\right\| \leq \left\|\left(A^{-1}B\right)^{-1}\right\|\left\|A^{-1}\right\| \leq \frac{\left\|A^{-1}\right\|}{1-\|A^{-1}\|\,\|A-B\|}$$

Considerando poi che

$$B^{-1}-A^{-1}=A^{-1}(A-B)\,B^{-1}$$

per la (2.18) si ha:

$$\left\|B^{-1}-A^{-1}\right\|=\left\|A^{-1}(A-B)\,B^{-1}\right\| \leq \left\|A^{-1}\right\|\|A-B\|\left\|B^{-1}\right\| \leq \frac{\left\|A^{-1}\right\|^{2}\|A-B\|}{1-\|A^{-1}\|\,\|A-B\|}$$

■

Capitolo 3

Calcolo differenziale negli spazi di Banach

3.1 Introduzione

In questo capitolo esponiamo le nozioni più elementari del calcolo differenziale negli spazi di Banach (cfr. [12] per una trattazione completa dell'argomento). Le definizioni di derivata secondo Frèchet e secondo Gateâux non sono altro che la naturale estensione agli spazi di Banach rispettivamente dei concetti classici di derivata e derivata direzionale per le applicazioni tra spazi in $\mathbb{R}^n$. Infine introduciamo anche i concetti di derivate successive e formula di Taylor per operatori tra spazi di Banach, utili nel seguito.

Notazione 3.1.1 *Nei seguenti paragrafi consideriamo sempre* $(X, \|\cdot\|)$ *e* $(Y, \|\cdot\|_Y)$ *spazi di Banach.*

3.2 Differenziabilità secondo Frèchet

Definizione 3.2.1 *Sia $\Omega \subset X$ un aperto e sia $f : \Omega \subset X \to Y$. f è differenziabile secondo Frèchet in $x_0 \in \Omega$ se esiste un'applicazione lineare continua $A : X \to Y$ tale che:*

$$\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - Ah\|_Y}{\|h\|} = 0 \tag{3.1}$$

L'applicazione A è detta differenziale secondo Frèchet in x_0.

Proposizione 3.2.2 *Il differenziale secondo Frèchet in un punto è unico.*

Dimostrazione. Supponiamo per assurdo che esistano A_1e A_2 con $A_1 \neq A_2$ tali che:

$$\begin{aligned} \lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - A_1 h\|_Y}{\|h\|} = 0 \\ \lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - A_2 h\|_Y}{\|h\|} = 0 \end{aligned} \tag{3.2}$$

allora si ha:

$$\|(A_1 - A_2) h\|_Y \leq \|f(x_0 + h) - f(x_0) - A_1 h\|_Y + \|f(x_0 + h) - f(x_0) - A_2 h\|_Y$$

quindi

$$\frac{\|(A_1 - A_2) h\|_Y}{\|h\|} \leq \frac{\|f(x_0 + h) - f(x_0) - A_1 h\|_Y}{\|h\|} + \frac{\|f(x_0 + h) - f(x_0) - A_2 h\|_Y}{\|h\|}$$

passando al limite per le (3.2) si ha:

$$\lim_{h \to 0} \frac{\|(A_1 - A_2) h\|_Y}{\|h\|} = 0 \Leftrightarrow \forall \varepsilon > 0 \; \exists \delta > 0 : \frac{\|(A_1 - A_2) h\|_Y}{\|h\|} < \varepsilon, \forall h : \|h\| < \delta$$

essendo A_1, A_2 lineari continue per la (2.10) si ha:

$$\forall \varepsilon > 0 \; \exists \delta > 0 : \|A_1 - A_2\|_{L(X,Y)} < \varepsilon, \forall h : \|h\| < \delta$$

cioè $A_1 = A_2$. ■

Definizione 3.2.3 *Sia $\Omega \subset X$ un aperto e sia $f : \Omega \subset X \to Y$ Frèchet differenziabile in ogni punto di Ω. L'applicazione $f' : \Omega \subset X \to L(X, Y)$ che associa ad ogni punto il differenziale secondo Frèchet di f calcolato in quel punto si dice derivata prima di f.*

Proposizione 3.2.4 *Se f è continua in x_0 allora il suo differenziale in x_0 è continuo.*

Dimostrazione. Posto:

$$\sigma(h) = f(x_0 + h) - f(x_0) - Ah \ , \ \sigma(0) = 0$$

essendo $\sigma(0) = 0$ segue che σ è continua. Poichè:

$$Ah = f(x_0 + h) - f(x_0) - \sigma(h)$$

si ha che per $h = 0$, A è continuo. ■

Proposizione 3.2.5 *Se f è differenziabile secondo Frèchet in x_0 allora f è continua in x_0.*

Dimostrazione. Consideriamo

$$\begin{aligned} 0 &\leq \|f(x_0 + h) - f(x_0)\|_Y = \|f(x_0 + h) - f(x_0) - Ah + Ah\|_Y \\ &\leq \|f(x_0 + h) - f(x_0) - Ah\|_Y + \|Ah\|_Y = \frac{\|f(x_0 + h) - f(x_0) - Ah\|_Y}{\|h\|} \|h\| + \|Ah\|_Y \end{aligned}$$

poichè f è Frèchet differenziabile in x_0 passando al limite per $h \to 0$ si ottiene la tesi:

$$\lim_{h \to 0} \|f(x_0 + h) - f(x_0)\|_Y = 0$$

■

Esempi di applicazioni Frèchet differenziabili

1. Un'applicazione costante è differenziabile secondo Frèchet in ogni punto del suo dominio ed ha derivata prima costantemente nulla.

2. $f \in L(X,Y)$ è Frèchet differenziabile in ogni punto del suo dominio e la sua derivata prima è costantemente uguale ad f

3. Un'applicazione bilineare continua

$$b : X \times Z \to Y \quad (X, Y, Z \text{ spazi di Banach})$$

è differenziabile secondo Frèchet in ogni punto di $X \times Z$ e la sua derivata in $(x_0, z_0) \in X \times Z$ è:

$$b'(x_0, z_0)(h, k) = b(h, z_0) + b(k, x_0)$$

3.3 Differenziabilità secondo Gateâux

Definizione 3.3.1 *Sia $\Omega \subset X$ un aperto e sia $f : \Omega \subset X \to Y$. f è differenziabile secondo Gateâux in $x_0 \in \Omega$ se esiste un'applicazione lineare continua $A : X \to Y$ tale che $\forall v \in X$*

$$\lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t} = Av \ .$$

L'applicazione A si dice differenziale secondo Gateâux in x_0.

Per l'unicità del limite si ha che:

Proposizione 3.3.2 *Il differenziale secondo Gateâux in un punto è unico.*

Definizione 3.3.3 *Sia $\Omega \subset X$ un aperto e sia $f : \Omega \subset X \to Y$ differenziabile secondo Gateâux in ogni punto di Ω. L'applicazione $f'_G : \Omega \subset X \to L(X, Y)$ che associa ad ogni punto il differenziale secondo Gateâux di f calcolato in quel punto si dice derivata secondo Gateâux di f.*

Proposizione 3.3.4 *Se $F : \Omega \subset X \to Y$ è Frèchet differenziabile in $x_0 \in \Omega$, allora è anche Gateâux differenziabile in x_0 e $F'_G(x_0) = F'(x_0)$*

Dimostrazione. Per ipotesi vale la (3.1) che è equivalente a:

$$\forall \varepsilon > 0 \ \exists \delta : \frac{\|F(x_0 + h) - F(x_0) - F'(x_0)h\|_Y}{\|h\|} < \varepsilon \quad \forall h : \|h\| < \delta \tag{3.3}$$

Consideriamo ora

$$\left\| \frac{F(x_0 + tv) - F(x_0) - F'(x_0)tv}{t} \right\|_Y = \frac{\|F(x_0 + tv) - F(x_0) - F'(x_0)tv\|_Y}{|t|}$$

posto $h = tv$

$$\left\| \frac{F(x_0 + tv) - F(x_0) - F'(x_0)tv}{t} \right\|_Y = \frac{\|F(x_0 + h) - F(x_0) - F'(x_0)h\|_Y}{\|h\|} \|v\| < \varepsilon \|v\|$$

dunque

$$\forall v \in X \quad \lim_{t \to 0} \frac{F(x_0 + tv) - F(x_0)}{t} - F'(x_0)v = 0$$

che equivale a

$$\forall v \in X \quad \lim_{t \to 0} \frac{F(x_0 + tv) - F(x_0)}{t} = F'(x_0)v$$

in definitiva $F'_G(x_0) = F'(x_0)$ ■

Osservazione 3.3.5 *La differenziabilità secondo Gateâux non implica la continuità e dunque non implica la differenziabilità secondo Frèchet.*

$Esempio\ di\ funzione\ Gateâux\ differenziabile\ in\ un\ punto\ ma\ non\ Frèchet\ differenziabile\ nello\ stesso\ punto$

$$f(x,y) = \begin{cases} \frac{x^4y^2}{(x^4+y^2)^2} & \text{se } (x,y) \neq (0,0) \\ 0 & \text{se } (x,y) = (0,0) \end{cases}$$

Osserviamo che la f è differenziabile secondo Gateâux in $(0,0)$, infatti:

$$\forall (x,y) \in \mathbb{R}^2 \quad \lim_{t\to 0} \frac{f(tx,ty) - f(0,0)}{t} = \lim_{t\to 0} \frac{t^4x^4t^2y^2}{t\left(t^4x^4 + t^2y^2\right)^2}$$

$$= \lim_{t\to 0} \frac{t^6x^4y^2}{t^5\left(t^2x^4 + y^2\right)^2} = \lim_{t\to 0} \frac{tx^4y^2}{\left(t^2x^4 + y^2\right)^2} = 0$$

Ma f non è continua in $(0,0)$ infatti la successione $\left(\frac{1}{n}, \frac{1}{n^2}\right)_{n\in\mathbb{N}}$ converge a $(0,0)$, tuttavia

$$f\left(\frac{1}{n}, \frac{1}{n^2}\right) = \frac{n^8}{4n^8} = \frac{1}{4}$$

ciò vuol dire che in un intorno dell'origine vi sono punti per cui

$\mid f(x,y) \mid > \varepsilon$, cioè che f non è continua nell'origine e dunque non essendo continua, non può essere differenziabile secondo Frèchet, per la proposizione 3.2.5.

