

# Risoluzione di sistemi di equazioni non lineari

Gerardo Toraldo

Università di Napoli Federico II

A.A. 2015-2016

## Metodi di Newton inesatti

Il metodo di Newton per il sistema di equazioni non lineare  $F(x) = 0$

$$x_{k+1} = x_k + s_k \text{ dove}$$

$$F'(x_k)s_k = -F(x_k) \Leftrightarrow r_k = F'(x_k)s_k + F(x_k) = 0$$

**Metodi Newton-inesatti:** (2) é risolto in maniera approssimata con un **residuo**

$$r_k = F'(x_k)s_k + F(x_k)$$

(possibilmente) **non nullo** (e passo  $s_k = F'(x_k)^{-1}r_k - F'(x_k)^{-1}F(x_k)$ ). Che accuratezza richiedere su  $r_k$ ? Si considera, come "misura di accuratezza" il rapporto

$$\eta_k = \frac{\|r_k\|}{\|F(x_k)\|} \simeq \frac{\|F(x_{k+1})\|}{\|F(x_k)\|}$$

### Metodo di Newton

$x_0 \in \mathbb{R}^n; k \leftarrow 0;$

**while** *not\_stopcondition*

    solve  $F'(x_k)s = F(x_k);$

$x_{k+1} \leftarrow x_k + s; k \leftarrow k + 1;$

**endwhile**

### Metodo di Newton Inesatto

$x_0 \in \mathbb{R}^n; k \leftarrow 0;$

**while** *not\_stopcondition*

**approx**solve  $F'(x_k)s = F(x_k);$

$x_{k+1} \leftarrow x_k + s; k \leftarrow k + 1;$

**endwhile**

## Teorema: Convergenza metodi NI

Supponiamo che  $\{\eta_k\}$  sia tale che  $\eta_k < t < 1$ , allora esiste  $\varepsilon > 0$  tale che, se  $\|x_0 - x^*\| < \varepsilon$  la successione  $\{x_k\}$  converge a  $x^*$  **q-linearmente** e, in particolare,

$$\|x_{k+1} - x^*\|_* < t\|x_k - x^*\|_* \quad \text{dove } \|y\|_* = \|F'(x^*)y\| \quad (1)$$

Osserviamo che, poiché  $F'(x_*)$  è non singolare, per ogni  $\gamma > 0$  esiste  $\rho > 0$  tale che, se  $\|x - x^*\| < \rho$  allora

$$\begin{aligned} \|F'(x) - F'(x^*)\| &< \gamma \\ \|F'(x)^{-1} - F'(x^*)^{-1}\| &< \gamma \\ \|F(x) - F'(x^*)(x - x_*)\| &< \gamma\|x - x_*\| \end{aligned}$$

e, di conseguenza

$$\|F(x_k)\| = \|[F(x_k) - F'(x^*)e_k] + F'(x^*)e_k\| \leq 2\gamma\|e_k\|$$

Inoltre,

$$\frac{1}{\mu}\|y\| \leq \|y\|_* \leq \mu\|y\| \quad \text{dove } \mu = \max\{\|F'(x^*)\|, \|F'(x^*)^{-1}\|\}$$

$$\begin{aligned}
 F'(x^*)e_{k+1} &= F'(x^*)(e_k + s_k) = F'(x^*) [e_k + F'(x_k)^{-1}r_k - F'(x_k)^{-1}F(x_k)] \\
 &= F'(x^*)F'(x_k)^{-1} [F'(x_k)e_k + r_k - F(x_k)] = \\
 &F'(x^*)F'(x_k)^{-1} \left[ r_k + \underbrace{(F'(x_k) - F'(x^*))e_k}_{\text{error term}} + \underbrace{(-F(x_k) + F'(x^*)e_k)}_{\text{error term}} \right];
 \end{aligned}$$

$$\begin{aligned}
 \|F'(x^*)F'(x_k)^{-1}\| &= \|F'(x^*)(F'(x_k)^{-1} - F'(x^*)^{-1}) + I\| \\
 &\leq 1 + \|F'(x^*)\| \cdot \|F'(x_k)^{-1} - F'(x^*)^{-1}\| \leq 1 + \gamma\mu
 \end{aligned}$$

