

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

---

SCUOLA DI SCIENZE  
Corso di Laurea in Matematica

**IL METODO DI NEWTON  
E LE SUE VARIANTI  
PER SISTEMI DI EQUAZIONI NON  
LINEARI**

Tesi di Laurea in Analisi Numerica

**Relatore:**  
Chiar.mo Prof.  
VALERIA SIMONCINI

**Presentata da:**  
CHIARA MANZOLI

**Correlatore:**  
Chiar.mo Dr.  
MARGHERITA PORCELLI

Seconda Sessione  
Anno Accademico 2014/2015



# Indice

<b>Introduzione</b>	<b>4</b>
<b>1 Metodi numerici per equazioni non lineari scalari</b>	<b>9</b>
1.1 Il problema . . . . .	9
1.2 Metodo di Newton o metodo delle tangenti . . . . .	9
1.3 Varianti del metodo di Newton . . . . .	14
1.3.1 Metodo di Newton stazionario . . . . .	14
1.3.2 Metodo delle secanti . . . . .	15
1.3.3 Metodo di Newton alle differenze . . . . .	17
1.4 Confronto tra i metodi . . . . .	18
<b>2 Metodi numerici per sistemi di equazioni non lineari</b>	<b>21</b>
2.1 Metodo di Newton . . . . .	23
2.1.1 Costruzione del modello lineare . . . . .	23
2.1.2 Algoritmo . . . . .	24
2.1.3 Analisi di convergenza . . . . .	25
2.2 Metodo di Broyden . . . . .	27
2.2.1 Costruzione del modello lineare . . . . .	27
2.2.2 Algoritmo . . . . .	30
2.2.3 Analisi di convergenza . . . . .	31
<b>3 Implementazione Matlab</b>	<b>39</b>
3.1 Criteri d'arresto . . . . .	39
3.2 Implementazione degli algoritmi . . . . .	41
3.3 Problema test . . . . .	43
3.4 Un esempio applicativo . . . . .	48
<b>Considerazioni finali</b>	<b>49</b>
<b>A</b>	<b>53</b>
A.1 Velocità di convergenza . . . . .	53
A.2 Norme . . . . .	54
A.2.1 Norme vettoriali . . . . .	54
A.2.2 Norme matriciali . . . . .	55

INDICE

---

A.3 Nozioni di analisi per funzioni in più variabili . . . . . 57

# Introduzione

In questa tesi vengono trattati alcuni dei principali metodi numerici per il calcolo di soluzioni di *sistemi non lineari*, di cui le *equazioni non lineari scalari* rappresentano un caso particolare più semplice. In particolare, la tesi è incentrata sul metodo di Newton e su alcune delle sue varianti che rappresentano gli algoritmi più efficaci noti in letteratura per la risoluzione di problemi non lineari.

I sistemi non lineari costituiscono il modello matematico per numerosi problemi di grande interesse applicativo come per esempio problemi di equilibrio chimico o la discretizzazione di problemi differenziali (di tipo ordinario o alle derivate parziali).

L'uso di metodi numerici per problemi non lineari è fondamentale in quanto permette di trovare soluzioni approssimate di sistemi non lineari anche quando la ricerca di soluzioni esatte risulterebbe troppo difficile o addirittura impossibile con gli strumenti dell'analisi matematica. Ad esempio, considerando il caso delle equazioni scalari polinomiali, un "limite" algebrico può essere espresso dal Teorema di Abel-Ruffini, il quale recita

*“Non esiste nessuna formula per le radici di una generica equazione polinomiale di grado quinto (o superiore) in funzione dei coefficienti del polinomio, usando solo le normali operazioni algebriche (addizione, sottrazione, moltiplicazione, divisione) e l'applicazione di radicali (radici quadratiche, cubiche, ecc).”.*

## Il problema e le caratteristiche generali dei metodi

Innanzitutto si considerano i sistemi non lineari composti da una sola equazione, e cioè le equazioni non lineari scalari. Il problema in questo caso consiste in:

Data una funzione non lineare scalare  $f : \mathbb{R} \rightarrow \mathbb{R}$ , lo scopo è quello di trovare uno scalare  $x^* \in \mathbb{R}$  tale che

$$f(x^*) = 0. \tag{1}$$

Il valore  $x^*$  è detto “zero”, “radice” o “soluzione” dell'equazione (1).

Molto spesso però ci si trova in situazioni più complicate in cui, invece di dover soddisfare una sola equazione in una incognita, devono essere soddisfatte simultaneamente  $n$  equazioni in  $n$  incognite, cioè si deve risolvere un sistema non lineare. In termini matematici:

Date  $n$  funzioni scalari  $f^i : \mathbb{R}^n \rightarrow \mathbb{R}$ , per  $i = 1, \dots, n$ , si cerca un vettore  $x^* \in \mathbb{R}^n$  tale che:

$$\begin{cases} f^1(x^*) = 0 \\ \vdots \\ f^n(x^*) = 0. \end{cases} \quad (2)$$

Analogamente al caso scalare, il vettore  $x^*$  è detto “zero”, “radice” o “soluzione” del sistema (2). Tale sistema può anche essere espresso in maniera compatta tramite l'utilizzo di una funzione  $n$ -dimensionale. Infatti se si considera la funzione vettoriale  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , con

$$F = (f^1, \dots, f^n)^T,$$

il sistema in (2) si può riscrivere come

$$F(x^*) = 0. \quad (3)$$

I metodi numerici che permettono la risoluzione dei problemi (1) e (3) appartengono alla classe dei metodi detti *iterativi*, ovvero, assegnato un vettore iniziale  $x_0$ , generano una successione di vettori  $\{x_k\}_{k \geq 0}$ . Un metodo iterativo è *convergente* se la successione  $\{x_k\}_{k \geq 0}$  generata converge ad una soluzione  $x^*$  del problema, cioè  $\lim_{k \rightarrow \infty} x_k = x^*$ ; è *localmente convergente* se la convergenza dipende in modo critico dalla vicinanza dell'approssimazione iniziale  $x_0$  a  $x^*$ , *globalmente convergente* altrimenti.

Tutti i metodi che verranno trattati nei prossimi capitoli sono caratterizzati da una convergenza di tipo locale e sebbene sotto opportune ipotesi garantiscono tutti di trovare una soluzione del problema (1) o (3), ogni volta è importante valutare quale sia il metodo migliore da utilizzare in base alle caratteristiche del problema specifico da risolvere. La scelta di un metodo rispetto ad un altro si basa su due principali caratteristiche dei metodi: la *velocità di convergenza* e il *costo computazionale*.

La velocità di convergenza indica quanto rapidamente i valori  $x_k$  ottenuti ad ogni iterazione si avvicinano a  $x^*$ . In ordine crescente, la velocità di convergenza può essere di tipo lineare, superlineare o con ordine di convergenza  $p$ , dove  $p > 1$  (quadratica se  $p = 2$ ). La definizione dettagliata è rimandata in Appendice A Definizione A.1.1.

In generale, il costo computazionale per la risoluzione di un sistema non lineare è quantificato in base al numero di valutazioni di funzioni che è richiesto ad ogni iterazione. Questo è un aspetto fondamentale, in quanto dover calcolare il valore di una funzione in molti punti potrebbe rappresentare una grossa difficoltà. Ad esempio questo può accadere quando la funzione

## INTRODUZIONE.

---

non è nota in forma analitica, ma è il risultato di procedure sperimentali o simulazioni numeriche. Chiaramente a questo va aggiunto il costo, talvolta elevato, legato alle operazioni lineari come i prodotti matrice-vettore o la risoluzione di sistemi lineari.

Notiamo inoltre che per definizione, il procedimento generato da un metodo iterativo non ha termine ed è perciò necessario introdurre dei criteri di arresto che permettono di interrompere il procedimento ad una certa iterazione  $k$  quando  $x_k$  è sufficientemente vicino a  $x^*$  (vedi Capitolo 3).

La tesi è organizzata nel modo seguente. Nel Capitolo 1 vengono considerate le equazioni non lineari (1) e descritti il metodo di Newton e le sue varianti: metodo di Newton stazionario, metodo delle secanti e metodo di Newton alle differenze. Il metodo di Newton e il metodo delle secanti vengono anche analizzati dal punto di vista teorico. Il Capitolo 2 è incentrato sulla descrizione e sull'analisi di convergenza del metodo di Newton e dei metodi Quasi-Newton (in particolare del metodo di Broyden) per il problema (3). Infine il Capitolo 3 è dedicato all'analisi implementativa dei metodi descritti nel Capitolo 2 e alla presentazione di alcuni esperimenti numerici su semplici problemi test.

## Notazioni

Per l'intera stesura sono state assunte le seguenti notazioni:

- $f : \mathbb{R} \rightarrow \mathbb{R}$  indica una funzione scalare;
- $f' : \mathbb{R} \rightarrow \mathbb{R}$  indica la derivata prima di  $f$ ;
- Lo spazio delle matrici di dimensione  $n \times n$  a coefficienti in  $\mathbb{R}$  è stato indicato con  $\mathbb{R}^{n \times n}$ ;
- Per ogni  $x \in \mathbb{R}^n$  o  $A \in \mathbb{R}^{n \times n}$ ,  $x^T$  e  $A^T$  indicano rispettivamente il vettore e la matrice trasposta;
- $I \in \mathbb{R}^{n \times n}$  indica la matrice identità di dimensione  $n \times n$ ;
- Per ogni vettore  $x \in \mathbb{R}^n$ ,  $x^i$  indica la  $i$ -esima componente per  $i = 1, \dots, n$ ;
- Per ogni matrice  $A = (a_{ij})$  con  $i, j = 1, \dots, n$ , le scritture  $a_{i.}$  e  $a_{.j}$  indicano rispettivamente la  $i$ -esima riga e la  $j$ -esima colonna di  $A$ ;
- $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  indica una funzione  $n$ -dimensionale;
- Data  $J : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ ,  $J(x)$  indica la matrice Jacobiana della funzione  $F$  in  $x$  (vedi Appendice A);
- L'insieme delle funzioni lipschitziane con costante di Lipschitz  $L$  su un insieme  $\mathcal{D} \in \mathbb{R}^n$  viene indicato con  $Lip_L(\mathcal{D})$  (vedi Appendice A);
- $B(x^*, r) := \{x \in \mathbb{R}^n \mid \|x^* - x\| < r\}$  indica la palla  $n$ -dimensionale con centro  $x^*$  e raggio  $r$ .



# Capitolo 1

## Metodi numerici per equazioni non lineari scalari

### 1.1 Il problema

In questo capitolo vengono analizzati alcuni metodi numerici per il calcolo di zeri di equazioni non lineari scalari, cioè data una funzione non lineare scalare  $f : \mathbb{R} \rightarrow \mathbb{R}$  si cerca un'approssimazione dello scalare  $x^*$  tale che

$$f(x^*) = 0. \quad (1.1)$$

In generale, gli zeri delle equazioni possono essere di due tipologie:

**Definizione 1.1.1.** *Lo scalare  $x^*$  è detto zero semplice del problema (1.1) se  $f'(x^*) \neq 0$ .*

*Sia  $q \in \mathbb{N}$ ,  $q > 1$ , lo scalare  $x^*$  è detto zero di molteplicità  $q$  per il problema (1.1) se  $f(x^*) = f'(x^*) = \dots = f^{(q-1)}(x^*) = 0$  e  $f^{(q)}(x^*) \neq 0$ .*

In questo capitolo si utilizzano le seguenti ipotesi cosiddette “standard” sul problema (1.1) [K].

Siano  $\mathcal{D} \subset \mathbb{R}$  un intervallo aperto e  $f : \mathcal{D} \rightarrow \mathbb{R}$ . Si suppone che:

- I1:  $f$  sia differenziabile in modo continuo su  $\mathcal{D}$ ;
- I2: il problema (1.1) abbia una soluzione  $x^* \in \mathcal{D}$ ;
- I3:  $f' \in Lip_L(\mathcal{D})$  ed esista  $\rho > 0$  tale che  $|f'(x)| \geq \rho$  per ogni  $x \in \mathcal{D}$ .

### 1.2 Metodo di Newton o metodo delle tangenti

Il metodo di Newton è il principale metodo numerico per l'approssimazione di radici di equazioni non lineari.

Data una funzione non lineare  $f$  e un'approssimazione iniziale  $x_0$ , il metodo di Newton genera una successione di approssimazioni  $\{x_k\}_{k \geq 0}$  costruendo, per ogni  $k$ , un modello lineare della funzione  $f$  in un intorno di  $x_k$  e approssimando la funzione con il modello stesso. Tale modello può essere costruito partendo dallo sviluppo di Taylor della funzione  $f$  in un punto  $x$  appartenente a un intorno dell'iterata corrente  $x_k$ , cioè

$$f(x) = f(x_k) + (x - x_k)f'(x_k) + (x - x_k)^2 \frac{f^{(2)}(x_k)}{2!} + o(|x - x_k|^3).$$

Troncando tale sviluppo al primo ordine si ottiene il seguente modello lineare:

$$m_k(x) = f(x_k) + (x - x_k)f'(x_k), \quad (1.2)$$

in un intorno sufficientemente piccolo di  $x_k$ .

Dato  $x_0$  come dato iniziale, la prima iterazione consiste nel calcolare  $x_1$  come lo zero del modello lineare  $m_k$  in (1.2) con  $k = 0$ , cioè nel risolvere l'equazione lineare scalare

$$m_0(x) = 0, \quad (1.3)$$

la quale conduce all'iterata successiva  $x_1$  della forma

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Analogamente si costruiscono le iterate successive  $x_2, x_3, \dots$

Dal punto di vista geometrico, (1.2) rappresenta la retta tangente alla funzione  $f$  nel punto  $(x_k, f(x_k))$  ed è per questo motivo che il metodo viene anche detto metodo delle tangenti. In particolare, risolvere (1.3) è equivalente a determinare il punto di intersezione tra la retta  $y = f(x_0) + f'(x_0)(x - x_0)$  e l'asse  $x$ .

In Figura 1.1 viene descritto graficamente come avviene la ricerca dei primi valori  $x_1$  e  $x_2$ . L'Algoritmo 1.2.1 descrive tali passaggi.

---

**Algoritmo 1.2.1:** Metodo di Newton

---

**Dati:**  $x_0$

**Per**  $k = 0, 1, 2, \dots$

**Poni**  $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$

---

Tale algoritmo risulta essere ben definito se  $f'(x_k) \neq 0$  per ogni  $k$ . Per quanto riguarda il costo computazionale, si può notare che ad ogni iterazione sono richieste la valutazione della funzione  $f$  e della sua derivata prima nel punto  $x_k$ .

Il seguente lemma fornisce un risultato chiave nell'analisi della convergenza del metodo di Newton.

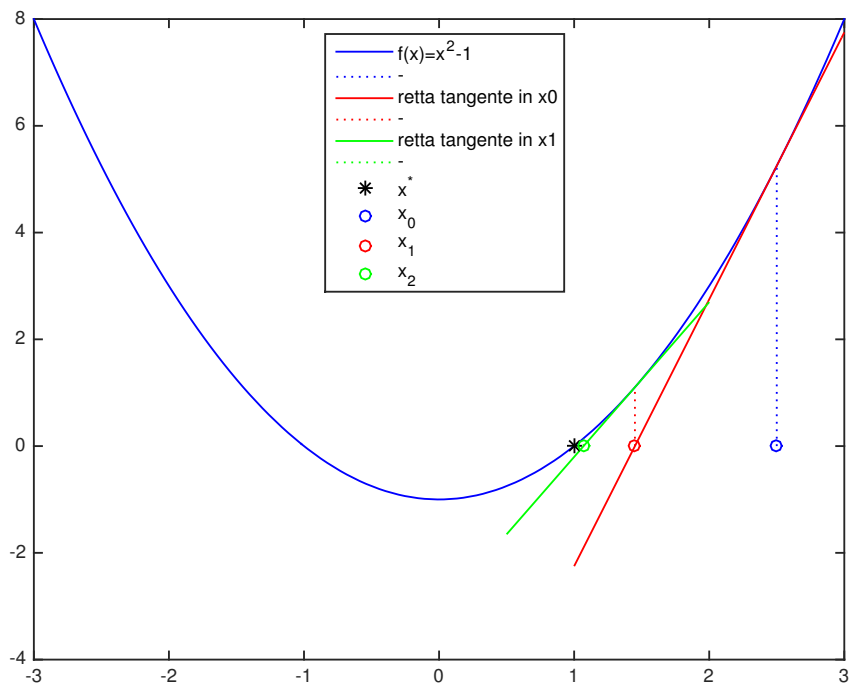


Figura 1.1: Rappresentazione grafica dei primi valori che si trovano col metodo di Newton applicato alla funzione  $f(x) = x^2 - 1$  e con approssimazione iniziale  $x_0 = 2.5$ .