Una funzione differenziabile secondo Gateâux è differenziabile su tutte le rette, dunque possiamo dare l'importante maggiorazione contenuta nel seguente teorema:

Teorema 3.3.6 *Sia $\Omega \subset X$ un aperto e sia $f : \Omega \subset X \to Y$ differenziabile secondo Gateâux in ogni punto di Ω. Siano $x_1, x_2 \in \Omega$ tali che il segmento che li unisce sia* *tutto contenuto in* Ω*. Allora:*

$$\|f(x_2) - f(x_1)\|_Y \leq \sup_{0 \leq t \leq 1} \left\|f'_G \circ \gamma(t)\right\|_{L(X,Y)} \|x_2 - x_1\| \tag{3.4}$$

dove $\gamma : t \longmapsto tx_2 + (1-t)\, x_1 \quad \forall t \in [0,1]$ *è il segmento che congiunge i punti* x_1, x_2.

Dimostrazione. Consideriamo la funzione $f \circ \gamma : [0,1] \to Y$, essa è differenzia-bile secondo Gateâux,con derivata:

$$(f \circ \gamma)'_G(t) = f'_G(\gamma(t))\, \gamma'(t) = f'_G(tx_2 + (1-t)\, x_1)\,(x_2 - x_1) \tag{3.5}$$

Sia ora $\xi \in Y^*$,cioè $\xi : Y \to \mathbb{R}$ lineare continua. La funzione reale

$$t \longmapsto \xi(f \circ \gamma(t)) = \langle \xi, f \circ \gamma(t) \rangle$$

è derivabile.

Applicando il teorema del valor medio si ha che:

$$\exists\, \theta \in [0,1] : (\langle \xi, f \circ \gamma(\theta) \rangle)'_G = \frac{\langle \xi,\ f \circ \gamma(1) - f \circ \gamma(0) \rangle}{1 - 0} \tag{3.6}$$

poichè $\xi'_G = \xi$, essendo ξ lineare, si ha dalla (3.6) che:

$$\exists\, \theta \in [0,1] : \left\langle \xi, (f \circ \gamma)'_G(\theta) \right\rangle = \langle \xi,\ f(x_2) - f(x_1) \rangle$$

per la (3.5)

$$\exists\, \theta \in [0,1] : \left\langle \xi, f'_G(\theta x_2 + (1-\theta)\, x_1)\,(x_2 - x_1) \right\rangle = \langle \xi,\ f(x_2) - f(x_1) \rangle$$

Risulta allora che:

$$\forall \xi \in Y^* \quad |\langle \xi,\ f(x_2) - f(x_1) \rangle| \leq \left(\sup_{0 \leq t \leq 1} \left\|f'_G \circ \gamma(t)\right\|_{L(X,Y)} \right) \|\xi\|_{Y^*} \|x_2 - x_1\| \tag{3.7}$$

Applicando il corollario 2.3.13 del teorema di Hahn-Banach a $f(x_2) - f(x_1) \in Y$ si ha:

$$\|f(x_2) - f(x_1)\|_Y = \sup_{\substack{\xi \in Y^* \\ \|\xi\|_{Y^*} \leq 1}} |\langle \xi, f(x_2) - f(x_1) \rangle| \tag{3.8}$$

dunque dalla (3.8) per la (3.7) otteniamo:

$$\begin{aligned} \|f(x_2) - f(x_1)\|_Y &\leq \sup_{\substack{\xi \in Y^* \\ \|\xi\|_{Y^*} \leq 1}} \left(\sup_{0 \leq t \leq 1} \left\| f'_G \circ \gamma(t) \right\|_{L(X,Y)} \right) \|\xi\|_{Y^*} \|x_2 - x_1\| \\ &\leq \left(\sup_{0 \leq t \leq 1} \left\| f'_G \circ \gamma(t) \right\|_{L(X,Y)} \right) \|x_2 - x_1\| \end{aligned}$$

■

Teorema 3.3.7 *Se $f : \Omega \subset X \to Y$ è differenziabile secondo Gateâux e se*

$$f'_G : \Omega \subset X \to L(X, Y)$$

è continua in x_0, allora f è Frèchet differenziabile in x_0 con

$$f'_G(x_0) = f'(x_0)$$

Dimostrazione. Applicando il teorema 3.3.6 alla funzione

$$\sigma(h) = f(x_0 + h) - f(x_0) - f'_G(x_0)\, h$$

considerando il segmento $\|h\| \leq \rho$ con $\rho > 0$ fissato e tenendo presente che $\sigma(0) = 0$ abbiamo:

$$\|\sigma(h) - \sigma(0)\|_Y \leq \sup_{\|h\| \leq \rho} \left\| f'_G(x_0 + h) - f'_G(x_0) \right\|_{L(X,Y)} \|h\|$$

$$\frac{\|\sigma(h)\|_Y}{\|h\|} \leq \sup_{\|h\| \leq \rho} \left\| f'_G(x_0 + h) - f'_G(x_0) \right\|_{L(X,Y)} \tag{3.9}$$

passando al limite per $h \to 0$ per la continuità di f'_G in x_0 si ha la tesi. ■

3.4 Derivate successive e Formula di Taylor

Definizione 3.4.1 *Sia $\Omega \subset X$ un aperto e sia $f : \Omega \subset X \to Y$ un'applicazione differenziabile secondo Frèchet in ogni punto di Ω e sia $f' : \Omega \subset X \to L(X,Y)$ la derivata prima di f. Supponiamo che f' sia differenziabile in Ω, allora l'applicazione $f'' : \Omega \subset X \to L(X, L(X,Y))$ che associa ad ogni punto il differenziale secondo Frèchet di f' calcolato in quel punto si dice derivata seconda di f. In particolare se f' è differenziabile in $x_0 \in \Omega$,allora l'applicazione $f''(x_0)$ si dice differenziale secondo di f in x_0.*

Procedendo per induzione su n si può definire il differenziale n-esimo di f.

Definizione 3.4.2 *Sia $\Omega \subset X$ un aperto e sia $f : \Omega \subset X \to Y$ un'applicazione differenziabile $n-1$ volte in $x_0 \in \Omega$. Se $x \longmapsto f^{(n-1)}(x)$ è differenziabile in x_0 diremo che f è differenziabile n volte in x_0 e chiameremo differenziale n-esimo di f in x_0 l'applicazione n-lineare $f^{(n)}(x_0)$ e derivata n-esima di f l'applicazione: $f^{(n)} : x \longmapsto f^{(n)}(x) \quad \forall x \in X$.*

Teorema 3.4.3 *Sia $\Omega \subset X$ un aperto convesso e sia $f : \Omega \subset X \to Y$ un'applicazione differenziabile n volte con continuità in $x_0 \in \Omega$ Allora $\forall x \in \Omega$:*

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2!} f''(x_0)(x-x_0, x-x_0) + ...+ \\ &+ \frac{1}{(n-1)!} f^{(n-1)}(x_0)\Big(\underbrace{x-x_0, ..., x-x_0}_{n-1 \ volte}\Big) + \\ &+ \frac{1}{(n-1)!} \int_0^1 (1-t)^{n-1} f^{(n)}(x_0)\Big(\underbrace{(1-t)(x-x_0), ..., (1-t)(x-x_0)}_{n\text{-}volte}\Big)\, dt \end{aligned}$$

l'ultimo termine può essere espresso anche come:

$$\frac{1}{n!}(1-t)^{n-1} f^{(n)}(x_0)\Big(\underbrace{x-x_0, ..., x-x_0}_{n\text{-}volte}\Big) + \epsilon(x_0, x-x_0)\Big(\underbrace{x-x_0, ..., x-x_0}_{n\text{-}volte}\Big)$$

dove $\epsilon\left(x_0, x-x_0\right)=$

$$=\frac{1}{(n-1)!}\int_0^1(1-t)^{n-1}\left(f^{(n)}\left(x-x_0\right)-f^{(n)}\left(x_0\right)\right)\left(\underbrace{(1-t)\left(x-x_0\right),...,(1-t)\left(x-x_0\right)}_{n\text{-}volte}\right)dt$$

Il seguente risultato è un corollario della formula di Taylor, che ci sarà utile nelle dimostrazioni dei risultati del prossimo capitolo:

Corollario 3.4.4 *Sia* $f:\Omega\subset X\rightarrow Y$ *un'applicazione Frèchet differenziabile con derivata continua in* Ω. *Allora* $\forall x, y\in\Omega$ *tali che il segmento congiungente i due punti sia tutto contenuto in* Ω,*si ha:*

$$f(x)-f(y)=\left[\int_0^1 f'(x+t(y-x))\,dt\right](x-y) \tag{3.10}$$

Capitolo 4

Metodo di Newton-Kantorovich

4.1 Introduzione

In questo capitolo daremo alcune nozioni di base sui metodi iterativi come l'ordine e il tipo di convergenza, (cfr. [4] per una trattazione più approfondita). I metodi iterativi devono la loro origine al problema di approssimare gli zeri di un operatore non lineare. L'idea di fondo che si utilizza è tradurre la ricerca degli zeri di un operatore in un problema di punto fisso. Quindi ci concentreremo sul metodo di Newton, che ha affascinato i matematici di ogni tempo (cfr [15]) non solo per le sue svariate applicazioni, ma soprattutto per la ricerca di teoremi che ne garantiscano la sicura convergenza. In particolare, facendo riferimento a [3] e a [9], focalizzeremo la nostra attenzione sul teorema di convergenza per il metodo di Newton negli spazi di Banach stabilito da Kantorovich nel 1948, che è certamente uno dei più importanti risultati dell'analisi non lineare e numerica classica. Il matematico russo, infatti, estende il metodo agli operatori non lineari tra spazi di Banach e dimostra un teorema di convergenza a posteriori per il metodo.

