SULLA CONVERGENZA DI CERTI METODI ITERATIVI (*)

di IGOR MORET (a Trieste) (**)

SOMMARIO. - *Si presenta un metodo per provare la convergenza di certi processi iterativi e lo si applica ai metodi di tipo Newtoniano della forma* $x_{n+1} = x_n - A(x_n)^{-1} F(x_n)$, *ottenendo al tempo stesso delle maggiorazioni a posteriori dell'errore in senso stretto.*

SUMMARY. - *A method to prove the convergence of certain iterative processes is presented and applied to Newton-Type Methods of the form* $x_{n+1} = x_n - A(x_n)^{-1} F(x_n)$, *moreover this allows to obtain sharp aposteriori error bounds.*

Introduzione. - Si voglia risolvere l'equazione

$$F(x) = 0$$

in uno spazio di Banach utilizzando allo scopo un classico metodo iterativo, basato sul Principio delle Contrazioni, che qui scriviamo nella forma

$$x_{n+1} = x_n - T(x_n) \tag{1}$$

L'applicazione di un tale metodo implica l'esame dei due seguenti aspetti fondamentali:

1) Stabilire condizioni tali da garantire la convergenza della successione generata dalla (1) ad x^* tale che $F(x^*) = 0$.

2) Ottenere una stima quanto più stretta possibile di $\| x_n - x^* \|$.

In letteratura, le indagini attorno a questi due problemi si trovano in prevalenza sviluppate separatamente. Con riferimento al

---

(*) Pervenuto in Redazione il 30 aprile 1983.

(**) Indirizzo dell'Autore: Istituto di Elettrotecnica ed Elettronica dell'Università degli Studi - Via Valerio, 10 - 34100 Trieste.

primo, si è rivelata particolarmente efficace, dal punto di vista dimostrativo, la tecnica delle «maggioranti non lineari», dettagliatamente illustrata in [5 cap. 12] e già utilizzata da Rheinboldt [9] per una teoria generale di convergenza di metodi della forma

$$x_{n+1} = x_n - A_n^{-1} F(x_n)$$

ossia per varianti del classico metodo di Newton

$$x_{n+1} = x_n - F'(x_n)^{-1} F(x_n)$$

nei quali A_n sostituisce $F'(x_n)$.

Il secondo aspetto, ossia l'analisi dell'errore, è stato sviluppato da diversi autori, prendendo prevalentemente in esame il metodo di Newton o sue varianti e ponendo l'accento in particolare sull'influsso degli errori aritmetici di arrotondamento. Citiamo fra gli altri Lancaster [3], Rockne [10], Gragg e Tapia [1] e, fra i lavori più recenti, Miel [4].

Lo spunto per questa nota viene da un articolo di Potra e Ptak [6], al quale si rifanno anche Lai e Wu in [2]. In esso viene messa a punto una tecnica dimostrativa, basata sul cosiddetto «metodo di induzione matematica non discreta», già descritto in [7], che permette la contemporanea risoluzione dei problemi 1) e 2) limitatamente al già citato metodo di Newton. Nel presente articolo questa tecnica viene ripresa e rielaborata, allo scopo di collegarla più strettamente ai processi iterativi del tipo (1) e trova applicazione nella analisi dei metodi di tipo Newtoniano della forma

$$x_{n+1} = x_n - A(x_n)^{-1} F(x_n).$$

In questi casi si dà, assieme ad un teorema di convergenza, che estende i risultati di Rheinboldt ([9] Theor. 4.3) e Lai e Wu ([2] Theor. 5.2), anche una maggiorazione dell'errore che può essere considerata, in base alle ipotesi assunte, la più stretta possibile.

1. - Sia b un numero reale positivo e sia $I_b = \{t : 0 \leqslant t < b\}$. Indichiamo quindi con C_b l'insieme delle funzioni $u(t)$, continue, di I_b in R e derivabili su $I_b - \{0\}$, tali che se $t > 0$ sia:

i) $u(t) > t$,
ii) $u'(t) > 1$.

Ogni funzione appartenente a C_b risulta dunque invertibile su I_b ed inoltre, detta $v(s)$ la sua inversa, la funzione $w(s) = s - v(s)$ è isotona nell'insieme $u(I_b)$.

