Sistemi di equazioni non lineari
Zeri di polinomi

Calcolo Numerico - Ing. Inf. - Lezione 11

Outline

Outline

La teoria dei metodi iterativi precedentemente esposta può essere estesa al caso in cui sia $f(x) : \mathbb{R}^n \to \mathbb{R}^n$, cioè al caso di un sistema di n equazioni (non lineari) in altrettante incognite

$$
\begin{gathered}
f_1(x_1, \ldots, x_n) = 0 \\
f_2(x_1, \ldots, x_n) = 0 \\
\vdots \\
f_n(x_1, \ldots, x_n) = 0
\end{gathered}
$$

Il sistema si può scrivere in una forma equivalente la quale consente di approssimare una soluzione $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T$ come punto fisso di una opportuna funzione di iterazione $\phi(x) : \mathbb{R}^n \to \mathbb{R}^n$ data da

$$\phi(x) = x - G(x)\, f(x)$$

dove $\phi = (\phi_1, \phi_2, \ldots, \phi_n)^T$, $x = (x_1, x_2, \ldots, x_n)^T$, $f = (f_1, f_2, \ldots, f_n)^T$ e $G(x)$ è una matrice $n \times n$ non singolare in un dominio $D \subseteq \mathbb{R}^n$ a cui appartiene α

Si considerano quindi metodi iterativi della forma

$$x^{(k+1)} = \phi(x^{(k)}), \qquad k = 0, 1, \dots,$$

per i quali si definisce l'ordine di convergenza come nel caso scalare cambiando nella definizione i valori assoluti in norme di vettori

Se esistono continue le derivate parziali prime delle funzioni ϕ_i, introducendo la matrice jacobiana

$$\Phi(x) = \begin{pmatrix} \frac{\partial \phi_1}{\partial x_1} & \dots & \frac{\partial \phi_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \phi_n}{\partial x_1} & \dots & \frac{\partial \phi_n}{\partial x_n} \end{pmatrix},$$

si può generalizzare il Teorema di convergenza locale

Teorema di convergenza locale in $\mathbb{R}^n$

Se α è un punto fisso di $\phi(x)$, condizione sufficiente per la convergenza ad α del metodo iterativo $x^{(k+1)} = \phi(x^{(k)})$ è che esistano due numeri positivi K e ρ, con $K < 1$, tali che si abbia

$$\|\Phi(x)\| \leq K, \quad \forall x \in D_\rho = \{x| \, \|x - \alpha\| \leq \rho\}$$

purché $x^{(0)}$ sia scelto in D_ρ; in tal caso α è l'unico punto fisso di ϕ in D_ρ

Per estendere il metodo di Newton ad un sistema non lineare, supponiamo che le funzioni f_i siano derivabili con continuità rispetto a ciascuna variabile e che la matrice jacobiana

$$J(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

sia non singolare in un dominio contenente nel suo interno una soluzione α del sistema

Specializzando la funzione di iterazione nella forma

$$\phi(x) \;=\; x \;-\; J^{-1}(x)\, f(x),$$

si ha il sistema $x = x - J^{-1}(x)f(x)$ equivalente al sistema $f(x) = 0$, da cui il metodo iterativo di **Newton-Raphson**

$$x^{(k+1)} \;=\; x^{(k)} - J^{-1}(x^{(k)})f(x^{(k)}) \qquad k = 0, 1, \ldots$$

Nell'uso pratico del metodo l'iterata $x^{(k+1)}$ si ricava dalla soluzione del sistema lineare

$$J(x^{(k)})d^{(k)} \;=\; -f(x^{(k)}) \qquad k = 0, 1, \ldots$$

dove $d^{(k)} = x^{(k+1)} - x^{(k)}$

Nel caso di una lenta variazione della matrice jacobiana $J(x)$, si può ricorrere ad un **metodo di Newton semplificato** della forma

$$J(x^{(0)})\, d^{(k)} = -f(x^{(k)}), \qquad k = 0, 1, \ldots$$

dove la matrice jacobiana è valutata una sola volta in una buona approssimazione iniziale $x^{(0)}$

Se il metodo di Newton semplificato converge, la convergenza è in generale lineare

Un modo di evitare il calcolo dell'intera matrice $J(x)$ ad ogni passo, consiste nel considerare l'i-esima equazione del sistema $f(x) = 0$ come una equazione nella sola incognita x_i ed applicare a ciascuna equazione del sistema il metodo di Newton per equazioni in una incognita.