Notazione 4.1.1 *Nei seguenti paragrafi utilizziamo la seguente notazione :*

* $(X, \|\cdot\|)$ *e* $(Y, \|\cdot\|_Y)$ *spazi di Banach.*

* $B(x_0, R) = \{x \in X : \|x - x_0\| < R\}$

cioè la palla aperta di centro x_0 *raggio* $R > 0$

* $\overline{B(x_0, r)} = \{x \in X : \|x - x_0\| \leq r\}$

cioè la palla chiusa di centro x_0 *raggio* $r > 0$

4.2 Metodi iterativi e ordine di convergenza

Spesso si preferisce trasformare il problema della ricerca degli zeri di un operatore continuo e non lineare nella ricerca dei punti fissi di un altro operatore opportunemente costruito.

Ad esempio, $F : D \subseteq X \to Y$ è un operatore continuo e non lineare, x_* è soluzione dell'equazione:

$$F(x) = 0$$

se e soltanto se x_* è un punto fisso per l'operatore $T = I - F$, cioè x_* è soluzione di

$$T(x) = x \ .$$

In generale l'operatore T può essere della forma:

$$T = I - BF$$

dove B è un operatore lineare, limitato e invertibile.

Quindi fissato $x_0 \in D$ come approssimazione iniziale dello zero di F, si può costruire la successione:

$$x_n = T(x_{n-1}) = x_{n-1} - B(x_{n-1}) F(x_{n-1}), \ \forall n \in \mathbb{N}$$

Il procedimento computazionale così realizzato è detto in generale iterazione di punto fisso e se converge, converge ad un punto fisso di T cioè ad uno zero di F.

Casi particolari di metodo iterativi di questa forma sono:

- il metodo di Newton, nel quale $B(x) = [F'(x)]^{-1}$
- il metodo di Newton modificato,in cui $B(x) = [F'(x_0)]^{-1}$
- il metodo delle secanti, dove $B(x) = \int_0^1 F'((1-t)k + tx)dt$ con k fissato.

Sebbene la convergenza di un metodo iterativo sia certamente auspicabile teoricamente, perchè tale processo sia applicabile nella pratica è necessario valutare la velocità di convergenza. A tal fine una proprietà che caratterizza i metodi iterativi è l'ordine di convergenza.

Definizione 4.2.1 *Se* $\exists \ \ p > 0$ *tale che*

$$\limsup_{n \to \infty} \frac{\|x_{n+1} - x_*\|}{\|x_n - x_*\|^p} = c \ con \ c > 0 \tag{4.1}$$

allora p *è detto ordine di convergenza della successione* $(x_n)_{n \in \mathbb{N}}$ *e la costante* c *viene detta costante asintotica di convergenza.*

Se $p = 1$ *la convergenza è detta lineare.*

Se $p < 1$ *la convergenza è detta sublineare.*

Se $p > 1$ *la convergenza è detta superlineare.*

Osservazione 4.2.2 *Nel caso particolare* $p = 2$ *la convergenza è detta quadratica e se* $p = 3$ *la convergenza è detta cubica.*

L'ordine di convergenza di un metodo iterativo consente di valutarne l'efficacia e di poterlo così confrontare con altri metodi. Tuttavia nella scelta di un metodo non bisogna trascurarne mai i costi computazionali. Spesso infatti un'elevata velocità di convergenza implica ad ogni iterazione una mole di calcoli non indifferente, che potrebbe compromettere comunque l'efficienza del metodo scelto.
I teoremi di convergenza per un metodo iterativo possono generalmente essere classificati in due grandi categorie:

- teoremi di convergenza a priori
- teoremi di convergenza a posteriori

Un teorema di convergenza per un metodo iterativo si dice di convergenza a priori, quando supponendo che la soluzione esatta esista e che sia sufficientemente vicina all'approssimazione iniziale, il teorema garantisce la convergenza.
Invece viene detto di convergenza a posteriori, se non si suppone che la soluzione esista e le condizioni coinvolgono solo l'approssimazione iniziale.

4.3 Il Metodo di Newton-Raphson

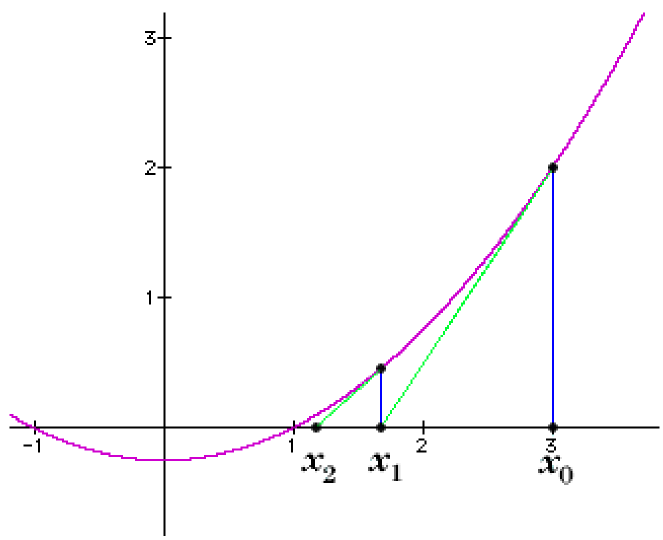
Nel 1669 Newton (1643-1727) propone un metodo per determinare un'approssimazione degli zeri di un'equazione algebrica non lineare. La sua idea originaria è considerevolmente più complicata e davvero molto distante dalla procedura che oggi porta il suo nome. La versione di Newton , infatti, più che riflettere un grande interesse per la soluzione delle equazioni non lineari tramite metodi numerici, appare come un espediente per risolvere un problema diverso. Data un'equazione $F(x, y) = 0$, Newton mirava ad esprimere y come serie di potenze di x. A tal fine egli propone nel manoscritto "Methodus fluxionum et serierum infinitarum" un esempio numerico del metodo, utilizzando un'equazione cubica. Al di là dello scopo che spinse Newton all'elaborazione del suo metodo, resta comunque il fatto che i suoi risultati realizzati nel 1669, furono pubblicati molto più tardi. Vent'anni dopo, senza essere mai venuto a conoscenza del lavoro di Newton, Raphson (1648-1715) propone una versione più semplice dell'algoritmo di Newton per la risoluzione di un'equazione algebrica. Tuttavia sia Newton che Raphson hanno elaborato indipendentemente procedure eqivalenti ma puramente algebriche, senza menzionare le derivate. Il metodo, passato alla storia come metodo di Newton-Raphson, appare per la prima volta nella forma in cui lo utilizziamo oggi soltanto nel 1831 in un'opera di Fourier.

Nel caso unidimensionale, il Metodo di Newton utilizza l'idea che localmente una funzione può essere approssimata dalla sua retta tangente.

Data una funzione non lineare $F : D \subseteq \mathbb{R} \to \mathbb{R}$ differenziabile con continuità in un aperto $D_0 \subseteq D$, partendo da un'approssimazione iniziale x_0 fissata in D_0 si approssima la F con la sua retta tangente in x_0, quindi lo zero dovrebbe trovarsi vicino al punto in cui la retta tangente interseca le ascisse. Tale punto x_1 dovrebbe essere un'approssimazione dello zero della F migliore di quanto non fosse x_0. Iterando tale procedimento si ha il seguente schema iterativo:

$$x_0, \quad x_n = x_{n-1} - \left[F'\left(x_{n-1}\right)\right]^{-1} F\left(x_{n-1}\right), \forall n \in \mathbb{N} .$$

L'iterata x_{n+1} è individuata dal punto di incontro dell'asse delle ascisse con la tangente alla curva $y = F(x)$ nel punto $(x_n, F(x_n))$, perciò si usa anche la denominazione di metodo delle tangenti per questa procedura numerica. Questa interpretazione geometrica fu introdotta per la prima volta da Mourraille nel 1768.