$$\begin{aligned}
 \|e_{k+1}\|_* &\leq (1 + \gamma\mu)(\|r_k\| + \|F'(x_k) - F'(x^*)\|\|e_k\| + \|F(x_k) - F'(x^*)e_k\|) \quad (2) \\
 &\leq (1 + \gamma\mu)(\eta_k\|F(x_k)\| + \gamma\|e_k\| + \gamma\|e_k\|) \leq (1 + \gamma\mu)(2\gamma\eta_k + 2\gamma)\|e_k\| \\
 &\leq (1 + \gamma\mu)(2\gamma\eta_k + 2\gamma)\mu\|e_k\|_*
 \end{aligned}$$

Lemma: (maggiorazione di  $\|F\|$ ).

Se

$$\alpha = \max\{\|F'(x_*)\| + \frac{1}{2\|F(x_*)^{-1}\|}, 2\|F(x_*)^{-1}\|\}$$

allora

$$\frac{1}{\alpha}\|x - x^*\| \leq \|F(x)\| \leq \alpha\|x - x^*\|$$

## Teorema: Velocità di Convergenza metodi NI

Supponiamo che la successione  $\{x_k\}$  generata da un metodo NI converga a  $x^*$ . La convergenza è superlineare se e solo se

$$\|r_k\| = o(\|F(x_k)\|) \text{ per } k \rightarrow \infty$$

**Dimostrazione:** supponiamo  $\|e_{k+1}\| = o(\|e_k\|)$

$$\begin{aligned} r_k &= \underbrace{F'(x_k)s_k + F(x_k)}_{= F(x_k)} = \underbrace{F'(x_k)e_{k+1} - F'(x_k)e_k + F(x_k) - F'(x^*)e_k + F'(x^*)e_k}_{= F(x_k) - F'(x^*)e_k} = \\ &= [F(x_k) - F'(x^*)e_k] + [(F'(x_k) - F'(x^*))e_k] + [(F'(x^*) + F'(x_k) - F'(x^*))e_{k+1}] \end{aligned}$$

e quindi

$$\|r_k\| \leq o(\|e_k\|) + o(1)\|e_k\| + [\|F'(x^*)\| + o(1)] o(\|e_k\|)$$

e quindi  $\|r_k\| = o(\|e_k\|) = o(\|F(x_k)\|) \quad k \rightarrow \infty$

Se  $\|r_k\| = o(\|F(x_k)\|)$ , ricordando ( 2) avremo

$$\begin{aligned} \|e_{k+1}\|_* &\leq (1 + \gamma\mu)(\|r_k\| + \|F'(x_k) - F'(x^*)\|\|e_k\| + \|F(x_k) - F'(x^*)e_k\|) \\ &= o(\|F(x_k)\|) + o(1)(\|e_k\|) + o(\|e_k\|) = o(\|e_k\|) \end{aligned}$$

## Teorema: Velocità di Convergenza metodi NI

Supponiamo che la successione  $\{x_k\}$  generata da un metodo NI converga a  $x^*$ . Se  $\eta_k \rightarrow 0$ , allora la convergenza è **q-superlineare**.

Il generico metodo di Newton inesatto per il sistema di equazioni non lineare  $F(x) = 0$  può essere scritto nella forma

$$x_{k+1} = x_k + s_k \quad \text{dove} \quad (3)$$

$$H_k s_k = -F(x_k) \quad (4)$$

In questo caso

$$r_k = F'(x_k) (-H_k F(x_k)) + F(x_k) \quad \text{e quindi}$$

$$\|r_k\| = \|F'(x_k) (-H_k F(x_k)) + F(x_k)\| \leq \|I - F'(x_k) H_k\| \cdot \|F(x_k)\|$$