Supposto che l'ordinamento delle equazioni sia tale che, in un dominio contenente α, si abbia

$$\frac{\partial f_i(x_1, \ldots, x_n)}{\partial x_i} \neq 0, \qquad i = 1, 2, \ldots, n$$

si ottiene il **metodo non lineare di Jacobi-Newton**

$$x_i^{(k+1)} = x_i^{(k)} - \frac{f_i(x_1^{(k)}, \ldots, x_n^{(k)})}{\frac{\partial f_i(x_1^{(k)}, \ldots, x_n^{(k)})}{\partial x_i}}, \quad i = 1, \ldots, n; \; k = 0, 1, \ldots$$

Una variante del precedente metodo si ottiene con il **metodo non lineare di Gauss-Seidel** che è della forma

$$x_i^{(k+1)} = x_i^{(k)} - \frac{f_i(x_1^{(k+1)}, \ldots, x_{i-1}^{(k+1)}, x_i^{(k)}, \ldots, x_n^{(k)})}{\frac{\partial f_i(x_1^{(k+1)}, \ldots, x_{i-1}^{(k+1)}, x_i^{(k)}, \ldots, x_n^{(k)})}{\partial x_i}},$$

$$i = 1, 2, \ldots, n, \; k = 0, 1, \ldots$$

dove si sono utilizzate, per il calcolo di $x_i^{(k+1)}$, le prime $i-1$ componenti già note dell'iterata $x^{(k+1)}$ in corso di calcolo

Una classe di metodi, mutuati dal metodo di Newton-Raphson e che possono essere interpretati come una estensione del metodo delle secanti in $\mathbb{R}^n$, si ottiene sostituendo la matrice $J(x)$ con una sua approssimazione discreta; si ha così l'evidente vantaggio di non dovere calcolare le n^2 derivate parziali costituenti $J(x)$

Per approssimare $J(x^{(k)})$ si può introdurre la matrice $\Delta(x^{(k)})$, reale di ordine n, i cui elementi sono approssimazioni delle derivate parziali $\frac{\partial f_i}{\partial x_j}$ (sulle dispense sono riportati alcuni esempi)

Outline

Studiamo il caso particolare in cui l'equazione da risolvere sia algebrica di grado $m \geq 2$, cioè della forma

$$(P(x)\ =)\ a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0\ =\ 0 \quad (a_m \neq 0)$$

con i coefficienti $a_i \in \mathbb{R}$, $i = 0, 1, \ldots, m$

Alle equazioni algebriche sono applicabili tutti i metodi visti nei paragrafi precedenti, ma le particolari proprietà della classe dei polinomi forniscono ulteriori procedimenti per la localizzazione e l'approssimazione delle radici

Uno strumento utile a tale scopo è l'utilizzo delle seguenti successioni

Definizione

Si dice **successione di Sturm** per il polinomio $P(x)$ con $a_i \in \mathbb{R}$, una successione di polinomi reali

$$P_0(x) := P(x),\ P_1(x), \ldots,\ P_k(x), \qquad (k \leq m)$$

per cui valgono le proprietà

1. $P_k(x)$ non ha zeri reali
2. $\alpha \in \mathbb{R}, P_r(\alpha) = 0$ per $1 \leq r \leq k-1$ implica $$P_{r-1}(\alpha)P_{r+1}(\alpha) < 0$$
3. $\alpha \in \mathbb{R}, P_0(\alpha) = 0 \implies P_0'(\alpha)P_1(\alpha) > 0$