Siano $u(t) \in C_b$, $0 < t_o < b$, $s_o = u(t_o)$. E' evidente che nell'intervallo chiuso $[u(0), s_o]$ la funzione $w(s)$ soddisfa le ipotesi del Lemma di Kantorovich ([5] p. 417) sicché la successione $\{s_n\}_{n=o}^{\infty}$ generata dalla forma iterativa

$$s_{n+1} = w(s_n) = s_n - v(s_n) \tag{1.1}$$

è decrescente e converge ad s^* tale che $v(s^*) = 0$, ossia $s^* = u(0)$.

Per ogni indice n poniamo $t_n = s_n - s_{n+1} = v(s_n)$, e quindi

$$t_{n+1} = v(s_{n+1}) = v(s_n - t_n) = v(u(t_n) - t_n), \tag{1.2}$$

o se si vuole

$$u(t_{n+1}) = u(t_n) - t_n. \tag{1.3}$$

Posto dunque

$$r(t) = v(u(t) - t), \tag{1.4}$$

in virtù della (1.2) restano verificate le seguenti relazioni:

$$t_{n+k} = r^k(t_n) \qquad (k > 0) \tag{1.5}$$

$$s_n - s^* = u(t_n) - u(0) = \sum_{i=n}^{\infty} t_i = \sum_{k=0}^{\infty} r^k(t_n) \tag{1.6}$$

dove si intenda $r^0(t) = t$, $r^k(t) = r(r^{k-1}(t))$ per $k \geqslant 1$.

Quanto detto mostra come si possano costruire quelle funzioni che Potra e Pták chiamano «rates of convergence», in [6].

Per il seguito X indicherà uno spazio di Banach, $T(x)$ una funzione continua di un dominio chiuso $G \subset X$ in X, e $t(x) = \|T(x)\|$.

Siano poi $x_o \in G$, $t_o \geqslant t(x_o)$, $u(t) \in C_b$ con $b > t_o$; in relazione a cui definiamo l'insieme:

$$U = U(x_o, T, u) = \{x \in G : t(x) \leq t_o \,;\, \|x - x_o\| \leq u(t_o) - u(t(x))\}. \tag{1.7}$$

Quindi per ogni $x \in U$ definiamo i due insiemi:

$$V(x) = \{x \in G : \|z - x\| \leq t(x)\}$$

$$W(x) = \{z \in G : t(z) \leq r(t(x))\}.$$

Lemma 1.1. *Sia* $x \in U, E(x) = V(x) \cap W(x) \neq \phi$ *e* $z \in E(x)$, *allora* $z \in U$.

Dim. Essendo $z \in V(x)$ avremo

$$\|z - x_o\| \leq \|z - x\| + \|x - x_o\| \leq \|x - x_o\| + t(x)$$

e poiché $z \in W(x)$ si ha $u(t(z)) \leq u(t(x)) - t(x)$, per cui

$$\|z - x_o\| \leq u(t_o) - u(t(x)) + t(x) \leq u(t_o) - u(t(z)).$$

Teorema 1.2. *Per ogni* $x \in U$ *sia* $z = x - T(x) \in W(x)$. *Allora la successione generata dalla*

$$x_{n+1} = x_n - T(x_n) \tag{1.8}$$

è convergente e, detto x^* il suo punto limite, per ogni indice n si ha

$$\| x_n - x^* \| \leq u(t(x_n)) - u(0) \tag{1.9}$$

Dim. Osserviamo per prima cosa che per ogni $x \in \mathrm{U}$ l insieme $\mathrm{E}(x)$ è non vuoto in quanto $z = x - T(x)$ vi appartiene.

Quindi, essendo, come ovvio, x_o (punto iniziale della (1.8)) appartenente ad U, in virtù del Lemma 1.1, induttivamente avremo per ogni indice n:

$$t(x_{n+1}) = \| x_{n+1} - x_{n+2} \| \leq r(t(x_n)) \tag{1.10}$$

Poiché inoltre la funzione $r(t)$ è monotona crescente su I_b, per ogni intero $k > 0$ risulterà

$$t(x_{n+k}) \leq r^k(t(x_n)) \qquad n = 0, 1, 2, \ldots\ldots \tag{1.11}$$

e pertanto, poiché in virtù della (1.5) si ha $\lim\limits_{k \to \infty} r^k(t_o) = 0$, la successione generata dalla (1.8) converge.