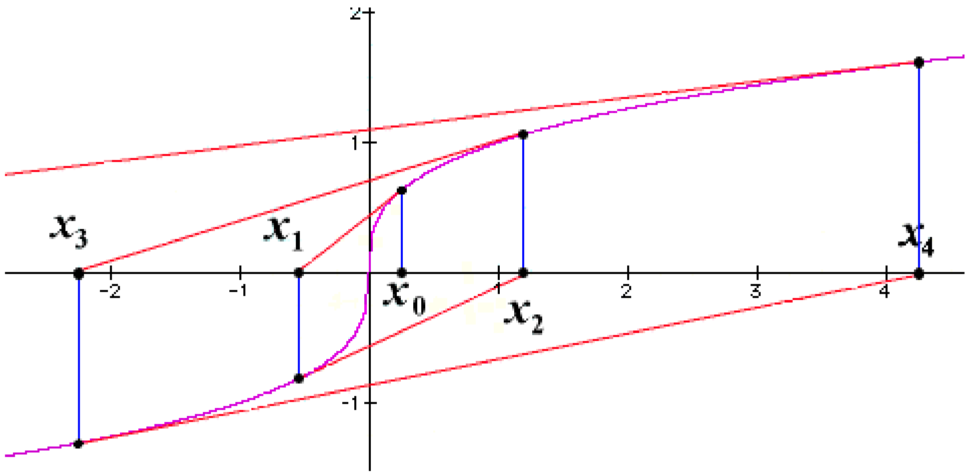
E' noto che lo svantaggio di questo metodo è che l'approssimazione iniziale deve essere scelta in un intorno sufficientemente piccolo della soluzione esatta allo scopo di garantire la convergenza. Già Newton nella sua prima opera al riguardo sottolinea che bisogna essere certi che la stima di partenza differisca dalla radice richiesta meno della sua decima parte, naturalmente una stima molto grossolana ma che comunque ci consente di comprendere come il problema della scelta di x_0 avesse già destato il suo interesse. Riguardo la convergenza del metodo di Newton nel caso unidimensionale, ci limiteremo a enunciare il seguente teorema che fornisce condizioni sufficienti alla convergenza nell'intervallo in cui consideriamo sia compresa la soluzione esatta dell'equazione.

Teorema 4.3.1 *Sia $F : [a, b] \subseteq \mathbb{R} \to \mathbb{R}$ una funzione derivabile due volte con continuità in $[a, b]$ con $F(a)F(b) < 0$ e si supponga:*

1. $F'(x) \neq 0 \ \forall x \in [a, b]$
2. $F''(x) \geq 0$ *oppure* $F''(x) \leq 0 \ \forall x \in [a, b]$
3. $\left|\frac{F(a)}{F'(a)}\right| \leq b - a$ *e* $\left|\frac{F(b)}{F'(b)}\right| \leq b - a$

allora $F(x)$ ha un solo zero α in $[a, b]$ ed il metodo di Newton converge ad α per ogni $x_0 \in [a, b]$.

Osservazione 4.3.2 *Osserviamo che le ipotesi 1 e 2 del teorema sono soddisfatte da tutte le fuzioni monotone concave o convesse su $[a, b]$. E' importante sottolineare che se la curva cambia concavità nell'intervallo al quale appartiene lo zero della funzione, il metodo di Newton non sempre converge, come si vede in questo esempio, in cui si entra in un ciclo infinito, perchè le approssimazioni si allontanano sempre di più dalla soluzione esatta che è l'origine.*



4.4 Il principio di maggiorazione

Nel 1951 Kantorovich stabilisce un importante risultato della teoria dei punti fissi: il principio di maggiorazione. Sulla base di questo principio egli semplifica e perfeziona la dimostrazione del teorema di convergenza a posteriori del metodo di Newton che già aveva realizzato tre anni prima. Il principio di maggiorazione ha segnato una svolta nella teoria dell'iterazione, infatti molti matematici se ne sono serviti per stabilire teoremi di convergenza per il metodo di Newton e per altri metodi iterativi da questo derivanti. Ad esempio Rheinboldt applicò il principio di maggiorazione di Kantorovich per dimostrare un teorema di convergenza per un metodo Newton-like da lui introdotto nel 1968. Questo risultato fu poi generalizzato da Dennis nel 1970 e perfezionato da Miel e Moiret in seguito.

Introduciamo quindi il principio di maggiorazione (teorema 4.4.3) e una sua applicazione utile nel seguito.

Definizione 4.4.1 *Sia* $T : B(x_0, R) \subset X \to X$ *non lineare e Frèchet differenziabile e* $d : [r_0, r_0 + r] \to \mathbb{R}$ *con* $r < R$ *una funzione derivabile tale che:*

$$\|Tx_0 - x_0\| \leq d(r_0) - r_0 \tag{4.2}$$

$$\|T'x\|_{End(X)} \leq d'(r) \quad \forall x \in X : \|x - x_0\| \leq r - r_0$$

allora si dice che l'equazione:

$$d(r) = r \tag{4.3}$$

maggiora

$$Tx = x \tag{4.4}$$

oppure che la funzione d *maggiora l'operatore* T.

Lemma 4.4.2 *Siano* $(x_n)_{n\in\mathbb{N}}$ *una successione a valori in* X *e* $(r_n)_{n\in\mathbb{N}}$ *una successione reale non negativa convergente a* r_*.

Se $(r_n)_{n\in\mathbb{N}}$ *maggiora* $(x_n)_{n\in\mathbb{N}}$, *cioè se:*

$$\|x_n - x_{n-1}\| \leq r_n - r_{n-1} \quad \forall n \in \mathbb{N} \tag{4.5}$$

Allora $\exists\, x_* \in X$ *tale che* $(x_n)_{n\in\mathbb{N}}$ *converge a* x_* *e vale:*

$$\|x_* - x_n\| \leq r_* - r_n \quad \forall n \in \mathbb{N} \tag{4.6}$$

Dimostrazione. Per ipotesi la successione reale $(r_n)_{n\in\mathbb{N}}$ è convergente a r_* dunque è di Cauchy cioè:

$$\forall \varepsilon > 0 \ \ \exists \upsilon \in \mathbb{N} : |r_{n+p} - r_n| \leq \varepsilon \ \ \forall n \geq \nu \ \forall p > 0$$

Consideriamo ora:

$$\|x_{n+p} - x_n\| = \left\| \sum_{k=0}^{p-1} (x_{n+p-k} - x_{n+p-k-1}) \right\| \le \sum_{k=0}^{p-1} \|x_{n+p-k} - x_{n+p-k-1}\|$$

utilizzando l'ipotesi (4.5)

$$\|x_{n+p} - x_n\| \le \sum_{k=0}^{p-1} \|x_{n+p-k} - x_{n+p-k-1}\| \le \sum_{k=0}^{p-1} (r_{n+p-k} - r_{n+p-k-1}) = r_{n+p} - r_n$$

poichè $(r_n)_{n\in\mathbb{N}}$ è convergente abbiamo allora che:

$$\|x_{n+p} - x_n\| \le r_{n+p} - r_n < \varepsilon$$

quindi $(x_n)_{n\in\mathbb{N}}$ è di Cauchy e poichè X è di Banach, allora $(x_n)_{n\in\mathbb{N}}$ è convergente.

Posto $x_* = \lim_{n\to\infty} x_n$ dalla (4.5) passando al limite per $n \to \infty$ si ottiene la tesi (4.6) ■

Teorema 4.4.3 ***(Principio di maggiorazione)*** *Sia $T : B(x_0, R) \subset X \to X$ non lineare e Frèchet differenziabile con continuità nella palla chiusa di centro x_0 e raggio r, $r < R$ e sia $d : [r_0, r_0 + r] \to \mathbb{R}$ una funzione derivabile su $[r_0, r_0 + r]$ che maggiora l'operatore T. Supponiamo inoltre che d abbia almeno un punto fisso s in $[r_0, r_0 + r]$, posto:*

$$x_0 \ , \ x_n = Tx_{n-1} \quad \forall n \in \mathbb{N} \tag{4.7}$$

$$r_0 \ , \ r_n = d(r_{n-1}) \quad \forall n \in \mathbb{N} \tag{4.8}$$

risulta che:

- l'operatore T ha un punto fisso x_*

- $(r_n)_{n\in\mathbb{N}}$ maggiora $(x_n)_{n\in\mathbb{N}}$

-$(x_n)_{n\in\mathbb{N}}$ converge a x_*e vale la (4.6)

Dimostrazione. Prima di tutto dimostriamo che $(r_n)_{n\in\mathbb{N}}$ converge al più piccolo punto fisso della funzione d in $[r_0, r_0 + r]$.

Per ipotesi d maggiora T, quindi dalla definizione (4.2) si ha che:

$$0 \leq \left\|T'x\right\|_{End(X)} \leq d'(r) \quad \forall x \in X : \|x - x_0\| \leq r - r_0,$$

ciò implica che d è crescente su $[r_0, r_0 + r]$. Allora la successione $(r_n)_{n\in\mathbb{N}}$ definita nella (4.8) è ben definita e crescente.

Sia $s \in [r_0, r_0 + r]$ un punto fisso di d cioè sia tale che $d(s) = s$. Dimostriamo per induzione su n che:

$$r_n \leq s \quad \forall n \in \mathbb{N} \tag{4.9}$$

$r_0 \leq s$ perché $s \in [r_0, r_0 + r]$.

Supponiamo ora che $r_k \leq s$ $\forall k : 0 \leq k \leq n$, allora $r_{n+1} = d(r_n) \leq d(s) = s$.