**Lemma 1.2.1.** *Siano  $\mathcal{D} \subset \mathbb{R}$  intervallo aperto e  $f : \mathcal{D} \rightarrow \mathbb{R}$  tale che  $f' \in Lip_L(\mathcal{D})$ . Allora  $\forall x, y \in \mathcal{D}$  vale*

$$|f(y) - f(x) - f'(x)(y - x)| \leq \frac{L|y - x|^2}{2}. \quad (1.4)$$

*Dimostrazione.* Per il teorema fondamentale del calcolo integrale vale l'uguaglianza

$$f(y) - f(x) = \int_x^y f'(z) dz,$$

da cui si ottiene

$$f(y) - f(x) - f'(x)(y - x) = \int_x^y (f'(z) - f'(x)) dz.$$

Facendo il seguente cambio di variabili  $z = x + t(y - x)$ ,  $dz = dt(y - x)$  con  $t \in (0, 1)$ , si ricava

$$f(y) - f(x) - f'(x)(y - x) = \int_0^1 (f'(x + t(y - x)) - f'(x))(y - x) dt.$$

Utilizzando la disuguaglianza triangolare applicata all'integrale e la lipschitzianità della funzione si ottiene la tesi. Infatti

$$|f(y) - f(x) - f'(x)(y - x)| \leq |y - x| \int_0^1 L|t(y - x)| dt \leq \frac{L|y - x|^2}{2}.$$

□

Il seguente risultato caratterizza la velocità di convergenza del metodo di Newton:

**Teorema 1.2.1.** [DS, Teorema 2.4.3] *Supponendo che valgano le ipotesi standard I1, I2, I3 allora esiste  $\eta > 0$  tale che se  $|x_0 - x^*| < \eta$  allora la successione  $\{x_k\}_{k \geq 0}$  generata dal metodo di Newton è ben definita, converge a  $x^*$  e*

$$|x_{k+1} - x^*| \leq \frac{L}{2\rho} |x_k - x^*|^2, \quad \forall k \geq 0, \quad (1.5)$$

ovvero la convergenza è quadratica.

*Dimostrazione.* Siano  $\tau \in (0, 1)$  e  $\tilde{\eta} > 0$  il raggio del più grande intorno di  $x^*$  contenuto in  $\mathcal{D}$ , si definisce  $\eta := \min\{\tilde{\eta}, \tau \frac{2\rho}{L}\}$ .

Considerando che  $f(x^*) = 0$  e che vale l'Algoritmo 1.2.1,  $\forall k \geq 0$  si può ricavare il seguente risultato:

$$\begin{aligned} x_{k+1} - x^* &= x_k - x^* - \frac{f(x_k)}{f'(x_k)} = x_k - x^* - \frac{f(x_k) - f(x^*)}{f'(x_k)} \\ &= \frac{1}{f'(x_k)} (f(x^*) - f(x_k) - f'(x_k)(x^* - x_k)), \end{aligned}$$

dove si è supposto  $f'(x_k) \neq 0$ . Applicando la (1.4) si ottiene

$$|x_{k+1} - x^*| \leq \frac{L}{2|f'(x_k)|} |x_k - x^*|^2, \quad \forall k \geq 0. \quad (1.6)$$

Notando che la (1.6) vale per ogni  $k \geq 0$ , si dimostra che la successione  $\{x_k\}_{k \geq 0}$  converge a  $x^*$  mostrando per induzione su  $k$  che  $|x_{k+1} - x^*| \leq \tau |x_k - x^*|$ , per ogni  $k$ .

Infatti, per  $k = 0$ , dato che per ipotesi vale  $|x_0 - x^*| < \eta \leq \tau \frac{2\rho}{L}$  e  $|f'(x_0)| \geq \rho$ , si applica la stima (1.6) e si ottiene

$$|x_1 - x^*| \leq \tau |x_0 - x^*| < \eta.$$

Supponendo che valga fino a  $k-1$ , si dimostra per  $k$ . Per ipotesi induttiva si sa che  $|x_k - x^*| \leq \tau |x_{k-1} - x^*| < \eta$  da cui deriva  $|f'(x_k)| \geq \rho$  per l'ipotesi I2, perciò dalla (1.6) si ha

$$|x_{k+1} - x^*| \leq \frac{L}{2|f'(x_k)|} \eta |x_k - x^*| \leq \tau |x_k - x^*| \leq \dots \leq \tau^{k+1} |x_0 - x^*| \xrightarrow{k \rightarrow \infty} 0.$$

Oltre alla convergenza della successione  $\{x_k\}_{k \geq 0}$  a  $x^*$ , si può dimostrare che tale successione è interamente contenuta in  $\mathcal{D}$ , infatti si ottiene  $|x_{k+1} - x^*| \leq \tau|x_k - x^*| < \eta$ . Così grazie alla stima (1.6) e sapendo per l'ipotesi *I2* che  $|f'(x_k)| \geq \rho, \forall k \geq 0$ , si ottiene la (1.5).  $\square$

Questo teorema assume che lo zero sia semplice, (ipotesi *I2*), e garantisce l'esistenza della successione delle approssimazioni  $\{x_k\}_{k \geq 0}$  e la sua convergenza quadratica a  $x^*$ .

Se invece lo zero non è semplice si ottiene il seguente risultato di convergenza più debole.

**Teorema 1.2.2.** *Siano  $\mathcal{D} \subset \mathbb{R}$  intervallo aperto,  $q \in \mathbb{N}$ ,  $q > 1$  e  $f : \mathbb{R} \rightarrow \mathbb{R}$  funzione derivabile sino all'ordine  $q+1$ . Se  $x^* \in \mathcal{D}$  è uno zero di molteplicità  $q$  per  $f(x) = 0$  allora esiste una costante  $\eta > 0$  tale che se  $|x_0 - x^*| < \eta$  allora la successione  $\{x_k\}_{k \geq 0}$  converge linearmente a  $x^*$  e*

$$|x_{k+1} - x^*| \leq \left(1 - \frac{1}{q}\right) |x_k - x^*|, \quad \forall k \geq 0. \quad (1.7)$$

*Dimostrazione.* Considerando come modello per  $f$  lo sviluppo di Taylor di  $f$  calcolato in un intorno del punto  $x^*$  troncato all'ordine  $q$ , si ottiene

$$f(x) = f(x^*) + (x-x^*)f'(x^*) + (x-x^*)^2 \frac{f''(x^*)}{2!} + \dots + (x-x^*)^q \frac{f^{(q)}(x^*)}{q!}. \quad (1.8)$$

Dato che per ipotesi  $x^*$  è uno zero di molteplicità  $q$ , la (1.8) diventa

$$f(x) = (x - x^*)^q \frac{f^{(q)}(x^*)}{q!}. \quad (1.9)$$

Tramite l'equazione (1.9) si calcola il valore di  $f(x_k)$  e di  $f'(x_k)$ , infatti

$$f(x_k) = (x_k - x^*)^q \frac{f^{(q)}(x^*)}{q!}, \quad (1.10)$$

e applicando le regole di derivazione, si ottiene

$$f'(x_k) = \frac{q(x_k - x^*)^{q-1} f^{(q)}(x^*)}{q!}. \quad (1.11)$$

Prendendo la formula generata dall'Algoritmo 1.2.1 per il calcolo di  $x_{k+1}$ , e sostituendo le espressioni ottenute in (1.10) e in (1.11) si ottiene la (1.7). Infatti  $\forall k \geq 0$ ,

$$x_{k+1} - x^* = x_k - x^* - \frac{(x_k - x^*)^q f^{(q)}(x^*)}{q(x_k - x^*)^{q-1} f^{(q)}(x^*)} = \left(1 - \frac{1}{q}\right) (x_k - x^*). \quad \square$$

Il Teorema 1.2.2 garantisce la convergenza lineare del metodo di Newton con fattore  $\left(1 - \frac{1}{q}\right)$  ad uno zero multiplo con molteplicità  $q$ . Sarà dunque tanto più lento quanto più  $q$  è grande.

### 1.3 Varianti del metodo di Newton

Come visto nel paragrafo precedente per poter applicare il metodo di Newton è necessario conoscere la derivata prima della funzione in quanto è richiesta la sua valutazione ad ogni iterazione. Questo però non è sempre possibile; infatti calcolare  $f'(x_k)$  ad ogni iterazione potrebbe avere un costo computazionale molto alto o nel peggior dei casi la derivata potrebbe non essere nota. Per questo motivo sono state apportate variazioni al metodo di Newton che permettono di trovare approssimazioni di soluzioni di equazioni non lineari senza utilizzare la derivata di  $f$  in modo esplicito ma bensì una sua approssimazione. Proprio la strategia di approssimazione della derivata nell'iterazione corrente caratterizza la variante del metodo di Newton:

- nel *metodo di Newton stazionario* si fissa  $f'(x_k) = f'(x_0)$  per ogni  $k \geq 0$ ;
- nel *metodo delle secanti* si usa il coefficiente della retta secante passante per i punti  $(x_{k-1}, f(x_{k-1}))$  e  $(x_k, f(x_k))$ ;
- nel *metodo alle differenze* si utilizza il rapporto incrementale tra i punti  $(x_k, f(x_k))$  e  $(x_k + h, f(x_k + h))$ , dove  $h$  è uno scalare vicino al valore 0.

#### 1.3.1 Metodo di Newton stazionario

Il metodo di Newton stazionario richiede la conoscenza del valore di  $f'$  solo in un punto. Infatti, dato  $x_0$  punto iniziale, i termini della successione  $\{x_k\}_{k \geq 0}$  vengono calcolati così:

---

**Algoritmo 1.3.1:** Metodo di Newton stazionario

---

**Dati:**  $x_0$

**Per**  $k = 0, 1, 2, \dots$

**Poni**  $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_0)}$

---

Si nota che l'algoritmo è ben definito se  $f'(x_0) \neq 0$ . La velocità di convergenza che caratterizza questo metodo è descritta dal seguente teorema:

**Teorema 1.3.1.** [K, Teorema 5.4.2] *Supponendo che valgano le ipotesi standard I1, I2, I3 allora esiste  $\eta > 0$  tale che se  $|x_0 - x^*| < \eta$  allora la successione  $\{x_k\}_{k \geq 0}$  generata dal metodo di Newton stazionario converge linearmente a  $x^*$ .*

### 1.3.2 Metodo delle secanti

Il metodo delle secanti può essere utilizzato quando non si conosce la derivata della funzione  $f$  e si basa sulla seguente osservazione: non potendo sfruttare la retta tangente in  $(x_k, f(x_k))$ , per la costruzione del modello lineare si considera la retta secante per  $(x_{k-1}, f(x_{k-1}))$  e  $(x_k, f(x_k))$ , il cui calcolo richiede la conoscenza della sola funzione  $f$ . Il metodo delle secanti è perciò basato sull'approssimazione della funzione  $f$  attraverso un modello lineare che richiede la conoscenza di  $f$  in due punti e pertanto l'algoritmo necessita due dati iniziali  $x_{-1}$  e  $x_0$ .

Dal punto di vista geometrico, alla prima iterazione si considera la retta secante alla funzione passante per i punti  $(x_{-1}, f(x_{-1}))$  e  $(x_0, f(x_0))$ , che ha coefficiente angolare

$$\alpha_0 = \frac{f(x_0) - f(x_{-1})}{x_0 - x_{-1}}.$$

Il modello lineare calcolato nel punto  $(x_0, f(x_0))$  ha dunque la seguente forma:

$$m_0(x) = f(x_0) + \alpha_0(x - x_0).$$

A questo punto si cerca la radice  $x_1$  di tale modello, ossia si impone che  $m_0(x) = 0$ , e il valore che si ottiene è

$$x_1 = x_0 - \frac{f(x_0)}{\alpha_0}.$$

Il valore  $x_1$  così ottenuto rappresenta un'approssimazione della radice  $x^*$  dell'equazione  $f(x) = 0$ . Iterando il procedimento appena descritto, si ottiene l'Algoritmo 1.3.2.

---

**Algoritmo 1.3.2:** Metodo delle secanti

---

**Dati:**  $x_{-1}, x_0$

**Per**  $k = 0, 1, 2, \dots$

<b>Calcola</b>	$\alpha_k = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$
<b>Poni</b>	$x_{k+1} = x_k - \frac{f(x_k)}{\alpha_k}$

---

Affinché l'algoritmo sia ben definito è necessario che  $\alpha_k \neq 0$ , per ogni  $k$ . Inoltre si può notare che all'interno di ogni iterazione è richiesto il calcolo del valore  $\alpha_k$  il quale implica un'unica valutazione di  $f$  in  $x_k$  poiché il valore di  $f(x_{k-1})$  viene recuperato dall'iterazione precedente.

Il seguente lemma fornisce una stima sull'errore  $|x_k - x^*|$  e mostra che questa dipende dall'errore al passo precedente e da quanto  $\alpha_k$  approssima bene  $f'(x_k)$ .

**Lemma 1.3.1.** *Supponendo che valgano le ipotesi standard I1, I2, I3 e considerando la successione generata dal metodo delle secanti  $\{x_k\}_{k \geq 0}$ , si definisce*

$$\Delta(x_k) := \alpha_k - f'(x_k). \quad (1.12)$$

*Allora esistono le costanti  $\zeta, \eta, \delta > 0$  tali che se  $|x^* - x_k| < \eta$  e  $|\Delta(x_k)| < \delta$  allora  $x_{k+1}$  è ben definito, cioè  $\alpha_k \neq 0$ , e soddisfa*

$$|x_{k+1} - x^*| \leq \zeta(|x_k - x^*|^2 + |\Delta(x_k)||x_k - x^*|). \quad (1.13)$$

Il metodo delle secanti è caratterizzato da una velocità di convergenza superlineare come descritto nel seguente risultato:

**Teorema 1.3.2.** [K, Teorema 5.4.7] *Supponendo che valgano le ipotesi standard I1, I2, I3 allora esiste  $\eta > 0$  tale che se  $x_{-1}$  e  $x_0$  sono punti distinti e appartengono all'intervallo di centro  $x^*$  e raggio  $\eta$  allora la successione  $\{x_k\}_{k \geq 0}$  generata dal metodo delle secanti converge superlinearmente a  $x^*$ .*

*Dimostrazione.* Sia  $\eta > 0$  tale che l'intervallo aperto  $(x^* - \eta, x^* + \eta) \subset \mathcal{D}$  e si suppone che tutti i termini della successione  $\{x_k\}_{k \geq 0}$  siano contenuti in  $\mathcal{D}$ . Nel caso in cui  $x_k = x^*$  per qualche  $k$  finito, allora si ottiene il risultato. Se invece  $\forall k \geq 0$  vale  $x_k \neq x_{k-1}$ , considerando  $\Delta(x_k)$  definito in (1.12) attraverso il teorema fondamentale del calcolo integrale si ottiene

$$\Delta(x_k) = \int_0^1 [f'(x_{k-1} + t(x_k - x_{k-1})) - f'(x_k)] dt.$$

Grazie alla disuguaglianza triangolare applicata all'integrale e alla lipschitzianità della funzione  $f'$ , si arriva alla disuguaglianza

$$|\Delta(x_k)| \leq L \frac{|x_k - x_{k-1}|}{2},$$

da cui aggiungendo e sottraendo  $x^*$ , si ottiene

$$|\Delta(x_k)| \leq L \frac{|x_k - x^* + x^* - x_{k-1}|}{2} \leq L \frac{|x_k - x^*| + |x_{k-1} - x^*|}{2}. \quad (1.14)$$

Sostituendo la (1.14) in (1.13) si arriva alla seguente disuguaglianza:

$$\begin{aligned} |x_{k+1} - x^*| &\leq \zeta(|x_k - x^*|^2 + L \frac{|x_k - x^*| + |x_{k-1} - x^*|}{2} |x_k - x^*|) \\ &= \zeta((1 + \frac{L}{2})|x_k - x^*|^2 + \frac{L}{2}|x_{k-1} - x^*||x_k - x^*|), \end{aligned}$$

da cui raccogliendo  $|x_k - x^*|$ , si ottiene

$$|x_{k+1} - x^*| \leq \zeta((1 + \frac{L}{2})|x_k - x^*| + \frac{L}{2}|x_{k-1} - x^*|)|x_k - x^*|. \quad (1.15)$$



Dato che  $|x_{k-1} - x^*|, |x_k - x^*| < \eta$ , riducendo al più  $\eta$  si può supporre che

$$\zeta(1 + L)\eta =: \beta < 1.$$

In questo modo si ricava

$$|x_{k+1} - x^*| < \beta|x_k - x^*|, \quad \forall k \geq 0,$$

e cioè la convergenza lineare di  $\{x_k\}_{k \geq 0}$  a  $x^*$ .