In particolare, la successione  $\eta_k$  è maggiorata dalla successione  $\rho_k$  dove

$$\rho_k = \|I - F'(x_k) H_k\|$$

### Teorema: Velocità di Convergenza metodi NI

Supponiamo che la successione  $\{x_k\}$  generata da un metodo NI converga a  $x^*$ . Se  $F'(x)$ , è Lipschitz continua in  $x^*$  allora la convergenza è **q-quadratica**, cioè esiste  $C > 0$  tale che

$$\|x_{k+1} - x^*\| \leq C\|x_k - x^*\|^2.$$

**Metodo di Newton Inesatto globalizzato**

$x_0 \in \mathfrak{R}^n$ ;  $k \leftarrow 0$ ;  $t \in (0, 1)$

**while**  $\|F(x_k)\| > tol$

trova  $\eta_k \in [0, 1)$  e  $s_k$  tali che

$$\|F(x_k) + F'(x_k)s_k\| \leq \eta_k \|F(x_k)\| \quad (\text{C1}) \ \&\&$$

$$\|F(x_k + s_k)\| \leq [1 - t(1 - \eta_k)] \|F(x_k)\| \quad (\text{C2})$$

$k \leftarrow k + 1$ ;  $x_{k+1} = x_k + s_k$

**endwhile**

Notiamo che, posto

$$\text{pred}_k = \|F(x_k)\| - \|F(x_k) + F'(x_k)s_k\|, \quad \text{ared}_k = \|F(x_k)\| - \|F(x_k + s_k)\|$$

$$(\text{C1}) \Rightarrow \text{pred}_k \geq (1 - \eta_k) \|F(x_k)\| \quad (5)$$

$$(\text{C2}) \Rightarrow \text{ared}_k \geq t(1 - \eta_k) \|F(x_k)\| \quad (6)$$

e quindi la condizione su  $s_k$  implica

$$\text{ared}_k \geq t(\text{pred}_k)$$

che può essere interpretata come una condizione di fedeltà del modello lineare di  $F$  ripetuto ad  $F$ .

**Metodo di Newton Inesatto con *Backtracking***

$x_0 \in \mathbb{R}^n$ ;  $k \leftarrow 0$ ;  $\eta_{max} \in [0, 1)$ ,  $0 < \theta_{min} < \theta_{max} < 1$

**while**  $\|F(x_k)\| > tol$

$\eta_k \in [0, \eta_{max}]$  calcola  $s_k$  tale che

$$\|F(x_k) + F'(x_k)s_k\| \leq \eta_k \|F(x_k)\|$$

**while**  $\|F(x_k + s_k)\| > [1 - t(1 - \eta_k)]\|F(x_k)\|$

scegli  $\Theta \in [\Theta_{min}, \Theta_{max}]$

$$s_k = \Theta s_k, \quad \eta_k = 1 - \Theta(1 - \eta_k)$$

**endwhile**

$$k \leftarrow k + 1; \quad x_{k+1} = x_k + s_k$$

**endwhile**

**Teorema: Convergenza NIB**

Supponiamo che la successione  $\{x_k\}$  generata da un metodo NIB abbia un punto di accumulazione  $x^*$ , con  $F'(x_*)$  invertibile, allora  $F(x_*) = 0$  e  $x_k \rightarrow x_*$ . Inoltre, per  $k$  sufficientemente grande, i valori iniziali di  $s_k$  ed  $\eta_k$  saranno accettabili.

Se la successione è limitata allora (in alternativa):

- Converge ad una soluzione  $x_*$
- Ammette punti di accumulazione con Jacobiano singolare.