Essendo r_n una successione crescente e limitata, essa converge e posto:

$$r_* = \lim_{n\to\infty} d(r_n) = \lim_{n\to\infty} r_{n+1}$$

poiché d è continua si ha:

$$r_* = \lim_{n\to\infty} d(r_n) = d\left(\lim_{n\to\infty} r_n\right) = d(r_*)$$

dunque r_* è punto fisso di d in $[r_0, r_0 + r]$ e dalla (4.9) risulta che r_* è il più piccolo punto fisso in $[r_0, r_0 + r]$.

Verifichiamo ora che x_n sia ben definita e convergente, a tal fine è sufficiente dimostrare che r_n maggiora x_n e che $x_n \in \overline{B(x_0, r)}$.

Procediamo per induzione su n:

Per $n = 1$ si ha:

$$\|x_1 - x_0\| = \|T(x_0) - x_0\| \leq d(r_0) - r_0 = r_1 - r_0 \leq (r + r_0) - r_0 = r.$$

Supponiamo che r_k maggiora x_k e $x_k \in B(x_0, R) \quad \forall k : 1 \leq k \leq n$ mostriamo che r_{n+1} maggiora x_{n+1} e $x_{n+1} \in \overline{B(x_0, r)}$.

Osserviamo che:

$$x_{n+1} - x_n = T(x_n) - T(x_{n-1}) = \int_0^1 T'(x_{n-1} + t(x_n - x_{n-1}))(x_n - x_{n-1})\, dt. \tag{4.10}$$

Consideriamo

$$\|x_{n-1} + t(x_n - x_{n-1}) - x_0\| \leq t\|x_n - x_{n-1}\| + \|x_{n-1} - x_0\|$$

$$= t\|x_n - x_{n-1}\| + \|x_{n-1} - x_{n-2} + x_{n-2} - ... - x_1 + x_1 - x_0\|$$

$$\leq t\|x_n - x_{n-1}\| + \|x_{n-1} - x_{n-2}\| + ... + \|x_1 - x_0\|$$

per ipotesi induttiva

$$\|x_{n-1} + t(x_n - x_{n-1}) - x_0\| \leq t(r_n - r_{n-1}) + (r_{n-1} - r_{n-2}) + ... + (r_1 - r_0) \tag{4.11}$$

$$= t(r_n - r_{n-1}) + r_{n-1} - r_0$$

Poiché per ipotesi d maggiora T dalla (4.2) segue che:

$$\left\|T'(x_{n-1} + t(x_n - x_{n-1}) - x_0)\right\| \leq d'(t(r_n - r_{n-1}) + r_{n-1}) \ , \tag{4.12}$$

quindi dalla (4.10) e dalla (4.12) otteniamo:

$$\begin{aligned}\|x_{n+1}-x_n\| &\leq \int_0^1 \left\|T'\left(x_{n-1}+t\left(x_n-x_{n-1}\right)\right)\right\| \|\left(x_n-x_{n-1}\right)\| \, dt \\ &\leq \int_0^1 d'\left(t\left(r_n-r_{n-1}\right)+r_{n-1}\right)\left(r_n-r_{n-1}\right) dt \\ &= d\left(t\left(r_n-r_{n-1}\right)+r_{n-1}\right)\Big]_{t=0}^{t=1} \\ &= d\left(r_n\right)-d\left(r_{n-1}\right)=r_{n+1}-r_n\end{aligned}$$

e inoltre :

$$\begin{aligned}\|x_{n+1}-x_0\| &= \|x_{n+1}-x_n+x_n-...-x_1+x_1-x_0\| \\ &\leq \|x_{n+1}-x_n\|+\|x_n-x_{n-1}\|+..+\|x_1-x_0\| \\ &\leq r_{n+1}-r_n+r_n-r_{n-1}+r_{n-1}+...+r_1-r_0 \\ &= r_{n+1}-r_0 \leq r_0+r-r_0=r\end{aligned}$$

ciò equivale a dire che $x_{n+1} \in \overline{B\left(x_0, r\right)}$.

Allora poiché $\left(x_n\right)_{n\in\mathbb{N}}$ è ben definita in $\overline{B\left(x_0, r\right)}$ e $\left(r_n\right)_{n\in\mathbb{N}}$ maggiora $\left(x_n\right)_{n\in\mathbb{N}}$, per il lemma 4.4.2 $\left(x_n\right)_{n\in\mathbb{N}}$ converge e detto $x_* = \lim_{n\to\infty} x_n$ si ha la (4.6). Inoltre per la continuità di T in $\overline{B\left(x_0, r\right)}$ si ha:

$$x_* = \lim_{n\to\infty} x_n = \lim_{n\to\infty} T(x_{n-1}) = T\left(\lim_{n\to\infty} x_{n-1}\right) = T\left(x_*\right)$$

cioè x_* è punto fisso di T. ■

Teorema 4.4.4 *Siano vere le ipotesi del teorema 4.4.3. Supponiamo che:*

$$d\left(r_0+r\right) \leq r_0+r \tag{4.13}$$

e che d ammetta un unico punto fisso r_*. Allora:

- T ammette un unico punto fisso x_*

- la successione delle approssimazioni successive di T converge a x_*,qualsiasi sia il punto iniziale $\widetilde{x} \in B(x_0\ ,r)$.

Dimostrazione. Poichè valgono le ipotesi del 4.4.3, per lo stesso teorema si ha che la successione (4.7) delle approssimazioni successive di T converge a x_* che è punto fisso per T. Consideriamo ora la successione $(s_n)_{n\in\mathbb{N}}$ delle approssimazioni successive di d, di punto iniziale $s_0 = r_0 + r$, cioè :

$$s_0\ ,\ s_n = d(s_{n-1}) \quad \forall n \in \mathbb{N} \tag{4.14}$$

Verifichiamo per induzione su $n$che $(s_n)_{n\in\mathbb{N}}$ è monotona decrescente.

Per $n = 1$ è vera, poiché utilizzando l'ipotesi (4.13) abbiamo che:

$$s_1 = d(s_0) = d(r_0 + r) \leq r_0 + r = s_0.$$

Supponiamo che $s_k \leq s_{k-1}\ \forall k : 1 \leq k \leq n$ dimostriamo che : $s_{n+1} \leq s_n$.

Utilizzando la crescenza di d:

$$s_{n+1} = d(s_n) \leq d(s_{n-1}) = s_n.$$

Sia $s \in [r_0, r_0 + r]$ un punto fisso di d, mostriamo che $(s_n)_{n\in\mathbb{N}}$ è limitata inferiormente da s. Procedendo ancora per induzione su n.

Per $n = 0$ è vera, poiché

$$s \leq r_0 + r = s_0.$$

Supponiamo che $s \leq s_k \ \forall k : 0 \leq k \leq n$. Quindi $s \leq s_n$, da cui essendo d crescente si ha che $d(s) \leq d(s_n)$, allora:

$$s = d(s) \leq d(s_n) = s_{n+1}.$$

Giacché $(s_n)_{n\in\mathbb{N}}$ è monotona decrescente limitata inferiormente da $s \in [r_0, r_0 + r]$, $(s_n)_{n\in\mathbb{N}}$ converge all'unico punto fisso r_* di d.

Entrambe le successioni (4.8) e (4.14) convergono, dunque a r_*.

Sia ora

$$y_0 \ , \ y_n = T(y_{n-1}) \quad \forall n \in \mathbb{N} \tag{4.15}$$

con $x_0 \neq y_0$ e diamo prova per induzione che

$$y_n \in B(x_0, r) \tag{4.16}$$

$$\|y_n - x_n\| \leq s_n - r_n. \tag{4.17}$$

Verifichiamo che la (4.17) valga per $n = 1$

$$\|y_1 - x_1\| = \|T(y_0) - T(x_0)\| = \left\| \int_0^1 T'(x_0 + t(y_0 - x_0))(y_0 - x_0)\, dt \right\|$$

poiché d maggiora T abbiamo che:

$$\|y_1 - x_1\| \leq \int_0^1 d'(t(s_0 - r_0) + r_0)(s_0 - r_0)\, dt = d(s_0) - d(r_0) = s_1 - r_1.$$

Inoltre:

$$\|y_1 - x_0\| = \|y_1 - x_1 + x_1 - x_0\| \leq \|y_1 - x_1\| + \|x_1 - x_0\|$$

$$\leq (s_1 - r_1) + (r_1 - r_0) = s_1 - r_0 \leq s_0 - r_0 = r$$

cioè vale anche la (4.16).

Supponiamo che la (4.17) e la (4.16) siano vere per k tale che $1 \leq k \leq n$ allora:

$$\|y_{n+1} - x_{n+1}\| = \|T(y_n) - T(x_n)\| = \int_0^1 \left\|T'(x_n + t(y_n - x_n))\right\| \|y_n - x_n\| \, dt$$

poiché d maggiora T si ha:

$$\|y_{n+1} - x_{n+1}\| \leq \int_0^1 d'(t(s_n - r_n) + r_n)(s_n - r_n) \, dt = d(s_n) - d(r_n) = s_{n+1} - r_{n+1}.$$

Inoltre

$$\|y_{n+1} - x_0\| = \|y_{n+1} - x_{n+1} + x_{n+1} - x_0\| \leq \|y_{n+1} - x_{n+1}\| + \|x_{n+1} - x_0\| \leq$$

$$\leq (s_{n+1} - r_{n+1}) + (r_{n+1} - r_0) = s_{n+1} - r_0 \leq s_0 - r_0 = r \ .$$

Poiché $(s_n)_{n\in\mathbb{N}}$ e $(r_n)_{n\in\mathbb{N}}$ convergono entrambe a r_* dalla (4.17) segue che anche $(y_n)_{n\in\mathbb{N}}$ converge a x_* .