Come conseguenza si ottiene la convergenza di tipo superlineare, infatti dalla disuguaglianza espressa in (1.15), si può scrivere

$$\frac{|x_{k+1} - x^*|}{|x_k - x^*|} \leq \zeta \left( \left(1 + \frac{L}{2}\right)|x_k - x^*| + \frac{L}{2}|x_{k-1} - x^*| \right) \rightarrow_{k \rightarrow \infty} 0.$$

□

### 1.3.3 Metodo di Newton alle differenze

Il metodo di Newton alle differenze approssima i valori di  $f'$  usando la definizione stessa di derivata. Infatti si costruisce un modello lineare che approssima la funzione usando il rapporto incrementale tra due punti molto vicini tra di loro.

Dato il punto iniziale  $x_0$  e un valore  $h_0$  “piccolo”, la prima iterazione consiste nel calcolare

$$\alpha_0^{h_0} = \frac{f(x_0 + h_0) - f(x_0)}{h_0},$$

che rappresenta il rapporto incrementale della funzione  $f$  tra i punti  $(x_0 + h_0, f(x_0 + h_0))$  e  $(x_0, f(x_0))$ . Poi si costruisce il modello lineare della funzione calcolato nel punto  $(x_0, f(x_0))$ , la cui espressione è

$$m_0(x) = f(x_0) + \alpha_0^{h_0}(x - x_0).$$

Per ricavare una nuova approssimazione  $x_1$  di  $x^*$ , si cerca la radice del modello lineare ottenendo

$$x_1 = x_0 - \frac{f(x_0)}{\alpha_0^{h_0}}.$$

Iterando questo procedimento, si costruisce una successione di valori  $\{x_k\}_{k \geq 0}$  che sotto opportune ipotesi converge a  $x^*$ .

L’algoritmo è descritto in Algoritmo 1.3.3.

Si ha la buona definizione dell’algoritmo se  $\alpha_k^{h_k} \neq 0$ , per ogni  $k$ .

Tale algoritmo richiede una scelta opportuna del valore da attribuire ad  $h_k$ . Da una parte, più  $h_k$  è vicino a 0 e più  $\alpha_k^{h_k}$  risulta una buona approssimazione della derivata in  $x_k$ , dall’altra  $h_k$  deve essere un valore abbastanza

---

**Algoritmo 1.3.3:** Metodo di Newton alle differenze

---

**Dati:**  $x_0, h_0$

**Per**  $k = 0, 1, 2, \dots$

{

**Calcola**  $\alpha_k^{h_k} = \frac{f(x_k+h_k)-f(x_k)}{h_k}$   
**Poni**  $x_{k+1} = x_k - \frac{f(x_k)}{\alpha_k^{h_k}}$

---

grande in modo che il calcolo del numeratore di  $\alpha_k^{h_k}$  non dia luogo al fenomeno dell'errore di cancellazione. Un buon compromesso viene raggiunto calcolando il valore di  $h_k$  nel modo seguente:

$$h_k = \begin{cases} \sqrt{\epsilon_m} x_k, & x_k \neq 0 \\ \sqrt{\epsilon_m}, & x_k = 0, \end{cases}$$

dove  $\epsilon_m$  indica la precisione di macchina.

Il metodo di Newton alle differenze richiede due valutazioni di funzione ad ogni iterazione; la velocità di convergenza è superlineare come descritto nel seguente teorema.

**Teorema 1.3.3.** [DS, Teorema 5.4.1] *Supponendo che valgano le ipotesi standard I1, I2, I3 allora esistono le costanti  $\lambda, \eta > 0$  tale che se  $\{h_k\}_{k \geq 0}$  è una successione reale con  $0 < |h_k| \leq \lambda$  tale che  $\lim_{k \rightarrow \infty} h_k = 0$  e se  $|x_0 - x^*| < \eta$  allora la successione  $\{x_k\}_{k \geq 0}$  definita dal metodo di Newton alle differenze converge superlinearmente a  $x^*$ .*

## 1.4 Confronto tra i metodi

La Tabella 1.1 riassume le caratteristiche dei vari metodi descritti nelle sezioni precedenti.

	$x_0$	num. di valutazioni	valutazione $f'$	velocità di convergenza
Newton	$x_0 \sim x^*$	2	sì	quadratica
Newton Staz.	$x_0 \sim x^*$	1	no	lineare
Newton Diff.	$x_0 \sim x^*$	2	no	superlineare
Secanti	$x_0 \sim x^*$	1	no	superlineare

Tabella 1.1: Classificazione dei metodi numerici per equazioni non lineari scalari con zeri semplici in base al dato iniziale, numero di valutazioni di funzione necessarie ad ogni iterazione, conoscenza di  $f'$  e velocità di convergenza.

Dalla Tabella 1.1 si deduce che se si è a conoscenza della derivata prima e la valutazione di  $f$  e  $f'$  non è costosa il metodo da preferire è il metodo di Newton in quanto converge più rapidamente rispetto agli altri metodi. Se invece la derivata prima non è nota e la valutazione della funzione ha un costo contenuto allora il metodo migliore è rappresentato dal metodo di Newton alle differenze, il quale possiede una velocità di convergenza superlineare. Nel caso in cui calcolare il valore della funzione risulti essere costoso, il metodo che richiede il minor numero di valutazioni e di iterazioni prima di giungere a una buona approssimazione di  $x^*$  è il metodo delle secanti.



## Capitolo 2

# Metodi numerici per sistemi di equazioni non lineari

In questo capitolo vengono analizzati due dei principali metodi numerici per sistemi non lineari, che rappresentano la generalizzazione al caso  $n$ -dimensionale del metodo delle tangenti e del metodo delle secanti. Analogamente al caso scalare, anche in più dimensioni i metodi si basano sull'approssimazione della funzione  $n$ -dimensionale attraverso un modello lineare.

Di seguito è riportato un breve riepilogo del caso scalare che verrà generalizzato al caso  $n$ -dimensionale. Sia  $x_k$  l'iterata corrente. Il procedimento ad ogni iterazione può essere riassunto così:

- 1) Si considera il modello lineare

$$m_k(x) = f(x_k) + \alpha_k(x - x_k), \quad (2.1)$$

che soddisfa  $m_k(x_k) = f(x_k)$  per ogni  $\alpha_k \in \mathbb{R}$ .

Nel caso in cui si stia utilizzando il metodo di Newton  $\alpha_k$  risulta essere pari alla quantità  $f'(x_k)$ . Se invece la derivata prima di  $f$  non è nota,  $\alpha_k$  risulta essere un valore che approssima  $f'(x_k)$ . In particolare nel metodo delle secanti si chiede che il modello soddisfi  $m_k(x_{k-1}) = f(x_{k-1})$ , cioè

$$f(x_{k-1}) = f(x_k) + \alpha_k(x_{k-1} - x_k),$$

che porta a

$$\alpha_k = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}.$$

- 2) Si calcola l'approssimazione successiva  $x_{k+1}$  imponendo  $m_k(x) = 0$ , cioè si pone

$$x_{k+1} = x_k - \frac{f(x_k)}{\alpha_k}.$$

Nel caso di più dimensioni, si risolve un sistema non lineare della forma

$$F(x) = 0, \quad (2.2)$$

con  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Come già anticipato, anche nel caso multidimensionale si costruisce un modello lineare, della stessa forma di (2.1):

$$M_k(x) = F(x_k) + A_k(x - x_k), \quad (2.3)$$

dove  $x_k$  rappresenta un vettore di lunghezza  $n$ , che si suppone essere il risultato ottenuto alla  $k$ -esima iterazione, e  $A_k$  è una matrice a coefficienti reali di dimensione  $n \times n$ . Tale modello inoltre soddisfa la condizione  $M_k(x_k) = F(x_k)$ , per ogni  $A_k \in \mathbb{R}^{n \times n}$ .

Il modo in cui viene definita la matrice  $A_k$  caratterizza i metodi proposti:

- i) Nel caso in cui la matrice corrisponda alla matrice Jacobiana, cioè

$$A_k = J(x_k),$$

il metodo risulta essere la generalizzazione del metodo di Newton scalare e il modello lineare risulta essere

$$M_k(x) = F(x_k) + J(x_k)(x - x_k).$$

Supposto che  $J(x_k)$  sia invertibile, l'iterata successiva si ottiene imponendo  $M_k(x) = 0$ , dunque

$$x_{k+1} = x_k - J(x_k)^{-1}F(x_k).$$

- ii) Se invece la matrice Jacobiana non è nota, la matrice  $A_k$  viene scelta come una sua approssimazione dando luogo a diversi metodi chiamati *metodi Quasi-Newton*. In questi metodi il modello lineare è rappresentato da (2.3) e il procedimento per ottenere le approssimazioni successive di  $x^*$  si ottiene imponendo l'annullamento del modello lineare, da cui si giunge a

$$x_{k+1} = x_k - A_k^{-1}F(x_k).$$

Il *metodo di Broyden* è un particolare metodo Quasi-Newton e rappresenta la generalizzazione del metodo delle secanti al caso  $n$ -dimensionale. Infatti, come nel caso unidimensionale, si chiede al modello che venga verificata l'equazione  $M_k(x_{k-1}) = F(x_{k-1})$ .

In questo capitolo verranno analizzati in modo più dettagliato i metodi di Newton multidimensionale e di Broyden.

## Le ipotesi standard

Per l'intero capitolo si utilizzano le seguenti ipotesi cosiddette standard sul problema (2.2).

Siano  $\mathcal{D} \subset \mathbb{R}^n$  un insieme aperto convesso e  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Si suppone che:

- I1:  $F$  sia differenziabile in modo continuo su  $\mathcal{D}$ ;
- I2: il problema (2.2) abbia una soluzione  $x^* \in \mathbb{R}^n$  ed esista  $r > 0$  tale che  $B(x^*, r) \subset \mathcal{D}$ ;
- I3:  $J \in Lip_L(B(x^*, r))$ ,  $J(x^*)$  sia invertibile ed esista  $\rho > 0$  tale che  $\|J(x^*)^{-1}\| \leq \rho$ .

## 2.1 Metodo di Newton

### 2.1.1 Costruzione del modello lineare

La costruzione del modello lineare avviene a partire dal Lemma A.3.1 che rappresenta la generalizzazione della formula di Taylor al prim'ordine del caso scalare dove l'errore è espresso in forma integrale. Il Lemma A.3.1 garantisce la validità della seguente formula:

$$F(x + s) = F(x) + \int_x^{x+s} J(z) dz, \quad (2.4)$$

dove  $F \in C^1(\mathcal{D})$  con  $\mathcal{D} \subset \mathbb{R}^n$ ;  $x, x + s \in \mathcal{D}$  e  $J(z)$  è la matrice Jacobiana della funzione  $F$  nel punto  $z \in \mathbb{R}^n$ .

Tale formula è fondamentale per ottenere l'approssimazione della funzione che si sta cercando in quanto l'integrale in (2.4) può essere approssimato nel modo seguente:

$$\int_x^{x+s} J(z) dz \simeq J(x)s,$$

dove  $J(x)s$  è un termine lineare. Sostituendo il risultato appena ottenuto in (2.4) si giunge a

$$F(x + s) \simeq F(x) + J(x)s,$$

e cioè a un'approssimazione lineare per  $F$ .

Come conseguenza di questo risultato si può definire il modello lineare di  $F$  calcolato in  $x$  nel modo seguente:

$$M(x + s) = F(x) + J(x)s. \quad (2.5)$$

Si nota che il modello lineare è una buona approssimazione di  $F$  in un intorno sufficientemente piccolo di  $x$ . Infatti la differenza tra  $M(x + s)$  e  $F(x + s)$

si può scrivere così:

$$F(x+s) - M(x+s) = \int_x^{x+s} [J(z) - J(x)] dz. \quad (2.6)$$

Inoltre ricordandosi l'ipotesi  $F \in C^1(\mathcal{D})$ ,  $J$  è continua e  $\|J(z) - J(x)\| \rightarrow 0$  se  $z \rightarrow x$ ,  $\forall z \in \mathcal{D}$ . Per il Lemma A.3.2 si può stimare l'entità della differenza tra funzione e modello espressa nella (2.6):

$$\begin{aligned} \|F(x+s) - M(x+s)\| &= \left\| \int_x^{x+s} [J(z) - J(x)] dz \right\| \\ &\leq \int_x^{x+s} \|J(z) - J(x)\| dz \leq \|s\| \sup_{z \in \mathcal{D}} \|J(z) - J(x)\| \xrightarrow{s \rightarrow 0} 0. \end{aligned}$$

Da questa stima si può notare che se  $s$  assume valori “piccoli” allora il modello lineare così come è stato definito in (2.5) risulta essere una buona approssimazione per la funzione  $F$ , cioè  $F(x+s) \simeq M(x+s)$ .

### 2.1.2 Algoritmo

Come già anticipato, il metodo di Newton è iterativo e ciò significa che, dato un valore iniziale  $x_0$ , l'algoritmo crea una successione di valori  $\{x_k\}_{k \geq 0}$  che converge ad uno zero del sistema non lineare, sotto opportune ipotesi su  $x_0$ . Tali valori sono in pratica le radici del modello lineare che ad ogni iterazione viene aggiornato, ossia viene calcolato nel valore ottenuto dall'iterazione precedente.

Quindi dato un sistema non lineare  $F(x) = 0$  e un punto iniziale  $x_0$ , per trovare un'approssimazione della radice del sistema è necessario calcolare il modello lineare  $M$ , definito in (2.5), nel punto  $x_0$  ottenendo così

$$M_0(x) = F(x_0) + J(x_0)(x - x_0).$$

Si prosegue poi imponendone l'annullamento, ossia  $M_0(x) = 0$ , che implica la risoluzione del seguente sistema lineare:

$$J(x_0)s_0 = -F(x_0),$$

cioè il calcolo del vettore

$$s_0 = -J(x_0)^{-1}F(x_0), \quad (2.7)$$

dove  $s_0 = x - x_0$ . Grazie a questo si trova  $x_1$ , nuova approssimazione di  $x^*$ , nel modo seguente:

$$x_1 = x_0 + s_0.$$



---

**Algoritmo 2.1.1:** Metodo di Newton

---

**Dati:**  $x_0$

**Per**  $k = 0, 1, 2, \dots$

**Risolvi**  $J(x_k)s_k = -F(x_k)$   
    **Poni**  $x_{k+1} = x_k + s_k$

---

Il vettore  $s_0$  in (2.7) è detto *passo di aggiornamento* o *passo di Newton*. L'algoritmo è descritto in Algoritmo 2.1.1.

Chiaramente la buona definizione dell'algoritmo dipende dall'invertibilità di  $J$  in  $x_k$ .

In modo equivalente il procedimento descritto può essere applicato lavorando sulle componenti della funzione  $F = (f^1, \dots, f^n)^T$ . Da questo punto di vista l'algoritmo consiste nel trovare un valore che sia simultaneamente uno zero del modello lineare  $(M_k)^i$  di ogni componente di  $F$ , dove  $\forall i = 1, \dots, n$

$$(M_k)^i(x) = f^i(x_k) + \nabla f^i(x_k)^T(x - x_k),$$

dove  $\nabla f^i(x_k)$  rappresenta il gradiente della  $i$ -esima componente di  $F$  calcolato in  $x_k$ .

**Osservazione 1.** *Se le equazioni che devono essere soddisfatte sono tutte lineari, il sistema è lineare e l'applicazione del metodo di Newton porta alla soluzione esatta alla prima iterazione.*

### 2.1.3 Analisi di convergenza

Il seguente lemma fornisce un risultato chiave nell'analisi della convergenza.

**Lemma 2.1.1.** [DS, Lemma 4.1.12] *Siano  $\mathcal{D} \subset \mathbb{R}^n$  un insieme aperto convesso,  $x \in \mathcal{D}$  e  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  una funzione differenziabile in modo continuo tale che  $J$  sia lipschitziana in  $x \in \mathcal{D}$  con costante di Lipschitz  $L$ . Si consideri  $\|\cdot\|$  una norma vettoriale e la norma matriciale indotta, allora  $\forall x + s \in \mathcal{D}$  vale*

$$\|F(x + s) - F(x) - J(x)s\| \leq \frac{L}{2}\|s\|^2.$$

La convergenza del metodo di Newton  $n$ -dimensionale è espressa dal seguente risultato, infatti:

**Teorema 2.1.1.** [DS, Teorema 5.2.1] *Supponendo che valgano le ipotesi standard I1, I2, I3 allora esiste  $\eta > 0$  tale che per ogni  $x_0 \in B(x^*, \eta)$  la successione  $\{x_k\}_{k \geq 0}$  generata dal metodo di Newton è ben definita, converge a  $x^*$  e*

$$\|x_{k+1} - x^*\| \leq \rho L \|x_k - x^*\|^2, \quad \forall k \geq 0, \quad (2.8)$$

*ovvero la convergenza è quadratica.*

*Dimostrazione.* Si sceglie  $\eta$  in modo tale che  $J(x)$  sia una matrice non singolare per ogni  $x \in B(x^*, \eta)$ , perciò si definisce  $\eta := \min \{r, \frac{1}{2\rho L}\}$ .