Infine, viste le (1.11) e (1.6) si ha:

$$\| x_n - x^* \| \leqslant \sum_{k=0}^{\infty} r^k(t(x_n)) = u(t(x_n)) - u(0).$$

OSSERVAZIONE. Da un punto di vista pratico sembra più logico, ai fini di una stima a posteriori dell'errore $\| x_n - x^* \|$, esprimere questo in relazione allo scarto $t(x_{n-1}) = \| x_{n-1} - x_n \|$ anziché a $t(x_n)$.

Valendo le ipotesi del teorema 1.1 si ha per ogni $n \geqslant 1$:

$$u(t(x_n)) \leq u(t(x_{n-1})) - t(x_{n-1})$$

e pertanto dalla (1.9) si ottiene

$$\| x_n - x^* \| \leq u(t(x_{n-1})) - t(x_{n-1}) - u(0). \tag{1.12}$$

Andiamo ora ad esaminare una particolare scelta di $u(t)$.

TEOREMA 1.3. *Siano a, c, d numeri reali non negativi con $a+c>0$, tali che la funzione di variabile reale*

$$u(t) = at + (a^2 t^2 + ct + d^2)^{1/2} \tag{1.13}$$

appartenga a C_b per qualche $b > t_0 \geqslant t(x_0)$ e per ogni $x \in \mathrm{U}$ (relativo alla (1.13)), posto $z = x - T(x)$, sia $z \in \mathrm{G}$ e

$$t(z) \leq \frac{t(x)\,(t(x) - 2(1-a)\,u(t(x)) + c)}{2a(u(t(x)) - t(x)) + c} \tag{1.14}$$

allora la successione generata dalla (1.8) converge e, detto x^* il limite, per ogni indice n valgono le (1.9) e (1.12).

Dim. L'inversa della (1.13) è

$$v(s) = (s^2 - d^2) \,/\, (2as - c) \tag{1.15}$$

ed è presto visto che il secondo membro della (1.14) rappresenta la funzione $r(t(x)) = v(u(t(x)) - t(x)$, sicché si verificano le ipotesi del teorema 1.2.

Il seguente risultato, che troverà applicazione nel successivo paragrafo, dà delle condizioni affinché la (1.13) definisca una funzione che appartiene a C_b.

TEOREMA 1.4. *Siano* $t_0 > 0, g > 0, 0 < h \leqslant 1$ *tali che*

$$m = 2\, g\, t_o \,/\, h^2 \leq 1 \tag{1.16}$$

e posto $e = 1 - m$, *sia* $d = h\, e^{1/2} \,/\, g$, *sia poi* $0 \leq a \leq 1$ *ed infine* $c = 2\,(1 - ah) \,/\, g$, *allora la funzione* (1.13) *appartiene a* C_b *con* $b > t_o$.

Dim. Notiamo che $u(t) > t$ se

$$p(t) = (2a - 1)\, t^2 + c\, t + d^2 > 0 \tag{1.17}$$

e che $u'(t) > 1$ se

$$q(t) = 4a^2\, t^2 (2a-1) + 4ct(2a-1) + c^2 - 4(1-a)^2\, d^2 > 0 \tag{1.18}$$

Si può facilmente constatare che se $1/2 \leq a \leq 1$ le (1.17) e (1.18) valgono per ogni $t > 0$.

Nel caso $0 \leqslant a < 1/2$, entrambi i polinomi $p(t)$ e $q(t)$ possiedono una ed una sola radice positiva che, come si desume dopo semplici calcoli è certamente maggiore di t_o, ove essi assumono valore positivo, da ciò si deduce l'esistenza di $b > t_o$ tale che $u(t) \in C_b$.

2. - Sia Y uno spazio di Banach ed $F : D \subset X \rightarrow Y$ una funzione differenziabile secondo Fréchet in un convesso $D_0 \subset D$. Sia poi $L(X, Y)$ lo spazio di Banach degli operatori lineari limitati da X in Y. Sia $A : D_0 \rightarrow L(X, Y)$ ed $\Omega(A)$ l'insieme degli $x \in D_0$ per cui $A(x)^{-1} \in L(Y, X)$ esiste. Supposto che $\Omega(A)$ sia non vuoto, per ogni $x \in \Omega(A)$ definiamo

$$\Omega(x, A) = \underset{\substack{y \in D_0 \\ y \neq x}}{Sup} \frac{\| A(x)^{-1} (A(x) - A(y)) \|}{\| x - y \|}$$

$$\Gamma(x, A, F') = \underset{\substack{y, y' \in D_0 \\ y \neq y'}}{Sup} \frac{\| A(x)^{-1} (F'(y) - F'(y')) \|}{\| y - y' \|}$$