Da qui, abbiamo l'unicità del punto fisso di T qualunque sia il punto iniziale $\widetilde{x} \in B(x_0\,, r)$. Infatti se $\widetilde{x} \neq x_*$ è punto fisso di T, scelto proprio $\widetilde{x}$ come punto iniziale, la successione delle approsssimazioni successive di T è costantemente uguale a $\widetilde{x}$, ma per quello che abbiamo appena dimostrato essa converge a x_* , quindi x_* è unico. ■

4.5 Il Metodo di Newton-Kantorovich

Dopo Newton e Raphson, molti altri matematici si interessarono a questo metodo ottenendo interessanti miglioramenti e generalizzazioni. Tra di essi citiamo Mourraille, che nel 1768 pose il problema fondamentale della scelta dell'approssimazione iniziale, indispensabile per garantire la convergenza del metodo e introduce per la prima volta la familiare interpretazione geometrica nel caso reale. Nel 1818 Fourier, sotto opportune ipotesi, dimostra che la convergenza del metodo in $\mathbb{R}$ è almeno quadratica . Nel 1829 Cauchy è il primo ad ottenere un risultato di convergenza a posteriori stabilendo condizioni sufficienti a garantire la qudraticità della convergenza. Successivamente il suo risultato fu esteso da Ostrowski nel 1940, mentre Fine nel 1916 dimostra la convergenza a posteriori del metodo in $\mathbb{R}^k$o $\mathbb{C}^k$.

Infine Kantorovich pubblica nel 1939 "The method of successive approximation for functional analysis", sviluppando la sua teoria sui metodi iterativi per equazioni funzionali negli spazi di Banach e gettando le basi per un teorema di convergenza del metodo di Newton.

Consideriamo $F : D \subseteq X \to Y$ operatore non lineare Frèchet differenziabile in un aperto convesso $D_0 \subseteq D$. Il metodo di Newton-Kantorovich è dato dalla seguente:

$$x_0, \quad x_n = x_{n-1} - \left[F'(x_{n-1})\right]^{-1} F(x_{n-1}), \forall n \in \mathbb{N} \tag{4.18}$$

dove x_0 è un'approssimazione iniziale fissata in D_0.

Questo metodo iterativo è ottenuto in pratica linearizzando l'equazione

$$F(x) = 0. \tag{4.19}$$

In altre parole si approssima $F(x)$ con lo sviluppo di Taylor del primo ordine con centro x_0:

$$F(x) \approx L_n(x) = F(x_n) + F'(x_n)(x - x_n) \tag{4.20}$$

e poi si risolve l'equazione risultante $L_n(x) = 0$. Lo zero di $L_n(x)$ fornisce l'approssimazione successiva x_{n+1}. Ma è nel 1948 che il matematico russo, in [7], stabilisce il teorema di convergenza a posteriori per il metodo di Newton negli spazi di Banach, conosciuto anche come teorema di Newton-Kantorovich. Questo risultato che è una generalizzazione e un miglioramento del teorema di convergenza realizzato da Cauchy nel 1829, come già detto precedentemente sarà perfezionato tre anni dopo dallo stesso Kantorovich con l'introduzione del suo principio di maggiorazione.

Teorema 4.5.1 ***(di Newton-Kantorovich)*** *Sia $F : D \subseteq X \to Y$ un operatore non lineare tra spazi di Banach, Frèchet differenziabile due volte in un aperto convesso $D_0 \subseteq D$. Supponiamo che :*

- $\|F''(x)\|_{L(X,L(X,Y))} \leq k, \ \ \forall x \in D_0$
- *per $x_0 \in D_0$ fissato sia $F(x_0) \neq 0$ e $F'(x_0)$ invertibile.*

Assumiamo inoltre :

- $a := \left\| [F'(x_0)]^{-1} F(x_0) \right\|$
- $b := \left\| [F'(x_0)]^{-1} \right\|_{L(Y,X)}$
- $\xi := abk$

Se $2\xi \leq 1$, allora le iterate (4.18) del metodo di Newton sono ben definite e posto:

$$r_* := \frac{1-\sqrt{1-2\xi}}{bk} \quad , \quad r_{**} := \frac{1+\sqrt{1-2\xi}}{bk}$$

si ha che $x_n \in B(x_0, r_*)$ e che $(x_n)_{n\in\mathbb{N}}$ converge ad una soluzione x_* della (4.19).

- *Se $2\xi < 1$, la soluzione x_* è unica in $B(x_0, r_{**})$.*
- *Se $2\xi = 1$, la soluzione x_* è unica in $\overline{B(x_0, r_{**})}$.*

Inoltre vale la seguente stima dell'errore:

$$\|x_* - x_n\| \leq \frac{2a_n}{1+\sqrt{1-2\xi_n}} \leq 2^{1-n} (2\xi)^{2^n-1} a, \ \forall n \in \mathbb{N} \tag{4.21}$$

dove:

$$b_0 = b \ , \ a_0 = a \ , \ \xi_0 = \xi$$

$$b_n = \frac{b_{n-1}}{1-\xi_{n-1}} \ , \ a_n = \frac{a_{n-1}\xi_{n-1}}{2\left(1-\xi_{n-1}\right)} \ , \ \xi_n = a_n b_n k \ , \ \forall n \in \mathbb{N}.$$

Osserviamo che mentre nel teorema 4.3.1, la convergenza nel caso unidimensionale ha come ipotesi fondamentale che la funzione non cambi la sua concavità su tutto l'intervallo considerato, Kantorovich si libera di questa condizione imponendo però la limitatezza della derivata seconda.

Proponiamo ora la dimostrazione più semplice costruita da Kantorovich, cioè quella che utilizza il principio di maggiorazione (cfr [8] e [7]).Prima di tutto egli definisce una funzione, che indicheremo nel seguito con ϕ, le cui approssimazioni di Newton r_n costituiscono una successione che maggiora le x_n definite nella (4.18).

$$\phi\left(r\right) = \frac{1}{2}bkr^2 - r + a \tag{4.22}$$

$$r_0 = 0 \ , \ r_n = r_{n-1} - \frac{\phi\left(r_{n-1}\right)}{\phi'\left(r_{n-1}\right)} \ , \ \forall n \in \mathbb{N} \tag{4.23}$$

Osservazione 4.5.2 *Nell'ipotesi $2\xi \leq 1$, la funzione ϕ ammette un minimo non negativo con $r_{\min} = (bk)^{-1}$.*

Infatti tenendo presente che $\xi = abk$,dall'ipotesi $2\xi \leq 1$ segue che

$$\frac{1}{bk} \geq 2a > 0 \tag{4.24}$$

Ora essendo $\phi'\left(r\right) = bkr - 1 > 0 \Leftrightarrow r > \frac{1}{bk}$, si ha che $r_{\min} = (bk)^{-1}$ .Quindi per la (4.24) si ha $r_{\min} > 0$.

Essendo inoltre $\phi\left(0\right) = a > 0$ e $\lim_{r\to\infty}\phi\left(r\right) = +\infty$ esistono due zeri di ϕ e sono:

$$r_* := \frac{1-\sqrt{1-2\xi}}{bk} \quad , \quad r_{**} := \frac{1+\sqrt{1-2\xi}}{bk}$$

Ovviamente:

-se $2\xi < 1$*, allora* $r_* < r_{\min} < r_{**}$

-se $2\xi = 1$*, allora* $r_* = r_{\min} = r_{**}$

Suddividiamo la dimostrazione del teorema 4.5.1 in tre parti.

Dimostrazione. 1)ESISTENZA DI UNA SOLUZIONE E CONVERGENZA DEL METODO DI NEWTON MODIFICATO

Consideriamo l'operatore di Goursat T definito da:

$$T(x) = x - \left[F'(x_0)\right]^{-1} F(x)$$

e la funzione derivabile d :

$$d(r) = \phi(r) + r \tag{4.25}$$

dove ϕ è la funzione definita in (4.22).

Osserviamo che d maggiora T ,cioè che è verificata la (4.2):

- $\|T(x_0) - x_0\| = \left\|x_0 - [F'(x_0)]^{-1} F(x_0) - x_0\right\| = \left\|[F'(x_0)]^{-1} F(x_0)\right\|$

$= a = \phi(0) - 0 = d(r_0) - r_0$

- $\|T'(x)\| = \left\|I - [F'(x_0)]^{-1} F'(x)\right\| = \left\|[F'(x_0)]^{-1} [F'(x_0) - F'(x)]\right\|$

$\leq \left\|[F'(x_0)]^{-1}\right\| \|F'(x_0) - F'(x)\| = b \|F'(x_0) - F'(x)\|$

tenendo conto che $F''(x_0) = \lim\limits_{x \to x_0} \frac{F'(x) - F'(x_0)}{x - x_0}$

$\|T'(x)\| = b\|F'(x_0) - F'(x)\| \leq b\|F''(x_0)\| \|x - x_0\| \leq bk\|x - x_0\|$

$\leq bkr = d'(r)\ ,\ \ \forall x : \|x - x_0\| \leq r - r_0 = r$

Notiamo ancora che dall'osservazione 4.5.2 e dalla (4.25) segue:

$d(r_*) = \phi(r_*) + r_* = r_*$

$d(r_{**}) = \phi(r_{**}) + r_{**} = r_{**}$

cioè che r_* e r_{**}sono punti fissi di d, distinti se $2\xi < 1$, uguali e coincidenti se $2\xi = 1$. Sono soddisfatte, dunque, tutte le ipotesi del principio di maggiorazione, da cui segue che anche l'operatore T ammette un punto fisso x_*, che è proprio la soluzione della (4.19). Ciò significa che il metodo di Newton modificato:

$$x_0 \ , \ x_n = x_{n-1} - \left[F'(x_0)\right]^{-1} F(x_{n-1}) \, , \forall n \in \mathbb{N} \tag{4.26}$$

converge a x_*.