La dimostrazione del teorema avviene per induzione su  $k$ .

Per  $k = 0$ :

Come prima cosa si dimostra che la matrice  $J(x_0)$  è non singolare, infatti grazie alle ipotesi del teorema si possono svolgere i seguenti passaggi:

$$\begin{aligned} \|J(x^*)^{-1}[J(x_0) - J(x^*)]\| &\leq \|J(x^*)^{-1}\| \|J(x_0) - J(x^*)\| \\ &\leq \rho L \|x_0 - x^*\| \leq \rho L \eta \leq \frac{1}{2}. \end{aligned} \quad (2.9)$$

Come conseguenza del Lemma A.2.1 si ottiene l'invertibilità di  $J(x_0)$ , da cui deriva l'esistenza di  $J(x_0)^{-1}$  e la buona definizione di  $x_1$  calcolato attraverso l'algoritmo del metodo di Newton. Inoltre, sempre come conseguenza del Lemma A.2.1 e utilizzando la stima (2.9) si giunge alla seguente disuguaglianza:

$$\|J(x_0)^{-1}\| \leq \frac{\|J(x^*)^{-1}\|}{1 - \|J(x^*)^{-1}[J(x_0) - J(x^*)]\|} \leq 2\|J(x^*)^{-1}\| \leq 2\rho. \quad (2.10)$$

Dall'esistenza di  $x_1$  e dal fatto che  $F(x^*) = 0$  si può scrivere

$$\begin{aligned} x_1 - x^* &= x_0 - x^* - J(x_0)^{-1}F(x_0) = x_0 - x^* - J(x_0)^{-1}[F(x_0) - F(x^*)] \\ &= J(x_0)^{-1}[F(x^*) - F(x_0) - J(x_0)(x^* - x_0)]. \end{aligned}$$

Applicando la norma a queste quantità, utilizzando la stima ottenuta in (2.10) e il Lemma 2.1.1 si ottiene la (2.8) per  $k = 0$ , infatti

$$\|x_1 - x^*\| \leq \|J(x_0)^{-1}\| \|F(x^*) - F(x_0) - J(x_0)(x^* - x_0)\| \leq \rho L \|x_0 - x^*\|^2. \quad (2.11)$$

Dato che per ipotesi  $\|x_0 - x^*\| < \eta \leq \frac{1}{2\rho L}$ , dalla disuguaglianza (2.11) si ottiene  $\|x_1 - x^*\| \leq \frac{1}{2}\|x_0 - x^*\|$  e ciò significa che  $x_1 \in B(x^*, \eta)$ .

Si suppone ora che il teorema valga fino a  $k - 1$ , e si dimostra per  $k$ :

Per ipotesi induttiva si sa che:  $J(x_{k-1})^{-1}$  esiste,  $\|J(x_{k-1})^{-1}\| \leq 2\rho$ ,  $x_k$  esiste ed è ben definito,  $\|x_k - x^*\| \leq \rho L \|x_{k-1} - x^*\|^2$  da cui si arriva a dire che  $\|x_k - x^*\| \leq \frac{1}{2}\|x_{k-1} - x^*\|$ , e cioè  $x_k \in B(x^*, \eta)$ .

Sfruttando tali risultati, analogamente a quanto fatto per il caso  $k = 0$ , si ottiene l'esistenza e la buona definizione di  $x_{k+1}$  e la stima  $\|x_{k+1} - x^*\| \leq \rho L \|x_k - x^*\|^2$ , da cui deriva

$$\|x_{k+1} - x^*\| \leq \frac{1}{2}\|x_k - x^*\|. \quad (2.12)$$

Si conclude perciò che  $\forall k \geq 0$  vale la formula (2.12) la quale implica la convergenza della successione  $\{x_k\}_{k \geq 0}$  a  $x^*$  che risulta essere di tipo quadratica grazie alla formula (2.8).  $\square$

In Tabella 2.1 vengono riassunte le caratteristiche del metodo di Newton  $n$ -dimensionale.

<b>Vantaggi</b>
1) Velocità di convergenza quadratica se $J(x^*)$ è non singolare; 2) Soluzione esatta in una sola iterazione se $F$ è lineare.
<b>Svantaggi</b>
1) $x_0$ sufficientemente vicino a $x^*$ ; 2) Richiede $J(x_k)$ ad ogni iterazione; 3) Ogni iterazione richiede la risoluzione di un sistema lineare, la cui matrice dei coefficienti potrebbe essere singolare.

Tabella 2.1: Elenco dei vantaggi e degli svantaggi del metodo di Newton  $n$ -dimensionale.

## 2.2 Metodo di Broyden

In questa sezione sono riportati alcuni risultati validi per il caso generale dei metodi Quasi-Newton, necessari per l'analisi dello specifico metodo di Broyden.

### 2.2.1 Costruzione del modello lineare

La costruzione del modello lineare della funzione  $F$  avviene analogamente a quanto fatto nel caso scalare. Infatti, supponendo di essere giunti alla  $k$ -esima iterazione e di conoscere  $x_{k-1}$  e  $x_k$ , per i metodi Quasi-Newton la forma del modello lineare alla  $(k+1)$ -esima iterazione calcolato nel punto  $x$  è la seguente:

$$M_{k+1}(x) = F(x_{k+1}) + A_{k+1}(x - x_{k+1}), \quad (2.13)$$

dove  $A_{k+1}$  è una matrice di dimensione  $n \times n$ . Il metodo di Broyden richiede che  $M_{k+1}(x_k) = F(x_k)$  e cioè che soddisfi la relazione

$$A_{k+1}(x_{k+1} - x_k) = F(x_{k+1}) - F(x_k).$$

Tale equazione è detta *equazione delle secanti*. Utilizzando le notazioni  $s_k = x_{k+1} - x_k$  e  $y_k = F(x_{k+1}) - F(x_k)$ , l'equazione delle secanti può essere scritta nel modo seguente:

$$A_{k+1}s_k = y_k. \quad (2.14)$$

Data questa relazione si può però notare che, se  $n > 1$ , la matrice  $A_{k+1}$  non è univocamente determinata ma esiste un sottospazio affine di matrici di dimensione  $n(n-1)$  che soddisfano l'equazione in (2.14).

Considerando  $M_{k+1}$  e  $M_k$  i modelli lineari costruiti rispettivamente alla  $(k+1)$ -esima e  $k$ -esima iterazione, una buona scelta della matrice  $A_{k+1}$  risulta essere quella che minimizza la differenza tra i due modelli calcolati in  $x \in \mathbb{R}^n$ , cioè l'espressione

$$M_{k+1}(x) - M_k(x) = F(x_{k+1}) + A_{k+1}(x - x_{k+1}) - F(x_k) - A_k(x - x_k),$$

da cui aggiungendo e sottraendo il termine  $A_{k+1}x_k$  e usando la (2.14) si ottiene

$$M_{k+1}(x) - M_k(x) = (A_{k+1} - A_k)(x - x_k).$$

La quantità  $x - x_k$ , per ogni  $x \in \mathbb{R}^n$ , può essere vista come combinazione lineare di due direzioni ortogonali tra di loro e cioè

$$x - x_k = \alpha(x_{k+1} - x_k) + t = \alpha s_k + t,$$

dove  $t^T s_k = 0$ . Perciò il termine che deve essere minimizzato diventa:

$$M_{k+1}(x) - M_k(x) = \alpha(A_{k+1} - A_k)s_k + (A_{k+1} - A_k)t.$$

Il primo addendo, grazie all'equazione delle secanti, si può scrivere  $(A_{k+1} - A_k)s_k = y_k - A_k s_k$  e perciò tutte le quantità sono già determinate. Al contrario, si può agire sul secondo addendo  $(A_{k+1} - A_k)t$  rendendolo nullo, cioè scegliendo  $A_{k+1}$  in modo che

$$(A_{k+1} - A_k)t = 0, \quad \forall t : t^T s_k = 0. \quad (2.15)$$

Affinché venga soddisfatta la condizione (2.15) è necessario che la matrice  $A_{k+1} - A_k$  abbia rango 1 e che sia della forma di  $us_k^T$ , con  $u \in \mathbb{R}^n$ . Ora per mantenere la validità dell'equazione delle secanti, la quale implica  $(A_{k+1} - A_k)s_k = y_k - A_k s_k$ , deve valere

$$u = \frac{y_k - A_k s_k}{s_k^T s_k}.$$

Quindi si ottiene

$$A_{k+1} = A_k + \frac{(y_k - A_k s_k)s_k^T}{s_k^T s_k}. \quad (2.16)$$

Il seguente lemma mostra che la scelta di  $A_{k+1}$  in (2.16) comporta il minimo cambiamento della matrice  $A_k$ . Si consideri la seguente notazione:

$$Q(y, s) = \{B \in \mathbb{R}^{n \times n} \mid Bs = y\}$$

Si nota che  $A_{k+1} \in Q(y_k, s_k)$  dato che deve essere verificata l'equazione delle secanti (2.14).

**Lemma 2.2.1.** *Siano  $A, B \in \mathbb{R}^{n \times n}$  e  $s, y \in \mathbb{R}^n$  con  $s \neq 0$ . Allora per ogni norma matriciale  $\|\cdot\|$  che soddisfa la proprietà*

$$\left\| \frac{vv^T}{v^T v} \right\| = 1 \quad \forall v \in \mathbb{R}^n,$$

la soluzione del seguente problema di minimo

$$\min_{B \in Q(y, s)} \|B - A\|$$

è

$$A' = A + \frac{(y - As)s^T}{s^T s}.$$

*Dimostrazione.* Sia  $B \in Q(y, s)$ , allora

$$\begin{aligned} \|A' - A\| &= \left\| \frac{(y - As)s^T}{s^T s} \right\| = \left\| \frac{(B - A)ss^T}{s^T s} \right\| \\ &\leq \|B - A\| \left\| \frac{ss^T}{s^T s} \right\| \leq \|B - A\|. \end{aligned}$$

Da cui si ottiene la tesi.  $\square$

In particolare da questo lemma si ritrova la definizione della matrice  $A_{k+1}$  data in (2.16), infatti

$$A_{k+1} = A_k + \frac{(y_k - A_k s_k)s_k^T}{s_k^T s_k} = A_k + \frac{F(x_{k+1})s_k^T}{s_k^T s_k}. \quad (2.17)$$

Se il lemma viene applicato utilizzando la norma di Frobenius la soluzione del problema di minimo risulta essere unica, mentre se si considerano altre norme è possibile che esistano più soluzioni.

Si può notare che il calcolo di  $A_{k+1}$  si ottiene a partire dall'approssimazione  $A_k$  di  $J(x_k)$  al passo precedente, per questo motivo si parla di *aggiornamento* della matrice di approssimazione.

Il modo in cui la matrice  $A_k$  definita dal metodo di Broyden approssima la Jacobiana è descritto dal seguente risultato e verrà utilizzato nella dimostrazione di convergenza del metodo (vedi Teorema 2.2.3).

**Lemma 2.2.2.** [DS, Lemma 8.2.1] *Siano  $\mathcal{D} \subset \mathbb{R}^n$  un insieme aperto convesso,  $x_k, x_{k+1}, x^* \in \mathcal{D}$  con  $x_k \neq x^*$ ,  $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $A_k \in \mathbb{R}^{n \times n}$  e  $A_{k+1}$  definita come in (2.17). Supponendo che la matrice Jacobiana soddisfi la condizione*

$$\|J(x) - J(x^*)\| \leq L\|x - x^*\| \quad \forall x \in \mathcal{D},$$

*allora, considerando la norma di Frobenius o  $l_2$ , vale la seguente formula:*

$$\|A_{k+1} - J(x^*)\| \leq \|A_k - J(x^*)\| + \frac{L}{2}(\|x_{k+1} - x^*\|_2 + \|x_k - x^*\|_2). \quad (2.18)$$

*Dimostrazione.* Considerando l'aggiornamento della matrice  $A_k$  descritto in (2.17) e sottraendo  $J(x^*)$  da entrambi i membri si ottiene

$$A_{k+1} - J(x^*) = A_k - J(x^*) + \frac{(y_k - A_k s_k)s_k^T}{s_k^T s_k},$$

da cui aggiungendo e sottraendo il termine  $\frac{J(x^*)s_k s_k^T}{s_k^T s_k}$  si arriva a

$$A_{k+1} - J(x^*) = A_k - J(x^*) + \frac{(J(x^*)s_k - A_k s_k)s_k^T}{s_k^T s_k} + \frac{(y_k - J(x^*)s_k)s_k^T}{s_k^T s_k}$$

$$= (A_k - J(x^*)) \left[ I - \frac{s_k s_k^T}{s_k^T s_k} \right] + \frac{(y_k - J(x^*) s_k) s_k^T}{s_k^T s_k}. \quad (2.19)$$

Applicando ora le proprietà della norma all'equazione appena ottenuta, si può scrivere

$$\|A_{k+1} - J(x^*)\| \leq \|A_k - J(x^*)\| \left\| I - \frac{s_k s_k^T}{s_k^T s_k} \right\|_2 + \frac{\|y_k - J(x^*) s_k\|_2}{\|s_k\|_2}.$$

Dato che  $I - \frac{s_k s_k^T}{s_k^T s_k}$  è un proiettore ortogonale (vedi Appendice A), grazie al Lemma A.2.3 si ha  $\left\| I - \frac{s_k s_k^T}{s_k^T s_k} \right\|_2 = 1$ . Inoltre dal Lemma A.3.3 per cui vale

$$\|y_k - J(x^*) s_k\|_2 \leq \frac{L}{2} (\|x_{k+1} - x^*\|_2 + \|x_k - x^*\|_2) \|s_k\|_2$$

si conclude la dimostrazione.  $\square$

La condizione (2.18) è detta di “deterioramento controllato” intendendo che anche se l'approssimazione  $A_k$  peggiorasse, questo avverrebbe sufficientemente piano da poter garantire comunque la convergenza del metodo Quasi-Newton.

### 2.2.2 Algoritmo

L'algoritmo del metodo di Broyden prevede due dati iniziali: l'approssimazione iniziale  $x_0 \in \mathbb{R}^n$  e  $A_0$  approssimazione della matrice  $J(x_0)$ . Per il calcolo del valore  $x_1$  si considera il modello lineare definito come in (2.13) calcolato in  $x$  e cioè

$$M_0(x) = F(x_0) + A_0(x - x_0),$$

e se ne cerca la radice, imponendo  $M_0(x) = 0$ . Questo implica la risoluzione del sistema lineare  $A_0(x - x_0) = -F(x_0)$ , che indicando  $s_0 = (x - x_0)$  si può esprimere

$$A_0 s_0 = -F(x_0).$$

Si ottiene perciò il passo di aggiornamento  $s_0$  attraverso cui si trova  $x_1$ :

$$x_1 = x_0 + s_0.$$

È poi necessario aggiornare la matrice  $A_0$  per l'iterazione seguente: si trova  $y_0 = F(x_1) - F(x_0)$  e la matrice  $A_1$  della forma (2.17). L'algoritmo è descritto in Algoritmo 2.2.1.

La buona definizione dell'algoritmo è legata all'invertibilità della matrice  $A_k$ . Inoltre notiamo che questo algoritmo dipende dalla scelta dell'approssimazione iniziale di  $J_0$ . Una possibilità è considerare il valore della Jacobiana nel punto  $x_0$  o una sua approssimazione quando disponibili; in alternativa si può prendere come  $A_0$  la matrice identità.

---

**Algoritmo 2.2.1:** Metodo di Broyden

---

**Dati:**  $x_0, A_0$

**Per**  $k = 0, 1, 2, \dots$

<b>Risolvi</b>	$A_k s_k = -F(x_k)$
<b>Poni</b>	$x_{k+1} = x_k + s_k$
<b>Calcola</b>	$y_k = F(x_{k+1}) - F(x_k)$
<b>Aggiorna</b>	$A_{k+1} = A_k + \frac{(y_k - A_k s_k) s_k^T}{s_k^T s_k}$

---

### 2.2.3 Analisi di convergenza

In questa sezione vengono analizzati i teoremi che garantiscono la convergenza della successione generata dal metodo di Broyden e la velocità con cui tale successione converge a  $x^*$ .