## Implementazione di NIB: scelta dei parametri

$$\eta_{max} = 0.9; \quad \theta_{min} = 0.1; \quad \theta_{max} = 0.5; \quad t = 1e - 4$$

Termini forzanti e convergenza:

- $\eta_k = \eta < 1 \Rightarrow$  convergenza **lineare**
- $\eta_k \rightarrow 0 \Rightarrow$  convergenza **superlineare**
- $\eta_k = O(\|F(x_k)\|) \Rightarrow$  convergenza **quadratica**

Una possibile scelta per i termini forzanti:

$$\eta_k = \min\{\eta_{max}, \tilde{\eta}_k\} \text{ dove}$$

$$\tilde{\eta}_k = \frac{\| \|F(x_k)\| - \|F(x_{k-1}) + F'(x_{k-1})s_{k-1}\| \|}{\|F(x_{k-1})\|}$$

## Metodo delle secanti

Il metodo di Newton per il sistema di equazioni non lineare  $F(x) = 0$

$$x_{k+1} = x_k - (F'(x_k))^{-1}F(x_k) \quad (7)$$

è una immediata generalizzazione del metodo delle tangenti (Newton-Raphson) a funzioni di piú variabili (formalmente ottenuta sostituendo la derivata della funzione con lo Jacobiano del sistema). Nel caso di funzioni di una variabile, le varianti del metodo di Newton che non utilizzano le derivate (secanti, falsa posizione, ...)

$$x_{k+1} = x_k - \frac{b-a}{F(b)-F(a)}F(x_k) \quad (8)$$

sono formalmente ottenute da (7) sostituendo la derivata con una approssimazione alle differenze finite  $\frac{F(b)-F(a)}{b-a}$  (con  $a, b$  opportunamente scelti).

**Un tale procedimento non è generalizzabile a funzioni di piú variabili.** Occorre stabilire delle opportune condizioni da imporre alla approssimazione  $B_k$  della matrice Jacobiana  $F'(x_k)$

### Condizione secante

$$B_k s_k = y_k, \quad \text{dove } s_k = x_k - x_{k-1}, \quad y_k = F(x_k) - F(x_{k-1}) \quad (9)$$

La condizione (9) ammette infinite soluzioni (che costituiscono uno spazio affine di dimensione  $n^2 - n$ )

## Condizione secante: Broyden (good)

$$B_{k+1} = \operatorname{argmin}_{X \in \mathbb{R}^{n \times n}} \|X - B\|_F \quad \text{s.t. } Xs_k = y_k \quad \text{con } B = B_k \quad (10)$$

Il problema (10) nelle  $n^2$  variabili  $x_{ij}$  può essere formulato come <sup>1</sup>

$$\min \Phi(X) = XX - 2BX \quad \text{s.t. } Xs = y, \quad (11)$$

o, equivalentemente,

$$\min \Phi(X) = 0.5(XX - 2BX) \quad \text{s.t. } \Theta_i(X) = X^i s = y_i, \quad i = 1..n, \quad (12)$$

dove  $X^i$  è la  $i$ -ma riga di  $X$  e gli indici di  $s_k, y_k$  sono omissi per semplificare le notazioni. Le condizioni di Lagrange per (12) sono

$$\nabla \Phi(X) = X - B = \sum_i^n \lambda_i \nabla \Theta_i(X) \Leftrightarrow X^i - B^i = \lambda_i s \Leftrightarrow X - B = \lambda s^T \quad (13)$$

Per ricavare il vettore dei parametri di Lagrange in (13), moltiplicando ambo i membri per  $s$  e ricordando il vincolo  $Xs = y$  si ottiene

$$Xs - Bs = \lambda s^T s \Leftrightarrow y - Bs = \lambda s^T s \quad \text{da cui} \quad (14)$$

$$\lambda = \frac{y - Bs}{s^T s} \quad \text{e quindi } X = B + \lambda s^T = B + \frac{y s^T - (Bs) s^T}{s^T s} \quad (15)$$

<sup>1</sup>il prodotto matrice matrice va inteso come prodotto componente per componente:  
 $C = AB, \quad c_{ij} = a_{ij} b_{ij}$

## Condizione secante: Broyden (good)

$$B_k = B_{k-1} + \frac{y_k s_k^T - (B_{k-1} s_k) s_k^T}{s_k^T s_k} = B_{k-1} + p_k s_k^T \quad \text{dove} \quad p_k = \frac{y_k - B_{k-1} s_k}{s_k^T s_k} \quad (16)$$