2)UNICITA' DELLA SOLUZIONE

In $[r_*, r_{**})$ la funzione d ha un unico punto fisso r_* e $d(r) \leq r$ allora per il teorema 4.4.4, T ammette un unico punto fisso x_* nella palla $B(x_0, r)$. Dunque la soluzione della equazione (4.19) esiste ed è unica.

3)CONVERGENZA DELLE APPROSSIMAZIONI DEFINITE IN (4.18)

Consideriamo la successione $(r_n)_{n\in\mathbb{N}}$ definita in (4.23) che è una successione crescente e convergente a r_* Prendiamo ora in esame la successione $(x_n)_{n\in\mathbb{N}}$ definita nella (4.18). x_1 è ben definito e si ha:

$$\begin{aligned} \|x_1 - x_0\| &= \left\|x_0 - \left[F'(x_0)\right]^{-1} F(x_0) - x_0\right\| \\ &= \left\|\left[F'(x_0)\right]^{-1} F(x_0)\right\| = a = r_1 - r_0 = r_1 \leq r_* \end{aligned} \tag{4.27}$$

x_2 è ben definito se e solo se $F'(x_1)$ è invertibile. Verifichiamolo:

$$\left\|I - \left[F'(x_0)\right]^{-1} F'(x_1)\right\| = \left\|\left[F'(x_0)\right]^{-1}\left[F'(x_0) - F'(x_1)\right]\right\| \leq \left\|\left[F'(x_0)\right]^{-1}\right\| \left\|\left[F'(x_0) - F'(x_1)\right]\right\|$$

per il corollario 3.4.4 dalla (3.10):

$$\begin{aligned}
\left\|I - \left[F'(x_0)\right]^{-1} F'(x_1)\right\| &\le \left\|\left[F'(x_0)\right]^{-1}\right\| \left\| \left[F'(x_0) - F'(x_1)\right] \right\| \\
&= \left\|\left[F'(x_0)\right]^{-1}\right\| \left\| \int_0^1 F''(x_1 + t(x_0 - x_1))(x_1 - x_0)\, dt \right\| \\
&\le b \int_0^1 \left\|F''(x_1 + t(x_0 - x_1))\right\| dt \left\|x_1 - x_0\right\| \\
&\le b \int_0^1 k dt \left\|x_1 - x_0\right\| = bk \left\|x_1 - x_0\right\| = bkr_1 \le bkr_* \\
&= bk \left(\frac{1 - \sqrt{1 - 2\xi}}{bk}\right) = 1 - \sqrt{1 - 2\xi}
\end{aligned}$$

poiché per ipotesi $2\xi \le 1$ abbiamo:

$$\left\|I - \left[F'(x_0)\right]^{-1} F'(x_1)\right\| \le 1 \tag{4.28}$$

Dal corollario 2.4.3 e dalla (4.28), si ha che l'operatore

$I - \left(I - [F'(x_0)]^{-1} F'(x_1)\right) = [F'(x_0)]^{-1} F'(x_1)$ è invertibile e quindi $F'(x_1)$ è invertibile e vale la (2.14):

$$\begin{aligned}
\left\|\left\{\left[F'(x_0)\right]^{-1} F'(x_1)\right\}^{-1}\right\| &= \left\|\left[F'(x_1)\right]^{-1} F'(x_0)\right\| \\
&\le \frac{1}{1 - \left\|I - [F'(x_0)]^{-1} F'(x_1)\right\|} \\
&\le \frac{1}{1 - bkr_1} = \frac{\phi'(0)}{\phi'(r_1)} = \frac{\phi'(r_0)}{\phi'(r_1)}
\end{aligned}$$

Si ha quindi:

$$\left\|\left\{\left[F'(x_0)\right]^{-1} F'(x_1)\right\}^{-1}\right\| \le \frac{1}{1 - bkr_1} = \frac{\phi'(r_0)}{\phi'(r_1)} \tag{4.29}$$

cioè $x_2 = x_1 - [F'(x_1)]^{-1} F'(x_1)$.

Vogliamo dimostrare che $\|x_2 - x_1\| \leq r_2 - r_1$

$$\begin{aligned} \|x_2 - x_1\| &= \left\| \left[F'(x_1)\right]^{-1} F'(x_1) \right\| = \left\| \left[F'(x_1)\right]^{-1} F'(x_0) \left[F'(x_0)\right]^{-1} F'(x_1) \right\| \\ &\leq \left\| \left[F'(x_1)\right]^{-1} F'(x_0) \right\| \left\| \left[F'(x_0)\right]^{-1} F'(x_1) \right\| . \end{aligned} \tag{4.30}$$

Consideriamo la formula di Taylor di centro x_0 arrestata al primo ordine :

$$F(x_1) = F(x_0) + F'(x_0)(x_1 - x_0) + \int_0^1 F''(tx_1 + (1-t)x_0)(x_1 - x_0, (1-t)(x_1 - x_0))\, dt$$

moltiplichiamola per $[F'(x_0)]^{-1}$:

$$[F'(x_0)]^{-1} F(x_1) = [F'(x_0)]^{-1} F(x_0) + (x_1 - x_0) +$$

$$+ \int_0^1 [F'(x_0)]^{-1} F''(tx_1 + (1-t)x_0)(x_1 - x_0, (1-t)(x_1 - x_0))\, dt =$$

osservando che $x_0 - x_1 = [F'(x_0)]^{-1} F(x_0)$ si ha che:

$$\left[F'(x_0)\right]^{-1} F(x_1) = \int_0^1 \left[F'(x_0)\right]^{-1} F''(tx_1 + (1-t)x_0)(x_1 - x_0, (1-t)(x_1 - x_0))\, dt \tag{4.31}$$

Passando alle norme,tenendo conto della (4.27) e delle ipotesi del teorema, dalla (4.31) si ottiene che:

$$\left\| \left[F'(x_0)\right]^{-1} F(x_1) \right\| \leq bk(r_1)^2 \int_0^1 (1-t)dt = \frac{bk(r_1)^2}{2}. \tag{4.32}$$

Allora dalla (4.30) si ha per la (4.32) e per la (4.29) che:

$$\|x_2 - x_1\| \leq \frac{bk(r_1)^2}{2} \cdot \frac{1}{1 - bkr_1} = -\frac{\phi(r_1)}{\phi'(r_1)} = r_2 - r_1$$

e inoltre

$$\|x_2 - x_0\| \leq \|x_2 - x_1\| + \|x_1 - x_0\| \leq r_2 - r_1 + r_1 - r = r_2 \leq r_*.$$

Analogamente riusciamo a dimostrare che r_n maggiora x_n e che $x_n \in B(x_0, r_n)$, quindi che $(x_n)_{n\in\mathbb{N}}$ converge a x_*. ■

Osservazione 4.5.3 *Le stime dell'errore (4.21) del teorema di Kantorovich e le (4.6) del principio di maggiorazione sono difatti coincidenti come provato nel 1986 da Yamamoto (cfr. [14]). Infatti egli dimostra che:*

$$r_{n+1} - r_n = a_n$$

$$r_* - r_n = \frac{1 - \sqrt{1 - 2\xi_n}}{b_n k} = \frac{2a_n}{1 + \sqrt{1 - 2\xi_n}}$$

$$r_{**} - r_n = \frac{1 + \sqrt{1 - 2\xi_n}}{b_n k}$$

dalle quali sostituendo nella (4.6) si ha la (4.21).

Nel 1955 Bartle sostituisce, in modo naturale, l'ipotesi $\|F''(x)\|_{L(X,L(X,Y))} \leq k, \ \forall x \in D_0$ con la condizione più debole:

$$\left\|F'(x) - F'(x_0)\right\| \leq k \|x - x_0\|, \forall x \in D_0 \tag{4.33}$$

Riportiamo di seguito la nuova dimostrazione del teorema 4.5.1 data da Bartle in (cfr [1]), in quanto questa tecnica dimostrativa è ancora oggi una delle più diffuse.

Dimostrazione. Osserviamo che $\forall x \in D_0$ tale che :

$$\|x - x_0\| \leq \frac{1}{k \left\|[F'(x_0)]^{-1}\right\|} = \frac{1}{bk}$$

dall'ipotesi (4.33) si ha che:

$$\left\|F'(x) - F'(x_0)\right\| \leq \frac{1}{bk} \tag{4.34}$$

Essendo per ipotesi $F'(x_0)$ invertibile e poiché vale la (4.34), allora per il lemma di Banach (Lemma 2.4.4) si ha che anche $F'(x)$ è invertibile $\forall x \in D_0$ tale che $\|x - x_0\| \leq$

$\frac{1}{bk}$ e che:

$$\left\| \left[F'(x)\right]^{-1} \right\| \leq \frac{\left\| [F'(x_0)]^{-1} \right\|}{1 - \left\| [F'(x_0)]^{-1} \right\| \|F'(x_0) - F'(x)\|} \leq \frac{b}{1 - bk\|x - x_0\|}. \tag{4.35}$$

Verifichiamo ora per induzione su n che le approssimazioni di Newton (4.18) sono ben definite e che $(x_n)_{n\in\mathbb{N}}$ è maggiorata da $(r_n)_{n\in\mathbb{N}}$ definita in (4.23).