Per l'intera sezione si assume che  $x_{k+1} \neq x_k \forall k \geq 0$  e ciò, dato che le ipotesi garantiscono l'invertibilità di  $A_k$  (vedi Teorema 2.2.1) e  $x_{k+1} - x_k = -A_k^{-1}F(x_k)$ , è equivalente a dire che  $F(x_k) \neq 0$ . Facendo tale ipotesi si preclude perciò il caso banale in cui l'algoritmo calcoli la soluzione esatta  $x^*$ .

I successivi Teoremi 2.2.1 e 2.2.2 sono risultati che valgono per qualsiasi metodo Quasi-Newton. Per tale motivo sarà sufficiente assumere che i valori della successione  $\{x_k\}_{k \geq 0}$  siano della forma

$$x_{k+1} = x_k - A_k^{-1}F(x_k),$$

dove  $\{A_k\}_{k \geq 0}$  è una successione di matrici invertibili.

**Teorema 2.2.1.** *Supponendo che valgano le ipotesi standard I1, I2, I3 dove la condizione di Lipschitzianità in I3 è più debole, cioè vale*

$$\|J(x) - J(x^*)\| \leq L\|x - x^*\| \quad \forall x \in \mathcal{D},$$

*allora esistono le costanti  $\delta, \eta > 0$  tali che se*

$$x_0 \in B(x^*, \eta), \quad \|A_0 - J(x^*)\|_2 \leq \delta$$

*e la successione di matrici  $\{A_k\}_{k \geq 0}$  soddisfa la formula (2.18) allora la successione  $\{x_k\}_{k \geq 0}$  generata da un metodo Quasi-Newton è ben definita e converge almeno linearmente a  $x^*$ .*

*Dimostrazione.* Sia  $\|\cdot\|$  la norma vettoriale o matriciale  $l_2$ . Si indica  $e_k = x_k - x^*$ . Si scelgono le costanti  $\eta, \delta$  tali che

$$6\rho\delta \leq 1, \tag{2.20}$$

$$3L\eta \leq 2\delta. \tag{2.21}$$

La dimostrazione del teorema consiste nel mostrare per induzione su  $k$  la validità delle seguenti formule:

$$\|A_k - J(x^*)\| \leq (2 - 2^{-k})\delta, \quad (2.22)$$

$$\|e_{k+1}\| \leq \frac{\|e_k\|}{2}. \quad (2.23)$$

Per le ipotesi assunte, ad ogni iterazione vale la formula (2.18), che può essere scritta

$$\|A_k - J(x^*)\| \leq \|A_{k-1} - J(x^*)\| + \frac{L}{2}(\|e_k\| + \|e_{k-1}\|). \quad (2.24)$$

Per  $k = 0$ : la stima in (2.22) è verificata per ipotesi, infatti  $\|A_0 - J(x^*)\| \leq \delta \leq 2\delta$ , mentre la (2.23) si dimostra analogamente al caso induttivo.

Si suppone che le formule valgano fino a  $k - 1$  e si dimostrano per  $k$ :

Sostituendo l'ipotesi induttiva

$$\|A_{k-1} - J(x^*)\| \leq (2 - 2^{-(k-1)})\delta \quad \text{e} \quad \|e_k\| \leq \frac{\|e_{k-1}\|}{2},$$

all'interno della formula (2.24) si può scrivere

$$\|A_k - J(x^*)\| \leq (2 - 2^{-(k-1)})\delta + \frac{3}{4}L\|e_{k-1}\|,$$

dove grazie all'ipotesi  $\|e_0\| \leq \eta$  si sa che

$$\|e_{k-1}\| \leq \frac{\|e_{k-2}\|}{2} \leq \dots \leq 2^{-(k-1)}\eta,$$

e usando la formula (2.21) si ottiene la formula (2.22), infatti

$$\|A_k - J(x^*)\| \leq (2 - 2^{-(k-1)})\delta + \frac{3L}{4}2^{-(k-1)}\eta \leq (2 - 2^{-k})\delta.$$

Per la dimostrazione della formula (2.23), bisogna prima mostrare che  $A_k$  è invertibile in modo tale che l'iterazione sia ben definita.

Dall'ipotesi  $\|J(x^*)^{-1}\| \leq \rho$  e dalle formule (2.22), (2.20) si ottiene

$$\begin{aligned} \|J(x^*)^{-1}(A_k - J(x^*))\| &\leq \|J(x^*)^{-1}\| \|A_k - J(x^*)\| \\ &\leq \rho(2 - 2^{-k})\delta \leq 2\rho\delta \leq \frac{1}{3}. \end{aligned}$$

Così grazie al Lemma A.2.1,  $A_k$  è non singolare e

$$\|A_k^{-1}\| \leq \frac{\|J(x^*)^{-1}\|}{1 - \|J(x^*)^{-1}(A_k - J(x^*))\|} \leq \frac{\rho}{1 - \frac{1}{3}} = \frac{3}{2}\rho. \quad (2.25)$$



Come conseguenza si ottiene la buona definizione di  $x_{k+1}$ . Infatti

$$x_{k+1} = x_k - A_k^{-1}F(x_k),$$

da cui sottraendo  $x^*$  ad entrambi i membri e considerando che  $F(x^*) = 0$ , si ha

$$x_{k+1} - x^* = x_k - x^* - A_k^{-1}(F(x_k) - F(x^*)),$$

che equivale a

$$A_k e_{k+1} = A_k e_k - F(x_k) + F(x^*),$$

da cui aggiungendo e sottraendo  $J(x^*)e_k$  nel membro destro, si ottiene

$$A_k e_{k+1} = [-F(x_k) + F(x^*) + J(x^*)e_k] + (A_k - J(x^*))e_k.$$

Passando alle norme, dall'equazione appena ottenuta si giunge alla seguente stima:

$$\|e_{k+1}\| \leq \|A_k^{-1}\| \left[ \|-F(x_k) + F(x^*) + J(x^*)e_k\| + \|A_k - J(x^*)\| \|e_k\| \right]. \quad (2.26)$$

Grazie al Lemma 2.1.1, si ottiene

$$\|-F(x_k) + F(x^*) + J(x^*)e_k\| \leq \frac{L\|e_k\|^2}{2},$$

e sostituendo tale stima e le formule (2.22) e (2.25) in (2.26) si arriva a

$$\|e_{k+1}\| \leq \frac{3}{2}\rho \left[ \frac{L}{2}\|e_k\| + (2 - 2^{-k})\delta \right] \|e_k\|. \quad (2.27)$$

Per ipotesi induttiva,  $\|e_0\| \leq \eta$ , e per la (2.21) si ha:

$$\frac{L\|e_k\|}{2} \leq 2^{-(k+1)}L\eta \leq \frac{2^{-k}\delta}{3},$$

che sostituita in (2.27) e utilizzando la (2.20) si ottiene la formula (2.23), infatti

$$\|e_{k+1}\| \leq \frac{3}{2}\rho \left[ \frac{1}{3}2^{-k} + 2 - 2^{-k} \right] \delta \|e_k\| \leq 3\rho\delta \|e_k\| \leq \frac{\|e_k\|}{2}.$$

□

Questo risultato implica la convergenza lineare della successione  $\{x_k\}_{k \geq 0}$  sotto le condizioni prima indicate, ma esistono dei casi in cui la convergenza è più veloce e cioè superlineare. Per dimostrare ciò è fondamentale il seguente risultato:

**Teorema 2.2.2.** [DS, Teorema 8.2.4 (Dennis-Moré)] *Siano  $\mathcal{D} \subset \mathbb{R}^n$  un insieme aperto convesso,  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $J \in Lip_L(\mathcal{D})$ ,  $x^* \in \mathcal{D}$  e  $J(x^*)$  non singolare. Si suppone che per qualche  $x_0 \in \mathcal{D}$  la successione  $\{x_k\}_{k \geq 0}$  generata da un metodo Quasi-Newton sia interamente contenuta in  $\mathcal{D}$ . Inoltre si assume che valgano  $x_k \neq x^*$ , per ogni  $k$ , e  $\lim_{k \rightarrow \infty} x_k = x^*$ . Allora la successione  $\{x_k\}_{k \geq 0}$  converge superlinearmente a  $x^*$  in una qualche norma  $\|\cdot\|$  e  $F(x^*) = 0$  se e solo se*

$$\lim_{k \rightarrow \infty} \frac{\|(A_k - J(x^*))s_k\|}{\|s_k\|} = 0, \quad (2.28)$$

dove  $s_k = x_{k+1} - x_k$ .

*Dimostrazione.* Si definisce  $e_k = x_k - x^*$ . Supponendo che valga la formula (2.28) si dimostra che  $F(x^*) = 0$  e che  $\{x_k\}_{k \geq 0}$  converge superlinearmente a  $x^*$ . Considerando l'algoritmo dei metodi Quasi-Newton si ha:

$$A_k s_k + F(x_k) = 0,$$

da cui aggiungendo e sottraendo  $J(x^*)s_k$  si ottiene

$$(A_k - J(x^*))s_k + F(x_k) + J(x^*)s_k = 0.$$

Per cui sottraendo  $F(x_{k+1})$  da entrambi i membri l'uguaglianza diventa

$$-F(x_{k+1}) = (A_k - J(x^*))s_k + [-F(x_{k+1}) + F(x_k) + J(x^*)s_k]. \quad (2.29)$$

Applicando ora le proprietà della norma, dividendo per  $\|s_k\|$  e utilizzando il Lemma A.3.3 si arriva alla seguente disuguaglianza

$$\begin{aligned} \frac{\|F(x_{k+1})\|}{\|s_k\|} &\leq \frac{\|(A_k - J(x^*))s_k\|}{\|s_k\|} + \frac{\|-F(x_{k+1}) + F(x_k) + J(x^*)s_k\|}{\|s_k\|} \\ &\leq \frac{\|(A_k - J(x^*))s_k\|}{\|s_k\|} + \frac{L}{2}(\|e_k\| + \|e_{k+1}\|). \end{aligned}$$

Per ipotesi vale  $\lim_{k \rightarrow \infty} \|e_k\| = 0$ , inoltre usando la formula (2.28) e la disuguaglianza appena ricavata si ottiene

$$\lim_{k \rightarrow \infty} \frac{\|F(x_{k+1})\|}{\|s_k\|} = 0. \quad (2.30)$$

Inoltre per la continuità di  $F$  si ricava:

$$F(x^*) = \lim_{k \rightarrow \infty} F(x_k) = 0.$$

Dato che  $F(x^*)=0$  e grazie al Lemma A.3.4 esistono  $\alpha > 0$ ,  $k_0 \geq 0$  tali che

$$\|F(x_{k+1})\| = \|F(x_{k+1}) - F(x^*)\| \geq \alpha \|e_{k+1}\|, \quad (2.31)$$

per ogni  $k \geq k_0$ . Partendo da quanto si è ottenuto in (2.30), utilizzando la (2.31) e la relazione  $\|s_k\| \leq \|e_{k+1}\| + \|e_k\|$ , si ottiene

$$\begin{aligned} 0 &= \lim_{k \rightarrow \infty} \frac{\|F(x_{k+1})\|}{\|s_k\|} \geq \lim_{k \rightarrow \infty} \alpha \frac{\|e_{k+1}\|}{\|s_k\|} \\ &\geq \lim_{k \rightarrow \infty} \alpha \frac{\|e_{k+1}\|}{\|e_k\| + \|e_{k+1}\|} = \lim_{k \rightarrow \infty} \frac{\alpha r_k}{1 + r_k}, \end{aligned}$$

dove  $r_k = \frac{\|e_{k+1}\|}{\|e_k\|}$ . Questo risultato implica che

$$\lim_{k \rightarrow \infty} r_k = 0,$$

e quindi la convergenza superlineare.

Viceversa, supponendo che  $F(x^*) = 0$  e che la successione  $\{x_k\}_{k \geq 0}$  converga superlinearmente a  $x^*$  si dimostra la (2.28).

Siccome  $F(x^*) = 0$ , dal Lemma A.3.4 si ha l'esistenza di  $\beta > 0$ ,  $k_0 \geq 0$ , tali che

$$\|F(x_{k+1})\| = \|F(x_{k+1}) - F(x^*)\| \leq \beta \|e_{k+1}\|,$$

per ogni  $k \geq k_0$ . Sfruttando l'ipotesi della convergenza superlineare e la stima appena ottenuta si arriva a

$$0 = \lim_{k \rightarrow \infty} \frac{\|e_{k+1}\|}{\|e_k\|} \geq \lim_{k \rightarrow \infty} \frac{\|F(x_{k+1})\|}{\beta \|e_k\|} = \lim_{k \rightarrow \infty} \frac{1}{\beta} \frac{\|F(x_{k+1})\|}{\|s_k\|} \frac{\|s_k\|}{\|e_k\|}.$$

Siccome per il Lemma A.1.1 vale  $\lim_{k \rightarrow \infty} \frac{\|s_k\|}{\|e_k\|} = 1$ , la formula sopra ricavata implica la (2.30).

Applicando le proprietà della norma alla formula (2.29), dividendo per  $\|s_k\|$  e utilizzando il Lemma A.3.3 si ottiene

$$\begin{aligned} \frac{\|(A_k - J(x^*))s_k\|}{\|s_k\|} &\leq \frac{\|F(x_{k+1})\|}{\|s_k\|} + \frac{\| -F(x_{k+1}) + F(x_k) + J(x^*)s_k \|}{\|s_k\|} \\ &\leq \frac{\|F(x_{k+1})\|}{\|s_k\|} + \frac{L}{2} (\|e_k\| + \|e_{k+1}\|), \end{aligned}$$

da cui utilizzando la (2.30) e la proprietà  $\lim_{k \rightarrow \infty} \|e_k\| = 0$  si ottiene la tesi.  $\square$

Il teorema appena dimostrato descrive una condizione necessaria e sufficiente affinché ci sia una convergenza di tipo superlineare per un metodo Quasi-Newton.

È importante sottolineare che la condizione (2.28) non richiede che le matrici  $A_k$  convergano alla Jacobiana  $J(x^*)$ , ma esprime una condizione più debole.

Si può notare che siccome  $J$  è lipschitziana la formula (2.28) può essere scritta anche nel modo seguente:

$$\lim_{k \rightarrow \infty} \frac{\|(A_k - J(x_k))s_k\|}{\|s_k\|} = 0. \quad (2.32)$$

Un aspetto interessante è dato dall'interpretazione della formula (2.32). Infatti dato che  $s_k = -A_k^{-1}F(x_k)$ , la (2.32) è equivalente a

$$\lim_{k \rightarrow \infty} \frac{\|J(x_k)(s_k^N - s_k)\|}{\|s_k\|} = 0,$$

dove  $s_k^N = -J(x_k)^{-1}F(x_k)$  è il passo di Newton, e cioè il valore che si ottiene dalla risoluzione del sistema lineare alla  $k$ -esima iterazione del metodo di Newton. Quindi condizione necessaria e sufficiente affinché un metodo Quasi-Newton converga superlinearmente è che il passo  $s_k = -A_k^{-1}F(x_k)$  converga in norma e in direzione al passo di Newton  $s_k^N = -J(x_k)^{-1}F(x_k)$ .