- Con la (16) all'iterazione  $k$  viene calcolata una approssimazione della matrice Jacobiana  $F'(x_k)$  a partire da quella calcolata all'iterazione  $k - 1$  mediante un *update* di rango 1.
- La risoluzione di un sistema lineare è l'operazione piú onerosa per il calcolo del passo ad ogni iterazione;
- In generale la successione  $\{B_k\}$  non converge a  $F'(x^*)$ ;
- La formula (16) non preserva simmetria e pd.
- Convergenza superlineare:  $\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0$

**Metodo di Broyden (diretto) Good**

$$x_0 \in \mathfrak{R}^n; B_0 \in \mathfrak{R}^{n \times n}; k \leftarrow 1;$$

$$F_0 = F(x_0); s_1 \leftarrow -B_0^{-1} F_0;$$

$$x_1 \leftarrow x_0 + s_1, F_1 \leftarrow F(x_1);$$

$$y_1 = F_1 - F_0$$
**while**  $\|F(x_k)\| > tol$ 

$$B_k \leftarrow B_{k-1} + \frac{y_k s_k^T - (B_{k-1} s_k) s_k^T}{s_k^T s_k}$$

$$s_{k+1} \leftarrow -B_k^{-1} F_k;$$

$$x_{k+1} \leftarrow x_k + s_{k+1}; F_{k+1} = F(x_{k+1})$$

$$k \leftarrow k + 1; y_k = F_k - F_{k-1}$$
**endwhile**

## Condizione secante: Broyden (bad)

Idea: fare una update di una **matrice che approssimi l'inversa della matrice Jacobiana**.

## Formula di Sherman Morrison Woodbury

$$(A + vw^T)^{-1} = A^{-1} - \frac{A^{-1}vw^T A^{-1}}{1 + w^T A^{-1}v}$$

$$\begin{aligned} \left( A^{-1} - \frac{A^{-1}vw^T A^{-1}}{1 + w^T A^{-1}v} \right) (A + vw^T) &= I + A^{-1}vw^T - \frac{A^{-1}vw^T + A^{-1}vw^T A^{-1}vw^T}{1 + w^T A^{-1}v} = \\ I + \frac{A^{-1}vw^T + A^{-1}vw^T \overbrace{w^T A^{-1}v} - A^{-1}vw^T - A^{-1}v \overbrace{w^T A^{-1}v} w^T}{1 + w^T A^{-1}v} &= I \end{aligned}$$

Utilizzando SMW con  $B_{k-1} \rightarrow A$ ,  $p_k \rightarrow v$ ,  $s_k \rightarrow w$ ,  $H = (B)^{-1}$  si ha

$$H_k = (B_k)^{-1} = (B_{k-1} + p_k s_k^T)^{-1} = H_{k-1} - \frac{H_{k-1} p_k s_k^T H_{k-1}}{1 + s_k^T H_{k-1} p_k} = \quad (17)$$

$$H_{k-1} - \frac{(H_{k-1} y_k - s_k) s_k^T H_{k-1} / s_k^T s_k}{1 + \frac{s_k^T H_{k-1} y_k - s_k^T s_k}{s_k^T s_k}} = H_{k-1} + \frac{(s_k - H_{k-1} y_k) s_k^T H_{k-1}}{s_k^T H_{k-1} y_k} \quad (18)$$

## Condizione secante: Broyden (bad)

$$H_k = H_{k-1} + \frac{(s_k - H_{k-1}y_k)s_k^T H_{k-1}}{s_k^T H_{k-1}y_k} \quad (19)$$

- Con la (19) all'iterazione  $k$  viene calcolata una approssimazione dell'inversa della matrice Jacobiana a partire da quella calcolata all'iterazione  $k - 1$  mediante un *update* di rango 1.
- Un prodotto matrice vettore è il calcolo piú oneroso richiesto per il calcolo del passo ad ogni iterazione;
- In generale la successione  $\{H_k\}$  non converge a  $F'(x^*)^{-1}$ ;
- La formula (19) non preserva simmetria e pd.
- Convergenza superlineare:  $\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0$