LEMMA 2.1 *Sia* $x_0 \in \Omega(A)$ *e* $\Gamma(x_0, A, F') \leqslant g$. *Allora si ha*

$$\| A(x_0)^{-1} [F(y) - F(x) - F'(x)\,(y - x)] \| \leqslant (g/2)\, \| y - x \|^2, \; x, y \in D_0 \tag{2.1}$$

Dim. La (2.1) si ottiene direttamente dalla definizione di $\Gamma(x_0, A, F')$, applicando un noto risultato dell'Analisi (vedi ad es. [9] lemma 3.1 (c)) alla funzione $G(x) = A(x_0)^{-1} F(x)$.

LEMMA 2.2 *Valgano le ipotesi del lemma 2.1 e sia inoltre* $\Psi(x_0, A) \leqslant p$. *Allora per ogni* $z \in \mathrm{D}_0$ *tale che* $\| z - x_0 \| < 1/p$ *si ha che* $z \in \Omega(A)$ *ed inoltre*

$$\| A(z)^{-1} [F'(x) - A(x)] \| \leq \| A(x_0)^{-1} [F'(x) - A(x)] \| / (1 - p \| z - x_0 \|) \tag{2.2}$$

$$\| A(z)^{-1} [F(y) - F(x) - F'(x)\,(y-x)] \| \leq (g/2) \| y - x \|^2 / (1 - p \| z - x_0 \|) \tag{2.3}$$

per ogni $x, y \in \mathrm{D}_0$.

Dim. Poiché $\| A(x_0)^{-1} (A(x_0) - A(z)) \| \leqslant p \| z - x_0 \| < 1$, per il noto «lemma di perturbazione (vedi [8] Th. 9.1) si ha che $z \in \Omega(A)$ ed inoltre che $\| [I - A(x_0)^{-1} (A(x_0) - A(z))]^{-1} \| \leq 1/(1 - p \| z - x_0 \|)$. Da quest'ultima e dalla identità

$$A(z)^{-1} = [I - A(x_0)^{-1} (A(x_0) - A(z))]^{-1} A(x_0)^{-1}$$

segue la (2.2) ed infine, utilizzando la (2.1), la (2.3).

Il teorema seguente può essere considerato una estensione del Teorema 4.3 di Rheinboldt [9] (vedi anche Teor. 12.6.4 in [5]). Esso costituisce un'applicazione dei risultati del paragrafo precedente al caso $T(x) = A(x)^{-1} F(x)$ con $A(x) \in \mathrm{L}(\mathrm{X}, \mathrm{Y})$.

TEOREMA 2.3. *Sia* $x_0 \in \Omega(A)$, $g > 0$, $0 < h \leq 1$, $0 \leq a \leq 1$ *tali che*

$$\Gamma(x_0, A, F') \leq g, \tag{2.4}$$

$$\Psi(x_0, A) \leq ag, \tag{2.5}$$

$$\| A(x_0)^{-1} [F'(x) - A(x)] \| \leq (1-h) + (1-a)\, g \| x - x_0 \|, \tag{2.6}$$

Si abbia inoltre $\| A(x_0)^{-1} F(x_0) \| \leqslant t_0$, $m = 2 g t_0 / h^2 \leqslant 1$. *Considerata poi la funzione (1.13) con c e d definiti come nel teorema 1.4, sia* $R = u(t_0) - u(0) = h/g - d$ *e sia* $B = \{x : \| x - x_0 \| \leq R\} \subset \mathrm{D}_0$. *Allora resta definita la successione generata da*

$$x_{n+1} = x_n - A(x_n)^{-1} F(x_n) \quad n = 0, 1, 2, \ldots \tag{2.7}$$

essa rimane in B *e converge ad* x^* *tale che* $F(x^*) = 0$. *Inoltre valgono le relazioni*

$$\| x_n - x \| \leqslant u(\| x_n - x_{n+1} \|) - d. \tag{2.9}$$

$$\| x_n - x \| \leqslant u(\| x_{n-1} - x_n \|) - \| x_{n-1} - x_n \| - d. \tag{2.10}$$

Dim. Consideriamo l'insieme U definito dalla (1.7). Esso è incluso in B. Sia $x \in \mathrm{U}$ con $T(x) = A(x)^{-1} F(x) \neq 0$. Posto $z = x - T(x)$

proveremo che esiste $T(z)$ e vale la (1.14).