Sfruttando ora il Teorema 2.2.2 si riesce a dimostrare la convergenza superlineare della successione generata dal metodo di Broyden, infatti:

**Teorema 2.2.3.** *Supponendo che valgano le ipotesi standard I1, I2, I3 allora esistono le costanti  $\delta, \eta > 0$  tali che se*

$$x_0 \in B(x^*, \eta) \quad e \quad \|A_0 - J(x^*)\|_2 \leq \delta,$$

*allora la successione  $\{x_k\}_{k \geq 0}$  generata dal metodo di Broyden converge superlinearmente a  $x^*$ .*

*Dimostrazione.* Si definisce  $E_k = A_k - J(x^*)$  e sia  $\|\cdot\|$  la norma vettoriale  $l_2$ . Si nota che le ipotesi assunte nel Teorema 2.2.1 risultano essere più deboli rispetto a quelle attuali, per tale motivo i risultati ottenuti nel Teorema 2.2.1 rimangono validi. Così per dimostrare questo teorema è sufficiente mostrare che vale la condizione espressa in (2.28), la quale può essere scritta come

$$\lim_{k \rightarrow \infty} \frac{\|E_k s_k\|}{\|s_k\|} = 0. \quad (2.33)$$

Applicando la norma di Frobenius e le sue proprietà alla (2.19) si ottiene

$$\|E_{k+1}\|_F \leq \left\| E_k \left( I - \frac{s_k s_k^T}{s_k^T s_k} \right) \right\|_F + \frac{\|(y_k - J(x^*)s_k) s_k^T\|_F}{s_k^T s_k}. \quad (2.34)$$

All'interno della dimostrazione del Lemma 2.2.2, si è mostrato che

$$\left\| \frac{(y_k - J(x^*)s_k) s_k^T}{s_k^T s_k} \right\|_F \leq \frac{L}{2} (\|e_k\| + \|e_{k+1}\|).$$

Utilizzando questa stima, la (2.23) e il Lemma A.2.2, la (2.34) diventa

$$\|E_{k+1}\|_F \leq \|E_k\|_F - \frac{\|E_k s_k\|^2}{2\|E_k\|_F \|s_k\|^2} + \frac{3L}{4} \|e_k\|,$$

o equivalentemente

$$\frac{\|E_k s_k\|^2}{\|s_k\|^2} \leq 2\|E_k\|_F \left[ \|E_k\|_F - \|E_{k+1}\|_F + \frac{3L}{4} \|e_k\| \right].$$

Grazie alla dimostrazione del Teorema 2.2.1 si ha che  $\|E_k\|_F \leq 2\delta$  per ogni  $k \geq 0$  e  $\sum_{k=0}^{\infty} \|e_k\| \leq 2\eta$ . Introducendo questi risultati all'interno dell'ultima disuguaglianza si arriva a

$$\frac{\|E_k s_k\|^2}{\|s_k\|^2} \leq 4\delta \left[ \|E_k\|_F - \|E_{k+1}\|_F + \frac{3L}{4} \|e_k\| \right].$$

Facendo la somma in entrambi i membri per  $k$  da 0 a  $i$ , si ottiene

$$\begin{aligned} \sum_{k=0}^i \frac{\|E_k s_k\|^2}{\|s_k\|^2} &\leq 4\delta \left[ \|E_0\|_F - \|E_{i+1}\|_F + \frac{3L}{4} \sum_{k=0}^i \|e_k\| \right] \\ &\leq 4\delta \left[ \|E_0\|_F + \frac{3L\eta}{2} \right] \leq 4\delta \left[ \delta + \frac{3L\eta}{2} \right]. \end{aligned}$$

Siccome tale stima è verificata per ogni  $i \geq 0$ , si ha

$$\sum_{k=0}^{\infty} \frac{\|E_k s_k\|^2}{\|s_k\|^2} < \infty$$

e ciò implica la validità del limite (2.33) da cui la tesi.  $\square$



## Capitolo 3

# Implementazione Matlab

Questo capitolo è dedicato allo studio del comportamento pratico del metodo di Newton  $n$ -dimensionale e del metodo di Broyden con l'obiettivo di mostrare alcune caratteristiche dei metodi messe in evidenza nell'analisi teorica descritta nel capitolo precedente.

### 3.1 Criteri d'arresto

In questo paragrafo vengono analizzati alcuni possibili criteri d'arresto. È però necessario sottolineare che non esiste un criterio d'arresto adatto ad ogni situazione, ma la scelta dipende dal metodo risolutivo che si sta utilizzando e il problema da trattare.

Tutti i criteri prevedono l'introduzione di valori di soglia che vengono chiamati tolleranze e quantificano le richieste di accuratezza. I seguenti criteri sono descritti per il caso  $n$ -dimensionale, ma sono analoghi per  $n = 1$ . Nei criteri d'arresto in cui è richiesto l'utilizzo della norma generalmente si intende la  $\|\cdot\|_2$  o la  $\|\cdot\|_\infty$ .

Un criterio possibile è detto criterio di salvaguardia e agisce in modo tale che il procedimento non superi un numero massimo di iterazioni, indicato con  $K_{max}$ . Tale criterio assicura la terminazione dell'algoritmo ma non garantisce che il valore trovato sia una buona approssimazione della soluzione del problema.

Un altro criterio d'arresto adatto per gli algoritmi sulla ricerca di zeri di sistemi non lineari è il seguente:

$$\|F(x_k)\| < tollF, \quad (3.1)$$

dove  $tollF$  è la tolleranza. Tale criterio provoca perciò l'arresto all'iterata  $k$ -esima se si trova un vettore  $x_k$  tale che la quantità  $\|F(x_k)\|$  sia prossima allo zero e cioè  $F(x_k)$  vicino a  $F(x^*)$ . In generale però questo non implica la

vicinanza tra i vettori  $x_k$  e  $x^*$  che dipende dal comportamento della funzione  $F$  intorno a  $x^*$ .

Il lemma seguente permette l'introduzione di un criterio che evita l'arresto dell'algoritmo a un vettore  $x_k$  distante da  $x^*$ .

**Lemma 3.1.1.** *Supponendo che valgano le ipotesi standard I1, I2, I3 assunte per i sistemi non lineari, sia  $\delta > 0$  tale che per ogni  $x \in B(x^*, \delta)$  valga*

$$\frac{\|e\|}{2\|J(x^*)^{-1}\|} \leq \|F(x)\| \leq 2\|J(x^*)\| \|e\|, \quad (3.2)$$

dove  $e = x - x^*$ . Allora per ogni  $x \in B(x^*, \delta)$  vale

$$\frac{\|e\|}{4\|e_0\| \|J(x^*)\| \|J(x^*)^{-1}\|} \leq \frac{\|F(x)\|}{\|F(x_0)\|} \leq \frac{4\|J(x^*)\| \|J(x^*)^{-1}\| \|e\|}{\|e_0\|},$$

dove  $e_0 = x_0 - x^*$ .

*Dimostrazione.* Applicando la formula (3.2) alla quantità  $\frac{1}{\|F(x_0)\|}$ , si ottiene:

$$\frac{1}{2\|J(x^*)\| \|e_0\|} \leq \frac{1}{\|F(x_0)\|} \leq \frac{2\|J(x^*)^{-1}\|}{\|e_0\|}.$$

Facendo il rapporto tra le disuguaglianze in (3.2) e quelle appena ottenute si giunge alla tesi.  $\square$

Da questo lemma si conclude che se  $J(x^*)$  è ben condizionata, allora la quantità  $\frac{\|F(x)\|}{\|F(x_0)\|}$  risulta essere una buona approssimazione per la stima dell'errore.

Su questo risultato si basa il seguente criterio d'arresto:

$$\|F(x_k)\| < toll_r \|F(x_0)\| + toll_a, \quad (3.3)$$

dove  $toll_r$  e  $toll_a$  sono rispettivamente la tolleranza relativa e assoluta. Questo criterio ha la proprietà di adattarsi alle diverse situazioni, infatti se la quantità  $\|F(x_0)\|$  è piccola allora predomina il termine  $toll_a$  e il criterio d'arresto (3.3) equivale a (3.1) risultando così un criterio assoluto, mentre se la quantità  $\|F(x_0)\|$  è grande allora predomina il termine  $\|F(x_0)\|toll_r$  diventando un criterio relativo.

Per quanto riguarda i metodi che hanno una velocità di convergenza almeno superlineare esiste un ulteriore criterio d'arresto basato sulla vicinanza di due iterate successive. Infatti, grazie al Lemma A.1.1, l'errore di approssimazione  $\|x_k - x^*\|$  di cui ogni iterata è affetto, da un certo punto in poi può essere approssimato con la quantità  $\|x_k - x_{k-1}\|$ .

Il criterio d'arresto è il seguente:

$$\|x_k - x_{k-1}\| \leq toll_r \|x_k\| + toll_a.$$

Si può notare che, come il precedente criterio, anche questo è un criterio misto, ossia a seconda della quantità predominante esso agisce come un criterio d'arresto assoluto o relativo.



## 3.2 Implementazione degli algoritmi

In questa sezione vengono descritte in modo dettagliato le implementazioni Matlab degli algoritmi dei metodi analizzati nel Capitolo 2 che sono riportate di seguito in Algoritmo 3.2.1 e 3.2.2.

---

**Algoritmo 3.2.1:** Algoritmo del metodo di Newton

---

**Dati:**  $x_0, K_{max}, toll_a, toll_r$

1. Calcola  $F = F(x_0)$ ;
  2. Poni  $F_0 = F$ ;
  3. Poni  $it = 0$ ;
  4. Se  $\|F_0\| < toll_a$ , termina;
  5. **Fintanto che**  $it < K_{max}$  **o**  $\|F\| < toll_r\|F_0\| + toll_a$ :
    - (a) Calcola  $J = J(x)$ ;
    - (b) Risolvi  $Js = -F$ ;
    - (c) Poni  $x = x + s$ ;
    - (d) Calcola  $F = F(x)$ ;
    - (e) Poni  $it = it + 1$ .
- 

Negli esperimenti riportati nei paragrafi successivi è stato considerato  $K_{max} = 100$  e, per il criterio d'arresto (3.3), sono state usate le seguenti tolleranze:  $toll_a = 10^{-12}$ ,  $toll_r = 10^{-6}$ .

La condizione al passo (5) dei due algoritmi è realizzata mediante il comando **while**.

Per il metodo di Newton, la risoluzione del sistema lineare espresso nel punto (5b) è stata implementata attraverso l'utilizzo del *backslash*. Invece, per lo svolgimento del punto (5a) del metodo di Broyden sono state considerate due possibilità implementative. La prima, analoga a quella utilizzata nel metodo di Newton, ossia la risoluzione del sistema lineare attraverso l'utilizzo del *backslash*. La seconda possibilità invece consente di evitare la risoluzione di un sistema lineare che rappresenta un passo molto costoso nell'algoritmo. Questo è possibile grazie al lemma seguente, il quale fornisce un'espressione esplicita e di semplice implementazione per il calcolo dell'inversa della matrice dei coefficienti del sistema lineare.

---

**Algoritmo 3.2.2:** Algoritmo del metodo di Broyden

---

**Dati:**  $x_0, A_0, K_{max}, toll_a, toll_r$

1. Calcola  $F = F(x_0), A = A_0$ ;
  2. Poni  $F_{old} = F, F_0 = F$ ;
  3. Poni  $it = 0$ ;
  4. Se  $\|F_0\| < toll_a$ , termina;
  5. **Fintanto che**  $it < K_{max}$  **o**  $\|F\| < toll_r \|F_0\| + toll_a$ :
    - (a) Risolvi  $As = -F$ ;
    - (b) Poni  $F_{old} = F$ ;
    - (c) Poni  $x = x + s$ ;
    - (d) Calcola  $F = F(x)$ ;
    - (e) Calcola  $y = F - F_{old}$ ;
    - (f) Aggiorna  $A = A - \frac{(y-As)s^T}{s^T s}$ ;
    - (g) Poni  $it = it + 1$ .
- 

**Lemma 3.2.1.** [DS, Lemma 8.3.1 (Sherman-Morrison-Woodbury)] *Siano  $u, v \in \mathbb{R}^n$  e  $A \in \mathbb{R}^{n \times n}$  non singolare. Allora la matrice  $A + uv^T$  è non singolare se e solo se*

$$1 + v^T A^{-1} u \neq 0.$$

*Inoltre*

$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1} uv^T A^{-1}}{1 + v^T A^{-1} u}. \quad (3.4)$$

Considerando l'applicazione del Lemma 3.2.1 al procedimento del metodo di Broyden descritto nel Capitolo 2 e ricordandosi l'aggiornamento della matrice  $A_k$  espresso in (2.17), la formula (3.4) può quindi essere riscritta così:

$$A_{k+1}^{-1} = A_k^{-1} + \frac{(s_k - A_k^{-1} y_k) s_k^T A_k^{-1}}{s_k^T A_k^{-1} y_k}. \quad (3.5)$$

Perciò il Lemma 3.2.1, sotto opportune ipotesi, garantisce l'invertibilità della matrice  $A_{k+1}$  e suggerisce la formula (3.5) per il calcolo dell'inversa che necessita della sola conoscenza di  $A_k^{-1}$ , nota dall'iterazione precedente. Quindi, come già anticipato, la risoluzione del sistema lineare  $A_k s_k = -F(x_k)$  si può sostituire con un prodotto tra la matrice  $A_k^{-1}$  per il vettore  $-F(x_k)$ . L'utilizzo di questo risultato è molto importante soprattutto quando  $n$  è grande, in quanto consente di ridurre il numero di operazioni elementari da svolgere

all'interno di ogni iterazione e quindi la diminuzione del tempo di esecuzione del codice Matlab.

Grazie al risultato espresso dal Lemma 3.2.1, l'algoritmo del metodo di Broyden può essere trasformato nel seguente modo:

Dato  $B_0$  approssimazione iniziale di  $J(x_0)^{-1}$ , i punti (5a) e (5f) diventano:

$$(5a) \quad \text{Poni } s = -BF;$$

$$(5f) \quad \text{Aggiorna } B = B - \frac{(s - By)s^T B}{s^T By}.$$

### 3.3 Problema test

Si consideri il seguente sistema di due equazioni in due incognite:

$$\begin{cases} x^2 + y^2 - 2 = 0 \\ e^{x-1} + y^3 - 2 = 0, \end{cases}$$

che ha due radici  $x_1^* = (1, 1)^T$  e  $x_2^*$ , come è rappresentato in Figura 3.1.

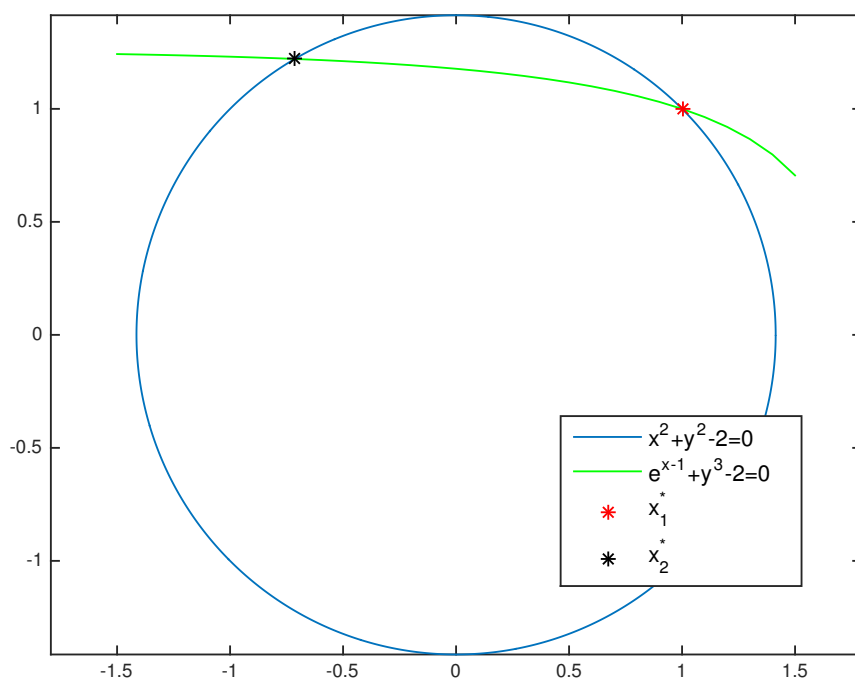


Figura 3.1

**A) Confronto tra il metodo di Newton e il metodo di Broyden**

Test 1) Considerando come dati iniziali  $x_0 = (1.2, 1.2)^T$  e  $A_0 = J(x_0)$ , i metodi producono i risultati contenuti nelle Tabelle 3.1 e 3.2.

$k$	$x_k$	$\ x_k - x_{k+1}\ $	$\ F(x_k)\ $	$\ e_k\ $
0	[1.2000; 1.2000]	2.61044025e-01	1.29451365e+00	2.82842712e-01
1	[0.9952; 1.0381]	3.66218326e-02	1.32842213e-01	3.84358759e-02
2	[0.9990; 1.0017]	1.99384442e-03	4.30572275e-03	1.99780033e-03
3	[1.0000; 1.0000]	3.98533060e-06	1.00675747e-05	3.98534914e-06
4	[1.0000; 1.0000]	1.85589056e-11	4.38745945e-11	1.85587560e-11

Tabella 3.1: Metodo di Newton per il Test 1).

$k$	$x_k$	$\ x_k - x_{k+1}\ $	$\ F(x_k)\ $	$\ e_k\ $
0	[1.2000; 1.2000]	2.61044025e-01	1.29451365e+00	3.84358759e-02
1	[0.9952; 1.0381]	2.76752806e-02	1.32842213e-01	1.31734800e-02
2	[0.9922; 1.0106]	1.08745458e-02	2.51284509e-02	2.42007555e-03
3	[0.9981; 1.0015]	2.25638781e-03	2.64108795e-03	1.64154298e-04
4	[0.9999; 1.0001]	1.61928983e-04	1.62560864e-04	2.22748666e-06
5	[1.0000; 1.0000]	2.21087499e-06	2.15812908e-06	1.66502138e-08

Tabella 3.2: Metodo di Broyden per il Test 1).