Dalla terza proprietà segue che gli zeri reali di $P(x)$ sono tutti semplici

Teorema di Sturm

Se si dispone di una successione di Sturm relativa a $P(x)$, il numero degli zeri del polinomio $P(x)$ nell'intervallo $a < x \leq b$ è dato da $V(a) - V(b)$, dove $V(x)$ è il numero delle variazioni di segno presenti nella successione dei valori non nulli assunti dai polinomi della successione nel punto x (la dimostrazione si trova sulle dispense)

Corollario

Se i polinomi della successione di Sturm sono $m+1$ (quindi $m = k$) la successione si dice **completa** e se inoltre tutti i polinomi hanno i coefficienti dei termini di grado massimo dello stesso segno, allora l'equazione $P(x) = 0$ ha m radici reali e distinte e viceversa

Costruzione della successione di Sturm

Per costruire effettivamente una successione di Sturm per il polinomio $P(x)$, si pone

$$P_0(x) := P(x)\,, \quad P_1(x) := P'(x)\,,$$

$$P_{r-1}(x) = P_r(x)Q_r(x) - P_{r+1}(x)\,, \; r = 1, 2, \ldots,$$

dove $Q_r(x)$ e $-P_{r+1}(x)$ sono rispettivamente quoziente e resto della divisione $P_{r-1}(x)/P_r(x)$, $r = 1, 2, \ldots$

Il processo esposto ha termine, poiché il grado dei polinomi decresce al crescere dell'indice e perciò per un certo $k \leq m$ risulta

$$P_{k-1}(x) \;=\; P_k(x)Q_k(x)$$

Costruzione della successione di Sturm

Il processo di costruzione della successione di Sturm è riconducibile al noto **algoritmo di Euclide** che fornisce il massimo comune divisore di $P_0(x)$ e $P_1(x)$, cioè si ha

$$P_k(x) \; = \; \text{M.C.D.} \; \{P(x), P'(x)\}$$

Segue che, nel caso in cui $P(x)$ e $P'(x)$ non abbiano zeri reali in comune, i polinomi $P_0(x), P_1(x), \ldots, P_k(x)$ formano una successione di Sturm per $P(x)$ come risulta dalla prossima slide

Costruzione della successione di Sturm

$P_k(x)$ non ha zeri reali e quindi la proprietà 1 risulta verificata

Se $P_r(\alpha) = 0$ con $1 \leq r \leq k-1$, allora si ha
$P_{r-1}(\alpha) = -P_{r+1}(\alpha) \neq 0$ e quindi è verificata la proprietà 2

$$P_{r-1}(\alpha)P_{r+1}(\alpha) \; < \; 0$$

Dalla definizione di $P_0(x)$ e di $P_1(x)$, se $P_0(\alpha) = 0$, segue la proprietà 3

$$P_0'(\alpha)P_1(\alpha) \; = \; (P'(\alpha))^2 \; > \; 0$$

Se $P(x)$ e $P'(x)$ hanno zeri reali in comune, questi sono tutti zeri di $P_k(x)$ con la stessa molteplicità che hanno come zeri di $P'(x)$ e la successione ottenuta con l'algoritmo di Euclide **non** fornisce una successione di Sturm

In tal caso si può verificare che la successione

$$\frac{P_0(x)}{P_k(x)}, \quad \frac{P_1(x)}{P_k(x)}, \ldots, \frac{P_k(x)}{P_k(x)},$$

è una successione di Sturm per il polinomio $P(x)/P_k(x)$ che ha tanti zeri semplici quanti sono gli zeri distinti di $P(x)$

La differenza $V(a) - V(b)$ valutata per la nuova successione fornisce il numero delle radici reali e distinte (indipendentemente dalla loro molteplicità) dell'equazione $P(x) = 0$ sull'intervallo $a < x \leq b$

Una successione di Sturm può essere usata per individuare un intervallo $[a, b]$ contenente una sola radice reale α di una equazione algebrica e quindi, con successivi dimezzamenti dell'intervallo come nel metodo di bisezione, si può approssimare α con qualsiasi accuratezza

L'approssimazione di uno zero di $P(x)$, che sia stato separato per esempio con una successione di Sturm, può essere fatta usando uno qualunque dei metodi esposti precedentemente