Per prima cosa, ricordando che $u(t_0) = h/g$, avremo

$$\| z-x_0 \| \leqslant \| x-x_0 \| + t(x) \leqslant u(t_0) - u(t(x)) + t(x) < h/g \leqslant 1/ag$$

sicché per il lemma 2.1 $A(z)^{-1}$ esiste. Avremo poi

$$A(z)^{-1}F(z) = A(z)^{-1}[F(z)-F(x)-F'(x)(z-x)+(F'(x)-A(x))(z-x)]$$

Quindi, applicando la (2.3) e la (2.2):

$$\| A(z)^{-1} F(z) \| \leqslant \frac{g\,t(x)^2 + [(1-h) + (1-a)\,g\,\| x-x_0 \|\,]\,t(x)}{2(1-ag\,\| z-x_0 \|)}$$

$$\leqslant \frac{t(x)\,[g\,t(x) + (1-h) + (1-a)\,g(u(t_0) - u(t(x)))]}{2\,[1 - ag(u(t_0) - u(t(x)) + t(x))]}$$

vale a dire la (1.14). Dunque per il teorema 1.2 la successione generata dalla (2.7) rimane in B e converge. Essendo poi per ogni n

$$\| A(x_0)^{-1} F(x_n) \| \leqslant t(x_n) + \| A(x_0)^{-1} (A(x_n) - A(x_0)) \| \, t(x_n)$$
$$\leqslant (1 + agR)\, t(x_n),$$

detto x^* il limite, sarà $F(x^*) = 0$. Le (1.9) e (1.12) danno infine le maggiorazioni (2.9) e (2.10).

Nel caso $A(x) = F'(x)$ le ipotesi assunte e le maggiorazioni a posteriori dell'errore ottenute per il metodo di Newton in [6] rientrano in quelle del teorema ora dimostrato.

Per concludere, osserviamo che le stime (2.9) e (2.10) possono considerarsi delle maggiorazioni dell'errore in senso stretto. Infatti, se valgono le ipotesi del teorema 1.3, per quanto visto nella dimostrazione del teorema 1.2, è facile verificare che, se scegliamo $F(x) = x^2 - d^2$ e $A(x) = 2ax + c$, tutte le ipotesi del teorema 2.3 sono verificate e le (2.9) e (2.10) valgono col segno di uguaglianza.

Considerazioni più dettagliate circa gli aspetti computazionali inerenti alle maggiorazioni d'errore sopra ottenute, appariranno in un prossimo lavoro.

BIBLIOGRAFIA

[1] GRAGG W. B., TAPIA R. A., *Optimal error bounds for the Newton-Kantorovich theorem.* SIAM J. Numer. Anal. 11, 10-13 (1974).

]2[LAI H. C., WU P. Y., *Error bounds of Newton Type process in Banach spaces.* Numer. Math. 39, 175-193 (1982).

[3] LANCASTTER P., *Error analysis for the Newton-Raphson method.* Numer. Math. 9, 55-68 (1966).

[4] MIEL G. J., *Unified error analysis for Newton-Type methods.* Numer. Math. 33, 391-396 (1979).

[5] ORTEGA J. M., RHEINBOLDT W. C., *Iterative Solution of Nonlinear Equations in Several Variables.* Academic Press, New York (1970).

[6] POTRA F. A., PTÁK V., *Sharp error bounds for Newton process.* Numer. Math. 34, 63-72 (1980).

[7] PTÁK V., *Non discrete Mathematical Induction and iterative existence proofs.* Linear Algebra and Appl. 13, 223-236 (1976).

[8] RALL L. B., *Computational Solution of Nonlinear Operator Equations.* Wiley & Sons, New York (1968).

[9] RHEINBOLDT W. C., *A unified convergence theory for a class of iterative processes.* SIAM J. Numer. Anal. 5, 42-63 (1968).

[10] ROCKNE J., *Newton's method under mild differentiability conditions with error analysis.* Numer. Math. 18, 401-412 (1972).