Si può notare che entrambi i metodi convergono alla soluzione  $x_1^*$  che è nota ed è quindi possibile calcolare l'errore  $e_k$  ad ogni iterazione. Inoltre dall'ordine di grandezza dell'errore si può vedere che la convergenza nel metodo di Newton è quadratica mentre nel metodo di Broyden è superlineare.

Test 2) Prendendo ora come dati iniziale  $x_0 = (0, 1)^T$  e  $A_0 = J(x_0)$  si ottengono i risultati riportati nelle Tabelle 3.3 e 3.4.

In questo caso si può notare che entrambi i metodi convergono alla soluzione  $x_2^*$ , il cui valore esatto non è noto e quindi non risulta possibile calcolare l'errore  $e_k$ . In ogni caso, i valori di  $\|F(x_k)\|$  e di  $\|x_k - x_{k+1}\|$  confermano la velocità di convergenza quadratica e superlineare dei due metodi (vedi Lemma 3.1.1 e Lemma A.1.1).

Test 3) Assumendo di avere come dati iniziali  $x_0 = (7, 7)^T$  e  $A_0 = J(x_0)$ , si ottengono i risultati esposti in Tabella 3.5.

Il metodo di Newton converge alla soluzione  $x_2^*$  come si vede in Tabella 3.5, invece il metodo di Broyden non converge ma si allontana dalla soluzione interrompendosi perché trova un  $A_k$  vicino alla singolarità.

$k$	$x_k$	$\ x_k - x_{k+1}\ $	$\ F(x_k)\ $
0	[0; 1.0000]	2.41154429e+00	1.18303694e+00
1	[-2.3591; 1.5000]	1.11696119e+00	5.98397955e+00
2	[-1.2630; 1.2855]	4.41158086e-01	1.26832665e+00
3	[-0.8253; 1.2303]	1.05124379e-01	1.95997802e-01
4	[-0.7205; 1.2214]	6.77670515e-03	1.11164042e-02
5	[-0.7138; 1.2209]	2.83428538e-05	4.62123401e-05
6	[-0.7137; 1.2209]	4.95486787e-10	8.08470102e-10

Tabella 3.3: Metodo di Newton per il Test 2).

$k$	$x_k$	$\ x_k - x_{k+1}\ $	$\ F(x_k)\ $
0	[0; 1.0000]	2.41154429e+00	1.18303694e+00
1	[-2.3591; 1.5000]	2.15701005e+00	5.98397955e+00
2	[-0.2248; 1.1878]	2.11629251e-01	5.39359169e-01
3	[-0.4364; 1.1828]	3.61764901e-01	4.24522078e-01
4	[-0.7919; 1.2500]	8.55789336e-02	2.24422595e-01
5	[-0.7145; 1.2134]	1.49551922e-02	3.77034792e-02
6	[-0.7046; 1.2245]	8.98897020e-03	1.83459378e-02
7	[-0.7129; 1.2211]	8.75560490e-04	1.32234869e-03
8	[-0.7137; 1.2209]	2.61231501e-05	4.58317926e-05

Tabella 3.4: Metodo di Broyden per il Test 2).

Come previsto dall'analisi di convergenza descritta nel Capitolo 2, dai tre casi riportati in questa sezione si può facilmente notare l'importanza della scelta del dato iniziale  $x_0$ , che condiziona fortemente la convergenza alla soluzione.

## B) Note sulla performance del metodo di Broyden

In questa sezione viene analizzato il comportamento della successione di matrici  $\{A_k\}_{k \geq 0}$  alla luce del Teorema 2.2.2.

Nei prossimi esempi verranno studiate le quantità

$$a_k := \frac{\|A_k - J(x^*)\|}{\|J(x^*)\|} \quad \text{e} \quad d_k := \frac{\|(A_k - J(x^*))s_k\|}{\|s_k\|},$$

dove  $a_k$  rappresenta l'errore relativo dell'approssimazione  $A_k$  rispetto a  $J(x^*)$  e  $d_k$  è la quantità che deve tendere a zero affinché il metodo di Broyden converga superlinearmente per il Teorema 2.2.2.

$k$	$x_k$	$\ x_k - x_{k+1}\ $	$\ F(x_k)\ $
0	[7.0000; 7.0000]	7.95170727e+00	7.50593251e+02
1	[8.0279; -0.8850]	2.62000283e+01	1.12668996e+03
2	[6.9757; 25.2939]	7.00750883e+01	1.65885089e+04
3	[-62.8678; 30.9864]	3.54994084e+01	3.01523863e+04
4	[-28.9040; 20.6583]	1.82293535e+01	8.90388279e+03
5	[-12.0247; 13.7738]	9.71431761e+00	2.63217237e+03
6	[-3.4619; 9.1860]	6.31282981e+00	7.78893717e+02
7	[2.0628; 6.1316]	4.32123978e+00	2.34829554e+02
8	[-1.7920; 4.1787]	2.45311417e+00	7.34432452e+01
9	[0.2507; 2.8204]	2.87916353e+00	2.17578007e+01
10	[-2.5087; 1.9990]	1.34789890e+00	1.02433411e+01
11	[-1.2591; 1.4939]	5.09084397e-01	2.31715317e+00
12	[-0.8009; 1.2719]	9.74387668e-02	3.41712716e-01
13	[-0.7165; 1.2231]	3.58527081e-03	1.34740840e-02
14	[-0.7137; 1.2209]	4.62035583e-06	2.30564270e-05

Tabella 3.5: Metodo di Newton per il Test 3).

Test i) Partendo dai dati iniziali  $x_0 = (1.2, 1.2)^T$  e  $A_0 = J(x_0)$ , rimangono validi i dati della Tabella 3.2 mentre le quantità di  $a_k$  e  $d_k$  sono riportate in Tabella 3.6.

$k$	$a_k$	$d_k$
0	2.61108789e-01	6.31589359e-01
1	7.38769461e-02	1.54073596e-01
2	4.46071961e-02	3.66567165e-02
3	4.18402979e-02	4.48138275e-03
4	4.17289451e-02	2.65666539e-04
5	4.16893134e-02	3.41433275e-06

Tabella 3.6: Risultati del Test i).

Dai risultati ottenuti si può vedere che i valori  $a_k$  sono dell'ordine di  $10^{-2}$ , mentre la quantità  $d_k$  tende a zero più velocemente.

Test ii) Prendendo ora come dati iniziali  $x_0 = (0.5, 0.5)^T$  e  $A_0 = J(x_0)$ , si ottengono i risultati in Tabella 3.7.

I valori di  $a_k$  raggiungono valori dell'ordine di  $10^{-1}$ , mentre quelli di  $d_k$  tendono a zero.

A conferma del Teorema 2.2.2, osserviamo che la convergenza  $A_k$  a  $J(x^*)$  (quando c'è) può essere molto lenta, mentre è molto più rapida la conver-

$k$	$x_k$	$\ x_k - x_{k+1}\ $	$\ F(x_k)\ $	$\ e_k\ $
0	[0.5000; 0.5000]	2.69258240e+00	1.96443744e+00	2.50000000e+00
1	[-0.5000; 3.0000]	2.85722294e+00	2.62444050e+01	4.73390037e-01
2	[1.4680; 0.9286]	7.72370076e-01	1.09205733e+00	4.20097195e-01
3	[0.7402; 0.6699]	3.45897739e-01	1.36686749e+00	1.89720953e-01
4	[1.0529; 0.8178]	6.82460004e-02	4.56702808e-01	1.57966415e-01
5	[1.0935; 0.8727]	3.06909140e-02	2.41279700e-01	1.28633016e-01
6	[1.0694; 0.8917]	1.66382927e-01	2.27529020e-01	3.80997491e-02
7	[0.9846; 1.0349]	4.52348773e-02	1.01397378e-01	7.14837285e-03
8	[1.0033; 0.9937]	6.71935530e-03	1.66782294e-02	4.29877504e-04
9	[1.0002; 0.9996]	4.23964412e-04	9.28299842e-04	6.15337408e-06
10	[1.0000; 1.0000]	5.57775938e-06	8.09113441e-06	5.80722715e-07

$k$	$a_k$	$d_k$
0	1.90637092e+00	7.72569872e+00
1	1.85095757e+00	7.50501358e+00
2	1.80562114e+00	2.38916402e-01
3	1.80964839e+00	7.71869037e-01
4	1.67920470e+00	6.62960593e-01
5	2.50424717e-01	5.51944563e-01
6	1.62157910e-01	2.15803656e-01
7	1.49147201e-01	8.70363579e-02
8	1.46522163e-01	2.01119739e-02
9	1.46461089e-01	1.10572605e-03
10	1.40188413e-01	1.30196043e-05

Tabella 3.7: Risultati del Test ii).

genza a zero di  $d_k$ . Per gli esempi in (A) e in (B), si ottengono risultati analoghi se  $A_0 = I$ .

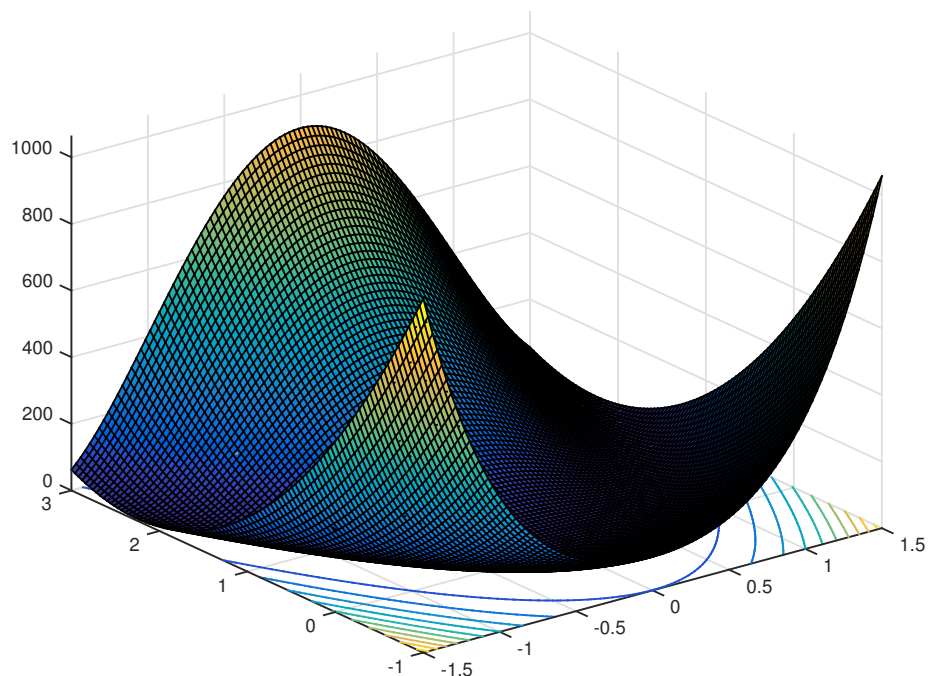
### C) Confronto tra le implementazioni del metodo di Broyden

Si considera ora un esempio [MGH] in cui sono stati analizzati i tempi di esecuzione dell'algoritmo di Broyden, implementato attraverso entrambe le opzioni descritte nella Sezione 3.2.

La funzione analizzata è la funzione di Rosenbrock (“funzione a banana”), molto famosa in letteratura e rappresentata in Figura 3.2 nel caso bidimensionale.

Per  $n \geq 2$ ,  $n \in \mathbb{N}$ , la funzione di Rosenbrock  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  è definita nel seguente modo:

$$\begin{aligned}
 F^{2i-1}(x) &= 10(x^{2i} - (x^{i-1})^2) \\
 F^{2i}(x) &= 1 - x^{2i-1},
 \end{aligned}$$

Figura 3.2: Funzione di Rosenbrock per  $n = 2$ .

per  $i = 1, \dots, n$ .

Nella Tabella 3.8 vengono riportati i risultati dei tempi di esecuzione in secondi ottenuti attraverso l'implementazione del metodo di Broyden. Le denotazioni delle colonne "Broyden" e "Broyden SM" indicano rispettivamente l'utilizzo del metodo di Broyden con il *backslash* e con la formula (3.5).

Mettendo a confronto le due opzioni implementative si può notare che, più  $n$  è grande, tanto più l'utilizzo della formula (3.5) riduce i tempi di esecuzione rispetto all'utilizzo del *backslash*. Infatti, il numero di iterazioni svolte è analogo e i tempi di esecuzione nella colonna "Broyden SM" sono minori rispetto a quelli della colonna "Broyden". Questo conferma quanto descritto in Sezione 3.2.

### 3.4 Un esempio applicativo

Di seguito è riportato un esempio che rappresenta un problema cinematico tratto da una collezione di problemi test provenienti da vari campi



$n$	Broyden	Broyden SM
100	0.0264	0.0177
500	0.1415	0.0723
1000	0.6400	0.2280
1500	1.7285	0.4701
2000	3.2058	0.8068
2500	4.9799	1.2238

Tabella 3.8: Tempo totale di esecuzione (secondi).

dell'ingegneria [F].

La funzione di riferimento è la seguente:

$$F = (f^1, \dots, f^8)^T : \mathbb{R}^8 \rightarrow \mathbb{R}^8,$$

tale che

$$\begin{aligned} f^1(x) &= 4.731 \cdot 10^{-3} x^1 x^3 - 0.3578 x^2 x^3 - 0.1238 x^1 + x^7 - 1.637 \cdot 10^{-3} x^2 \\ &\quad - 0.9338 x^4 - 0.3571; \\ f^2(x) &= 0.2238 x^1 x^3 + 0.7623 x^2 x^3 + 0.2638 x^1 - x^7 - 0.07745 x^2 - 0.6734 x^4 \\ &\quad - 0.6022; \\ f^3(x) &= x^6 x^8 + 0.3578 x^1 + 4.731 \cdot 10^{-3} x^2; \\ f^4(x) &= -0.7623 x^1 + 0.2238 x^2 + 0.3461; \\ f^5(x) &= (x^1)^2 + (x^2)^2 - 1; \\ f^6(x) &= (x^3)^2 + (x^4)^2 - 1; \\ f^7(x) &= (x^5)^2 + (x^6)^2 - 1; \\ f^8(x) &= (x^7)^2 + (x^8)^2 - 1. \end{aligned}$$

Tale sistema risulta avere sedici soluzioni, di cui una è rappresentata dal vettore

$$x^* = (0.6716, 0.7410, 0.9519, -0.3064, 0.9638, -0.2666, 0.4046, 0.9145)^T.$$

Il test è stato effettuato prendendo come dato iniziale una soluzione perturbata, cioè  $x_0 = x^* + 10^{-1}$ ; per il metodo di Broyden è stata considerata l'approssimazione iniziale  $A_0 = J(x_0)$ .

Si nota che il metodo di Newton non converge in quanto trova un  $x_k$  per il quale  $J(x_k)$  è vicino alla singolarità. Al contrario, il metodo di Broyden giunge alla soluzione  $x^*$  dopo solo otto iterazioni, come si vede dalla Tabella 3.9.

$k$	$\ e_k\ $
0	1.24618960e+00
1	6.34675349e-02
2	5.78415142e-03
3	3.80879762e-03
4	4.68499446e-04
5	1.08574475e-04
6	8.82957496e-05
7	8.77760764e-05
8	8.78149059e-05

Tabella 3.9: Risultati relativi al problema in Sezione 3.4.

# Considerazioni finali

In questa tesi sono stati descritti i principali metodi numerici per la risoluzione di sistemi non lineari. Tali metodi sono stati analizzati sia dal punto di vista teorico (analisi di convergenza locale) che pratico (algoritmo e implementazione).

Nella prima parte della tesi, è stato trattato il caso più semplice delle equazioni non lineari, per le quali si è analizzato il metodo di Newton e alcune sue varianti: il metodo di Newton stazionario, il metodo delle secanti e il metodo di Newton alle differenze.

Nella seconda parte, è stato studiato il principale metodo numerico per sistemi non lineari, il metodo di Newton  $n$ -dimensionale, e analizzate le sue proprietà di convergenza locale. Inoltre sono stati riportati alcuni risultati teorici validi in generale per i metodi Quasi-Newton, e tra questi è stato trattato nello specifico il metodo di Broyden. Il metodo di Newton  $n$ -dimensionale rappresenta il metodo più rapido per giungere alla soluzione di un sistema non lineare, ma richiede maggiori conoscenze sul problema. D'altra parte, il metodo di Broyden richiede meno informazioni sul problema ed ha una rapida velocità di convergenza benché inferiore rispetto a quella del metodo di Newton. Infine, sono state proposte implementazioni efficienti dei metodi che sono state applicate nella risoluzione di vari problemi. In particolare è stata messa in evidenza la criticità della scelta dell'approssimazione iniziale per i metodi trattati e il comportamento dell'errore vicino alla soluzione.