Per $n = 1$ la tesi è vera infatti:

$$\|x_1 - x_0\| = \left\| [F'(x_0)]^{-1} F(x_0) \right\| = a = r_1 \leq r_1 - r_0.$$

Supponiamo ora che la (4.5) valga $\forall k : 1 \leq k \leq n$, dimostriamo che:

$$\|x_{n+1} - x_n\| \leq r_{n+1} - r_n \tag{4.36}$$

Osserviamo che dalla definizione (4.18) si ha che:

$$F(x_{n-1}) = F'(x_{n-1})(x_{n-1} - x_n) \tag{4.37}$$

Allora:

$$\|x_{n+1} - x_n\| = \left\| \left[F'(x_n)\right]^{-1} F(x_n) \right\| = \left\| \left[F'(x_n)\right]^{-1} [F(x_n) - F(x_{n-1}) + F(x_{n-1})] \right\|$$

per la (4.37) e per la (4.35) si ha:

$$\begin{aligned} \|x_{n+1} - x_n\| &= \left\| \left[F'(x_n)\right]^{-1} \left[F(x_n) - F(x_{n-1}) - F'(x_{n-1})(x_n - x_{n-1})\right] \right\| \\ &\leq \frac{b}{1 - bk\|x_n - x_0\|} \left\| F(x_n) - F(x_{n-1}) - F'(x_{n-1})(x_n - x_{n-1}) \right\|. \end{aligned}$$

Usando (3.10) del corollario della formula di Taylor:

$$\|x_{n+1}-x_n\| \leq \frac{b}{1-bk\,\|x_n-x_0\|}\left\|F\left(x_n\right)-F\left(x_{n-1}\right)-F'\left(x_{n-1}\right)\left(x_n-x_{n-1}\right)\right\|$$

$$=\frac{b}{1-bk\,\|x_n-x_0\|}\left\|\int_0^1 F'\left(x_{n-1}+t(x_n-x_{n-1}\right)dt\left(x_n-x_{n-1}\right)-F'\left(x_{n-1}\right)\left(x_n-x_{n-1}\right)\right\|$$

$$\leq \frac{b}{1-bk\,\|x_n-x_0\|}\int_0^1\left\|F'\left(x_{n-1}+t(x_n-x_{n-1}\right)-F'\left(x_{n-1}\right)\right\|\|x_n-x_{n-1}\|\,dt.$$

Per la (4.33):

$$\|x_{n+1}-x_n\| \leq \frac{b}{1-bk\,\|x_n-x_0\|}\int_0^1 k\,\|x_{n-1}+tx_n-tx_{n-1}-x_{n-1}\|\,\|x_n-x_{n-1}\|\,dt$$

$$=\frac{bk}{1-bk\,\|x_n-x_0\|}\int_0^1 t\,\|x_n-x_{n-1}\|^2\,dt=\frac{bk\,\|x_n-x_{n-1}\|^2}{2\left(1-bk\,\|x_n-x_0\|\right)}$$

per l'ipotesi induttiva otteniamo che:

$$\|x_{n+1}-x_n\| \leq \frac{bk\left(r_n-r_{n-1}\right)^2}{2\left(1-bk\,\|x_n-x_0\|\right)}. \tag{4.38}$$

Osserviamo adesso che usando ancora l'ipotesi induttiva abbiamo che:

$$\|x_n-x_0\|=\|x_n-x_{n-1}+x_{n-1}-..-x_1+x_1-x_0\|\leq\|x_n-x_{n-1}\|+..+\|x_1-x_0\|\leq$$

$$\leq r_n-r_{n-1}+r_{n-1}-..-r_1+r_1-r_0=r_n-r_0=r_n$$

quindi dalla (4.38):

$$\|x_{n+1}-x_n\| \leq \frac{bk\left(r_n-r_{n-1}\right)^2}{2\left(1-bkr_n\right)}. \tag{4.39}$$

Servendoci della definizioni (4.23) mostriamo che:

$$\frac{bk\left(r_n-r_{n-1}\right)^2}{2\left(1-bkr_n\right)}=r_{n+1}-r_n \tag{4.40}$$

$$\begin{aligned}
\frac{bk\left(r_n-r_{n-1}\right)^2}{2} &= \frac{bk\left(r_n\right)^2}{2}+\frac{bk\left(r_{n-1}\right)^2}{2}-bkr_nr_{n-1}\\
&= \frac{bk\left(r_n\right)^2}{2}+\frac{bk\left(r_{n-1}\right)^2}{2}-\left(r_{n-1}-\frac{\phi\left(r_{n-1}\right)}{\phi'\left(r_{n-1}\right)}\right)r_{n-1}bk\\
&= \frac{bk\left(r_n\right)^2}{2}+\frac{bk\left(r_{n-1}\right)^2}{2}-bk\left(r_{n-1}\right)^2+r_{n-1}bk\frac{\phi\left(r_{n-1}\right)}{\phi'\left(r_{n-1}\right)}\\
&= \frac{bk\left(r_n\right)^2}{2}-\frac{bk\left(r_{n-1}\right)^2}{2}+r_{n-1}bk\frac{\phi\left(r_{n-1}\right)}{\phi'\left(r_{n-1}\right)}\\
&= \phi\left(r_n\right)+r_n-a+a-r_{n-1}-\phi\left(r_{n-1}\right)+r_{n-1}bk\frac{\phi\left(r_{n-1}\right)}{\phi'\left(r_{n-1}\right)}\\
&= \phi\left(r_n\right)-\frac{\phi\left(r_{n-1}\right)}{\phi'\left(r_{n-1}\right)}-\phi\left(r_{n-1}\right)+r_{n-1}bk\frac{\phi\left(r_{n-1}\right)}{\phi'\left(r_{n-1}\right)}\\
&= \phi\left(r_n\right)+\frac{\phi\left(r_{n-1}\right)}{\phi'\left(r_{n-1}\right)}\left(bkr_{n-1}-1\right)-\phi\left(r_{n-1}\right)\\
&= \phi\left(r_n\right)+\frac{\phi\left(r_{n-1}\right)}{\phi'\left(r_{n-1}\right)}\phi'\left(r_{n-1}\right)-\phi\left(r_{n-1}\right)=\phi\left(r_n\right)
\end{aligned}$$

Infine la (4.36) segue dalla (4.39). ■

Invece se sostituiamo l'ipotesi $b := \left\|\left[F'\left(x_0\right)\right]^{-1}\right\|_{L(Y,X)}$con l'ipotesi più forte:

$$\left\|\left[F'\left(x_0\right)\right]^{-1}\right\|_{L(Y,X)} \leq b,\ \forall x \in D_0$$

il teorema di Kantorovich vale sotto l'ipotesi più debole $\xi < 2$. Questa versione modificata del teorema fu dimostrata da Mysovskii ed è nota come teorema di Newton-Mysovskii (cfr [10] ,[11]).

Bibliografia

[1] R.G.Bartle, Newton's Method in Banach Spaces, Proc.Am.Math.Soc.6, 1955

[2] H. Brezis, Analisi funzionale - Teoria ed applicazioni, Liguori Editore, Napoli, 1990.

[3] F. Cianciaruso, Il Metodo di Newton-Kantorovich ed alcune sue Applicazioni, Tesi di Dottorato, 2001

[4] W.Gautschi, Numerical Analysis: an introduction, Birkhäuser,Boston, 1997.

[5] M.Giaquinta, G.Modica, Analisi Matematica 3: strutture lineari e metriche, continuità, Pitagora Editrice, Bologna 2000.

[6] G.Gilardi, Analisi tre, McGraw-Hill, Milano, 1994.

[7] L.V.Kantorovich, On Newton's Method for functional equations, Dokl. Akad. Nauk. SSSR 76, 1948 (in russo)

[8] L.V.Kantorovich, The Majorant Principle and Newton's Method, Dokl. Akad. Nauk. SSSR 76, 1951 (in russo)

[9] L.V.Kantorovich,G.P.Akilov, Functional analysis, 2nd ed., Pergamon Press,

Oxford, 1982

[10] I.Mysovskii, On tha convergence of L.V.Kantorovich's method of solution of functional equation and its application, Dokl.Akad.Nauk SSSr 70,1950 (in russo)

[11] J.M.Ortega,W.C.Rheinboldt, Iterative solution of nonlinear equation in several variables, Academic Press, New York, 1970

[12] G.Prodi, A.Ambrosetti, Analisi Non Lineare, I Quaderno, Pubblicazione della classe di Scienze della Scuola Normale di Pisa, 1973.

[13] A.E.Taylor, D.C.Lay, Introduction to functional Analysis,Wiley, New York,1958.

[14] T.Yamamoto, A method for finding sharp error bounds for Newton's Method under the Kantorovich assumption, Numer.Math.49, 1986

[15] T.Yamamoto, Historical developments in convergence analysis for Newton's and Newton's like Methods, Journal Computational.and Applied.Mathemathics 124, 2000