## Matrici di Broyden: invertibilità

### Matrix Determinant Lemma (MDL)

Siano  $A \in \mathbb{R}^{n \times n}$ ,  $u, v \in \mathbb{R}^n$ , con  $A$  non singolare. Allora

$$\det(A + uv^T) = (1 + v^T A^{-1} u) \det(A)$$

### Matrici di Broyden: invertibilità

$$B_{k-1} \text{ invertibile} \Rightarrow B_k \text{ invertibile sse } s_k^T B_{k-1}^{-1} y_k \neq 0$$

$$H_{k-1} \text{ invertibile} \Rightarrow H_k \text{ invertibile sse } s_k^T H_{k-1} y_k \neq 0$$

#### Dimostrazione:

$$\det(B_k) = \det\left(B_{k-1} + \frac{y_k s_k^T - (B_{k-1} s_k) s_k^T}{s_k^T s_k}\right) = \det\left(B_{k-1} + p_k s_k^T\right) \left[p_k = \frac{y_k - B_{k-1} s_k}{s_k^T s_k}\right]$$

$$\text{Applicando MDL si ha } \det(B_k) = \det(B_{k-1}) \left(1 + s_k^T B_{k-1}^{-1} \frac{y_k - B_{k-1} s_k}{s_k^T s_k}\right)$$

$$= \det(B_{k-1}) \left(1 + \frac{s_k^T B_{k-1}^{-1} y_k}{s_k^T s_k} - 1\right) = \det(B_{k-1}) \frac{s_k^T B_{k-1}^{-1} y_k}{s_k^T s_k}$$

Broyden per  $F(x) = Ax - b$

$$B_k = B_{k-1} + \frac{y_k s_k^T - (B_{k-1} s_k) s_k^T}{s_k^T s_k} = B_{k-1} + p_k s_k^T \quad \text{dove} \quad p_k = \frac{y_k - B_{k-1} s_k}{s_k^T s_k}$$

Se  $F(x) = Ax - b$ , allora  $y_k = A s_k$  e, posto  $E_{k-1} = B_{k-1} - A$ , si ha

$$p_k = \frac{(A - B_{k-1}) s_k}{s_k^T s_k} = -\frac{E_{k-1} s_k}{s_k^T s_k}, \quad \text{e quindi} \quad B_k = B_{k-1} - E_{k-1} \frac{s_k s_k^T}{s_k^T s_k}$$

$$E_k = E_{k-1} \left( I - \frac{s_k s_k^T}{s_k^T s_k} \right) \Rightarrow \|E_k\|_2 \leq \|E_{k-1}\|_2 \quad \text{essendo} \quad \left\| I - \frac{s_k s_k^T}{s_k^T s_k} \right\|_2 = 1$$

Al crescere delle iterazioni, le matrici di Broyden "si avvicinano" ad  $A$ .

## Convergenza dei Metodi di Broyden

### Broyden: convergenza finita per $Ax = b$ (Gay 1979)

Nel caso in cui sia  $F(x) = Ax - b$ , il metodo di Broyden *Good* [Bad] **converge in al piú  $2n$  iterazioni**, purché la matrice  $B_0 [H_0]$  sia non singolare e durante le iterazioni sia soddisfatta la condizione  $s_k^T B_{k-1}^{-1} y_k \neq 0$  [ $s_k^T H_{k-1} y_k \neq 0$ ]

### Broyden: convergenza per $F \in C^1$

Nel caso in cui  $F(x)$  sia di classe  $C^1$ ,  $x_*$  sia tale che  $F(x_*) = 0$ ,  $\det(F'(x_*)) \neq 0$  con  $F'(x)$  Lipschitz continua in un intorno di  $x_*$ , e inoltre si verifichi  $s_k^T B_{k-1}^{-1} y_k \neq 0$  [ $s_k^T H_{k-1} y_k \neq 0$ ]

il metodo di Broyden *Good* [Bad] è localmente convergente intorno a  $x_*$ , con velocità di convergenza superlineare.