In conclusione, i metodi trattati rappresentano strumenti utili per la risoluzione di problemi non lineari sia concreti, derivanti soprattutto dalle discipline fisiche, chimiche e ingegneristiche, che teorici.



# Appendice A

## A.1 Velocità di convergenza

Vi sono principalmente tre tipi di velocità e sono riportate nella seguente definizione:

**Definizione A.1.1.** *Siano  $\{x_k\}_{k \geq 0} \subset \mathbb{R}^n$  e  $x^* \in \mathbb{R}^n$ , tali che la successione  $\{x_k\}_{k \geq 0}$  converga a  $x^*$ .*

- *Se esiste una costante  $c \in [0, 1)$  e un intero  $\tilde{k} \geq 0$  tale che  $\forall k \geq \tilde{k}$  vale*

$$\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|,$$

*allora si dice che la successione  $\{x_k\}_{k \geq 0}$  converge linearmente a  $x^*$ .*

- *Se esiste una successione  $\{c_k\}_{k \geq 0} \rightarrow_{k \rightarrow \infty} 0$  tale che*

$$\|x_{k+1} - x^*\| \leq c_k \|x_k - x^*\|,$$

*allora si dice che la successione  $\{x_k\}_{k \geq 0}$  converge superlinearmente a  $x^*$ .*

- *Se esistono le costanti  $c \geq 0$ ,  $p > 1$  ed un intero  $\tilde{k} \geq 0$  tali che  $\forall k \geq \tilde{k}$  vale*

$$\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|^p,$$

*allora si dice che la successione  $\{x_k\}_{k \geq 0}$  converge a  $x^*$  con ordine almeno  $p$ . In particolare se  $p = 2, 3$  la convergenza è detta rispettivamente quadratica o cubica.*

Il seguente risultato mostra una caratteristica dei metodi numerici in cui la successione  $\{x_k\}_{k \geq 0}$  converge almeno superlinearmente a  $x^*$ , che si trova alla base di un criterio d'arresto e dell'analisi della convergenza del metodo di Broyden.

**Lemma A.1.1.** *Sia  $\{x_k\}_{k \geq 0}$  una successione in  $\mathbb{R}^n$ . Supponendo che  $\{x_k\}_{k \geq 0}$  converga almeno superlinearmente a  $x^* \in \mathbb{R}^n$ , allora per qualsiasi norma  $\|\cdot\|$  vale*

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x_k\|}{\|x_k - x^*\|} = 1.$$

*Dimostrazione.* A partire dalla quantità  $\|x_k - x_{k-1}\|$ , si possono ottenere le seguenti stime:

Utilizzando la disuguaglianza triangolare si ha

$$\|x_k - x_{k-1}\| = \|x_k - x^* + x^* - x_{k-1}\| = \|e_k - e_{k-1}\| \leq \|e_k\| + \|e_{k-1}\|, \quad (\text{A.1})$$

dove  $e_k = x_k - x^*$ . Utilizzando poi la disuguaglianza  $\|x - y\| \geq \left| \|x\| - \|y\| \right|$  e supponendo che per  $k$  sufficientemente grande valga  $\|e_{k-1}\| > \|e_k\|$ , si ottiene

$$\|x_k - x_{k-1}\| \geq \|e_{k-1}\| - \|e_k\|. \quad (\text{A.2})$$

Unendo (A.1) e (A.2), si ottiene la seguente catena di disuguaglianze:

$$1 - \frac{\|e_k\|}{\|e_{k-1}\|} \leq \frac{\|x_k - x_{k-1}\|}{\|e_{k-1}\|} \leq 1 + \frac{\|e_k\|}{\|e_{k-1}\|}.$$

Grazie alla definizione di convergenza superlineare, si ha

$$\lim_{k \rightarrow \infty} \frac{\|x_k - x_{k-1}\|}{\|e_{k-1}\|} = 1.$$

□

Ciò significa che da un certo  $k$  in poi la quantità  $\|x_k - x_{k-1}\|$  approssima  $\|x_{k-1} - x^*\|$ , quantità che in generale non è nota poiché comprende il vettore  $x^*$  di cui a priori non si conosce il valore.

## A.2 Norme

### A.2.1 Norme vettoriali

**Definizione A.2.1.** Una norma su  $\mathbb{R}^n$  è una funzione  $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$  che, per ogni  $v, w \in \mathbb{R}^n$ , soddisfa le seguenti proprietà:

- i)  $\|v\| \geq 0$ , e  $\|v\| = 0$  se e solo se  $v$  rappresenta il vettore nullo;
- ii)  $\|\alpha v\| = |\alpha| \cdot \|v\|$  per ogni  $\alpha \in \mathbb{R}$ ;
- iii)  $\|v + w\| \leq \|v\| + \|w\|$ .

In generale le norme più utilizzate sono:  
per  $v = (v^1, \dots, v^n)^T \in \mathbb{R}^n$ ,

$$\begin{aligned} \text{norma supremum o } l_\infty \quad \|v\|_\infty &:= \max_{1 \leq i \leq n} |v^i|; \\ \text{norma } l_1 \quad \|v\|_1 &:= \sum_{i=1}^n |v^i|; \\ \text{norma } l_2 \text{ o euclidea} \quad \|v\|_2 &:= \left( \sum_{i=1}^n (v^i)^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Le norme  $l_1$  e  $l_2$  sono esempi di norme  $l_p$  che vengono definite nel seguente modo:

per  $1 \leq p < \infty$ ,

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

### A.2.2 Norme matriciali

Analogamente ai vettori, anche le matrici possono essere quantificate attraverso le norme.

**Definizione A.2.2.** [HJ] Una norma su  $\mathbb{R}^{n \times n}$  è una funzione  $\|\cdot\| : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  che, per ogni  $A, B \in \mathbb{R}^{n \times n}$ , soddisfa le seguenti proprietà:

- i)  $\|A\| \geq 0$ , e  $\|A\| = 0$  se e solo se  $A$  rappresenta la matrice nulla;
- ii)  $\|\alpha A\| = |\alpha| \cdot \|A\|$  per ogni  $\alpha \in \mathbb{R}$ ;
- iii)  $\|A + B\| \leq \|A\| + \|B\|$ ;
- iv)  $\|AB\| \leq \|A\| \cdot \|B\|$ .

Se la matrice è vista come operatore essa può essere misurata attraverso la norma indotta dalla norma vettoriale  $\|\cdot\|$ . Una definizione naturale di norma indotta sulla matrice  $A$  è

$$\|A\| = \max_{v \in \mathbb{R}^n, v \neq 0} \left\{ \frac{\|Av\|}{\|v\|} \right\}. \quad (\text{A.3})$$

Per ogni norma vettoriale  $l_p$  la norma in (A.3) viene indicata con  $\|A\|_p$ . Perciò le norme matriciali indotte dalle norme  $l_1, l_2, l_\infty$  risultano essere le seguenti:

$$\|A\|_1 = \max_{1 \leq j \leq n} \{\|a_{.j}\|_1\};$$

$$\|A\|_2 = \max_{v \in \mathbb{R}^n, v \neq 0} \left\{ \frac{\|Av\|_2}{\|v\|_2} \right\};$$

$$\|A\|_\infty = \max_{1 \leq i \leq n} \{\|a_{i.}\|_1\}.$$

Oltre alle norme indotte, si può definire un'ulteriore norma matriciale detta norma di Frobenius, che risulta essere la norma  $l_2$  di  $A$  scritta come vettore, infatti è definita nel seguente modo:

$$\|A\|_F = \left( \sum_{j=1}^n \sum_{i=1}^n a_{ij}^2 \right)^{\frac{1}{2}}.$$

Altre proprietà delle norme sono espresse dal seguente teorema:

**Teorema A.2.1.** *Siano  $\|\cdot\|$  e  $\|\|\cdot\|\|$  delle norme matriciali su  $\mathbb{R}^{n \times n}$ . Allora esistono delle costanti  $\alpha, \beta$  tali che*

$$\alpha\|A\| \leq \|\|A\|\| \leq \beta\|A\|,$$

per ogni  $A \in \mathbb{R}^{n \times n}$ . In particolare,

$$n^{-\frac{1}{2}}\|A\|_F \leq \|A\|_2 \leq \|A\|_F,$$

e, per  $p = 1$  o  $p = \infty$ ,

$$n^{-\frac{1}{2}}\|A\|_p \leq \|A\|_2 \leq n^{\frac{1}{2}}\|A\|_p.$$

Per ogni matrice  $A, B \in \mathbb{R}^{n \times n}$ , vale

$$\|AB\|_F \leq \min\{\|A\|_2\|B\|_F, \|A\|_F\|B\|_2\}.$$

Inoltre, per ogni  $v, w \in \mathbb{R}^n$ , vale

$$\|vw^T\|_F = \|vw^T\|_2 = \|v\|_2 \cdot \|w\|_2.$$

Di seguito sono espressi alcuni risultati che sono necessari per la dimostrazione di alcuni teoremi enunciati nei capitoli precedenti.

**Lemma A.2.1.** [DS, Teorema 3.1.4] *Siano  $A, B, E \in \mathbb{R}^{n \times n}$  e  $\|\cdot\|$  una norma su  $\mathbb{R}^{n \times n}$  tale che  $\|I\| = 1$ . Se  $\|E\| < 1$ , allora  $(I - E)^{-1}$  esiste e vale*

$$\|(I - E)^{-1}\| \leq \frac{1}{1 - \|E\|}.$$

Se  $A$  è una matrice non singolare e  $\|A^{-1}(B - A)\| < 1$ , allora  $B$  risulta essere una matrice non singolare e vale la seguente stima:

$$\|B^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}(B - A)\|}.$$

**Lemma A.2.2.** [DS, Lemma 8.2.5] *Siano  $s \in \mathbb{R}^n$ ,  $s \neq 0$ ,  $E \in \mathbb{R}^{n \times n}$  e  $\|\cdot\|$  la norma vettoriale  $l_2$ . Allora*

$$\left\| E \left( I - \frac{ss^T}{s^T s} \right) \right\|_F = \left( \|E\|_F^2 - \left( \frac{\|Es\|}{\|s\|} \right)^2 \right)^{\frac{1}{2}} \leq \|E\|_F - \frac{1}{2\|E\|_F} \left( \frac{\|Es\|}{\|s\|} \right)^2.$$

**Definizione A.2.3.** *Sia  $W \subset \mathbb{R}^n$  un sottospazio di dimensione  $k$  con  $k \leq n$ .  $P \in \mathbb{R}^{n \times n}$  si dice proiettore ortogonale per il sottospazio  $W$  se  $P$  è una matrice simmetrica,  $\text{rango}(P) = k$  e  $P^2 = P$ .*

**Lemma A.2.3.** *Sia  $P \in \mathbb{R}^{n \times n}$  un proiettore ortogonale allora  $\|P\|_2 = 1$ .*



### A.3 Nozioni di analisi per funzioni in più variabili

**Definizione A.3.1.** Sia  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $F = (f^1, \dots, f^n)$  una funzione continua con  $f^i : \mathbb{R}^n \rightarrow \mathbb{R}$ . Si dice che  $F$  è una funzione differenziabile in modo continuo in  $x \in \mathbb{R}^n$  se ogni sua componente  $f^i$ ,  $i = 1, \dots, n$  è differenziabile in modo continuo in  $x$ .

La derivata prima di  $F$  in  $x$  è rappresentata da una matrice di dimensione  $n \times n$  detta Jacobiana e definita nel seguente modo:

$$(J(x))_{ij} = \left( \frac{\partial f^i}{\partial x^j}(x) \right),$$

per ogni  $i, j = 1, \dots, n$ .

Si dice che  $F$  è differenziabile in modo continuo in un insieme aperto  $\mathcal{D} \subset \mathbb{R}^n$  se  $F$  è differenziabile in modo continuo in ogni punto  $x \in \mathcal{D}$ , in tal caso si scrive  $F \in C^1(\mathcal{D})$ .

**Lemma A.3.1.** [DS, Lemma 4.1.9] Siano  $\mathcal{D} \subset \mathbb{R}^n$  un insieme aperto convesso e  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  una funzione tale che  $F \in C^1(\mathcal{D})$ . Per ogni  $x, x+p \in \mathcal{D}$  vale

$$F(x+p) - F(x) = \int_0^1 [J(x+tp)p] dt = \int_x^{x+p} J(z) dz.$$

**Lemma A.3.2.** [DS, Lemma 4.1.10] Sia  $G : \mathcal{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ , dove  $\mathcal{D}$  è un insieme aperto convesso tale che  $x, x+p \in \mathcal{D}$ . Per qualsiasi norma considerata sullo spazio  $\mathbb{R}^{n \times n}$ , se  $G$  è integrabile su  $[x, x+p]$  vale la formula

$$\left\| \int_0^1 [G(x+tp)p] dt \right\| \leq \int_0^1 \|G(x+tp)p\| dt.$$

**Definizione A.3.2.** Siano  $\mathcal{D} \subset \mathbb{R}^n$  aperto,  $\|\cdot\|$  una norma su  $\mathbb{R}^n$  e  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $F$  si dice lipschitziana su  $\mathcal{D}$  se esiste una costante  $L$ , detta costante di Lipschitz, tale che valga

$$\|F(x) - F(y)\| \leq L\|x - y\|, \quad \forall x, y \in \mathcal{D}.$$

In tal caso si scrive  $F \in Lip_L(\mathcal{D})$ .

**Definizione A.3.3.** Siano  $G : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ ,  $x \in \mathbb{R}^n$ ,  $\|\cdot\|$  una norma su  $\mathbb{R}^n$  e  $\|\|\cdot\|\|$  una norma su  $\mathbb{R}^{n \times n}$ .  $G$  si dice lipschitziana in  $x$  se esiste un aperto  $\mathcal{D} \subset \mathbb{R}^n$ ,  $x \in \mathcal{D}$ , e una costante  $L$  tale che  $\forall v \in \mathcal{D}$  vale

$$\|\|G(v) - G(x)\|\| \leq L\|v - x\|. \quad (\text{A.4})$$

La costante  $L$  è detta costante di Lipschitz per  $G$  in  $x$ . Se la formula (A.4) vale per ogni  $x \in \mathcal{D}$  allora  $G \in Lip_L(\mathcal{D})$ .

**Lemma A.3.3.** [DS, Lemma 4.1.15] *Siano  $\mathcal{D} \subset \mathbb{R}^n$  un insieme aperto convesso,  $x \in \mathcal{D}$  e  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  una funzione differenziabile in modo continuo tale che  $J$  sia lipschitziana in  $x \in \mathcal{D}$  con costante di Lipschitz  $L$ . Si consideri  $\|\cdot\|$  una norma vettoriale e la norma matriciale indotta, allora  $\forall u, v \in \mathcal{D}$  vale*

$$\|F(v) - F(u) - J(x)(v - u)\| \leq L \frac{\|v - x\| + \|u - x\|}{2} \|v - u\|.$$

**Lemma A.3.4.** [DS, Lemma 4.1.16] *Siano  $\mathcal{D} \subset \mathbb{R}^n$  un insieme aperto convesso,  $x \in \mathcal{D}$  e  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  una funzione differenziabile in modo continuo tale che esista  $J(x)^{-1}$  e  $J$  sia lipschitziana in  $x \in \mathcal{D}$  con costante di Lipschitz  $L$ . Si consideri  $\|\cdot\|$  una norma vettoriale e la norma matriciale indotta, allora esistono  $\epsilon > 0$ ,  $0 < \alpha < \beta$  tali che*

$$\alpha\|v - u\| \leq \|F(v) - F(u)\| \leq \beta\|v - u\|,$$

*per ogni  $u, v \in \mathcal{D}$  tali che  $\max\{\|v - x\|, \|u - x\|\} \leq \epsilon$ .*

# Bibliografia

- [DS] J. E. Dennis, R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equation*, Prentice Hall, Englewood Cliffs, NJ, 1983;
- [K] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, Society for Industrial and Applied Mathematics, SIAM, 1995;
- [NW] J. Nocedal, S. J. Wright, *Numerical Optimization*, Springer Series in Operations Research, 1999;
- [HJ] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 2013;
- [MGH] J.J. Moré, B. S. Garbow, K. E. Hillstom, *Testing Unconstrained Optimization Software*, ACM Transactions on Mathematical Software (TOMS), 1981;
- [F] C. A. Floudas et altri, *Handbook of Test Problems in Local and Global Optimization*, Kluwer Academic Publishers, Nonconvex Optimization and its Applications, 33, 1999.