In particolare, si può utilizzare il metodo di Newton

$$x_{n+1} = x_n - \frac{P(x_n)}{P'(x_n)},$$

che richiede ad ogni passo il calcolo di $P(x_n)$ e $P'(x_n)$

Si consideri la matrice tridiagonale hermitiana

$$T = \begin{pmatrix} a_1 & \overline{b}_2 & & \\ b_2 & \ddots & \ddots & \\ & \ddots & \ddots & \overline{b}_n \\ & & b_n & a_n \end{pmatrix}$$

con $b_i \neq 0$, $i = 2, 3, \ldots, n$

Il suo polinomio caratteristico $P(\lambda) = \det(A - \lambda I)$ può essere calcolato per ricorrenza ottenendo $P(\lambda) = P_n(\lambda)$ dove

$$P_j(\lambda) = (a_j - \lambda)P_{j-1}(\lambda) - \mid b_j \mid^2 P_{j-2}(\lambda), \qquad j = 2, 3, \ldots, n$$

con $P_0(\lambda) = 1$ e $P_1(\lambda) = a_1 - \lambda$

Teorema

Nelle ipotesi fatte su T, la successione

$$(-1)^n P_n(\lambda),\ (-1)^{n-1} P_{n-1}(\lambda), \dots, -P_1(\lambda),\ P_0(\lambda),$$

è una successione di Sturm per $P(\lambda)$ e gli zeri di $P(\lambda)$ sono distinti

Esempio 1

Studiare l'equazione algebrica

$$x^4 - 2\,x^3 + 3\,x^2 - 4\,x + 1 \;=\; 0$$

indicando a quale insieme appartengono le soluzioni

La successione di Sturm calcolata utilizzando l'algoritmo di Euclide è

- $P_0(x) \;=\; x^4 - 2x^3 + 3x^2 - 4x + 1$
- $P_1(x) \;=\; 2x^3 - 3x^2 + 3x - 2$
- $P_2(x) \;=\; -3x^2 + 9x - 2$
- $P_3(x) \;=\; -8x + 3$
- $P_4(x) \;=\; -61$

Esempio 1

Costruiamo il seguente quadro

$-\infty$	0	$+\infty$
$+$	$+$	$+$
$-$	$-$	$+$
$-$	$-$	$-$
$+$	$+$	$-$
$-$	$-$	$-$
3	3	1

ottenuto dal segno assunto dai polinomi dalla successione per $x \to -\infty$, $x = 0$, $x \to +\infty$ e l'ultima riga riporta i valori $V(-\infty)$, $V(0)$ e $V(+\infty)$

Esempio 1

Conclusione

Dal precedente quadro si deduce che, essendo

$$V(-\infty) - V(+\infty) = 2$$

e risultando $P_4(x)$ uguale ad una costante ($\neq 0$), l'equazione data ha 2 soluzioni reali distinte e 2 soluzioni complesse coniugate

Esempio 2

Studiare l'equazione algebrica

$$x^4 - 2\,x^3 + 3\,x^2 - 4\,x + \lambda \;=\; 0\,, \qquad \lambda \in \mathbb{R}$$

indicando a quale insieme appartengono le soluzioni al variare del parametro λ

La successione di Sturm calcolata utilizzando l'algoritmo di Euclide è

- $P_0(x) \;=\; x^4 - 2x^3 + 3x^2 - 4x + \lambda$
- $P_1(x) \;=\; 2x^3 - 3x^2 + 3x - 2$
- $P_2(x) \;=\; -3x^2 + 9x - 4\lambda + 2$
- $P_3(x) \;=\; (2\lambda - 10)x + 3\lambda \qquad (\lambda \neq 5)$
- $P_4(x) \;=\; (\lambda - 2)(16\lambda^2 - 55\lambda + 100)$

Esempio 2 — $\lambda \in]-\infty, 0[$

In questo caso si ha

$-\infty$	0	$+\infty$
$+$	$-$	$+$
$-$	$-$	$+$
$-$	$+$	$-$
$+$	$-$	$-$
$-$	$-$	$-$
3	2	1

per cui $V(-\infty) - V(+\infty) = 2$, $V(-\infty) - V(0) = 1$ e $V(0) - V(+\infty) = 1$

L'equazione ha due soluzioni reali (una positiva e una negativa) e due soluzioni complesse coniugate

Esempio 2 — $\lambda = 0$

In questo caso si ha

$-\infty$	0	$+\infty$
$+$	0	$+$
$-$	$-$	$+$
$-$	$+$	$-$
$+$	0	$-$
$-$	$-$	$-$
3	2	1

per cui $V(-\infty) - V(+\infty) = 2$, $V(-\infty) - V(0) = 1$ e $V(0) - V(+\infty) = 1$

L'equazione ha due soluzioni reali (una uguale a zero e una positiva) e due soluzioni complesse coniugate

Esempio 2 — $\lambda \in]0, 2[$

In questo caso si ha

$-\infty$	0	$+\infty$
$+$	$+$	$+$
$-$	$-$	$+$
$-$	$\pm$	$-$
$+$	$+$	$-$
$-$	$-$	$-$
3	3	1

per cui $V(-\infty) - V(+\infty) = 2$, $V(-\infty) - V(0) = 0$ e $V(0) - V(+\infty) = 2$

L'equazione ha due soluzioni reali (positive) e due soluzioni complesse coniugate

Esempio 2 — $\lambda = 2$

In questo caso si ha

$-\infty$	0	$+\infty$
$+$	$+$	$+$
$-$	$-$	$+$
$-$	$-$	$-$
$+$	$+$	$-$
2	2	1

per cui $V(-\infty) - V(+\infty) = 1$, $V(-\infty) - V(0) = 1$ e $V(0) - V(+\infty) = 0$

L'equazione ha due soluzioni reali coincidenti ($\alpha_1 = \alpha_2 = 1$) e due soluzioni complesse coniugate

Esempio 2 — $\lambda \in]2, 5[$

In questo caso si ha

$-\infty$	0	$+\infty$
$+$	$+$	$+$
$-$	$-$	$+$
$-$	$-$	$-$
$+$	$+$	$-$
$+$	$+$	$+$
2	2	2

per cui $V(-\infty) - V(+\infty) = 0$, $V(-\infty) - V(0) = 0$ e $V(0) - V(+\infty) = 0$
L'equazione ha due coppie di soluzioni complesse coniugate

Esempio 2 — $\lambda = 5$

In questo caso si ha

$-\infty$	0	$+\infty$
$+$	$+$	$+$
$-$	$-$	$+$
$-$	$-$	$-$
$+$	$+$	$+$
2	2	2

per cui $V(-\infty) - V(+\infty) = 0$, $V(-\infty) - V(0) = 0$ e $V(0) - V(+\infty) = 0$

L'equazione ha due coppie di soluzioni complesse coniugate

Esempio 2 — $\lambda \in]5, +\infty[$

In questo caso si ha

$-\infty$	0	$+\infty$
$+$	$+$	$+$
$-$	$-$	$+$
$-$	$-$	$-$
$-$	$+$	$+$
$+$	$+$	$+$
2	2	2

per cui $V(-\infty) - V(+\infty) = 0$, $V(-\infty) - V(0) = 0$ e $V(0) - V(+\infty) = 0$

L'equazione ha due coppie di soluzioni complesse coniugate

Esempio 2

Conclusione

Riassumiamo i risultati:

- $\boxed{\lambda \in]-\infty, 0[}$ $\longrightarrow$ 2 soluzioni reali (pos. e neg.) e 2 soluzioni complesse coniugate
- $\boxed{\lambda = 0}$ $\longrightarrow$ $\alpha_1 = 0$, α_2 reale pos. e 2 soluzioni complesse coniugate
- $\boxed{\lambda \in]0, 2[}$ $\longrightarrow$ 2 soluzioni reali pos. e 2 soluzioni complesse coniugate
- $\boxed{\lambda = 2}$ $\longrightarrow$ $\alpha_1 = \alpha_2 = 1$ e 2 soluzioni complesse coniugate
- $\boxed{\lambda \in]2, +\infty[}$ $\longrightarrow$ 2 coppie di soluzioni complesse coniugate