Bollettino Unione Matematica Italiana

Sezione A – La Matematica nella Società e nella Cultura

Gaetano Zanghirati

Metodi paralleli per sistemi di equazioni non lineari bordati a blocchi

Bollettino dell'Unione Matematica Italiana, Serie 8, Vol. **1-A**—*La Matematica nella Società e nella Cultura* (1998), n.1S (Supplemento Tesi di Dottorato), p. 205–208.

Unione Matematica Italiana

<http://www.bdim.eu/item?id=BUMI_1998_8_1A_1S_205_0>

Bollettino dell'Unione Matematica Italiana, Unione Matematica Italiana, 1998.

La matematica nella Società e nella Cultura
Bollettino U. M. I.
(8) 1-A Suppl. (1998), pag. 205-208

Metodi paralleli per sistemi di equazioni non lineari bordati a blocchi.

GAETANO ZANGHIRATI

1. – Introduzione.

I metodi iterativi di tipo quasi–Newton risultano particolarmente adatti ad affrontare, su architetture parallele a memoria distribuita, il problema della risoluzione numerica di sistemi di equazioni non lineari differenziabili con matrice Jacobiana bordata a blocchi. In questo caso il sistema $F(x) = 0$ può, eventualmente dopo un riordinamento delle equazioni e delle variabili, essere scritto nella forma

$$\begin{cases} f_i(x_i, x_{q+1}) = 0; \qquad i = 1, \dots, q \\ f_{q+1}(x_1, \dots, x_q, x_{q+1}) = 0 \end{cases}$$

dove si è operato un *partizionamento* delle variabili e delle componenti della funzione $F\colon \mathbb{R}^n \mapsto \mathbb{R}^n$ in $q+1$ insiemi disgiunti. La matrice Jacobiana è costituita da un certo numero di blocchi quadrati, disgiunti, sulla diagonale principale, più un insieme finale di righe e colonne non nulle:

$$J(x) = \begin{bmatrix} A_1 & & & & B_1 \\ & A_2 & & & B_2 \\ & & \ddots & & \vdots \\ & & & A_q & B_q \\ C_1 & C_2 & \cdots & C_q & P \end{bmatrix}.$$

Problemi di questo tipo compaiono largamente nella discretizzazione di equazioni alle derivate parziali, nei metodi di collocazione e decomposizione del dominio, nel progetto di circuiti VLSI, in ingegneria strutturale, nella progettazione di reti di distribuzione ed in altri campi ancora. Date le grandi dimensioni tipiche di questi problemi, hanno particolare interesse metodi che tengono conto dalla struttura e del parallelismo insito nel sistema.

2. – L'approccio implicito.

Fondamentalmente, l'idea che si segue è quella di considerare, ad ogni iterazione, i q sottoinsiemi x_i di variabili come funzioni di x_{q+1} fissato. Sotto ipotesi poco restrittive, è lecito applicare il metodo di Newton per ricavare delle approssimazioni di $x_i(x_{q+1}^k)$ per $i = 1, \dots, q$. Sostituendo queste nel $(q+1)$-esimo sottoinsieme f_{q+1} si ottiene un sistema non lineare nelle sole variabili x_{q+1}^k, che può essere risolto con un metodo quasi–Newton con ricerca in linea, ottenendo così un metodo globalmente convergente. Con questo *disaccoppiamento* di variabili e funzioni, i q sistemi non lineari $f_i(x_i, x_{q+1}) = 0$ possono essere trattati indipendentemente e simultaneamente, dando origine a metodi *multilivello*, in cui si hanno iterazioni «interne» (quelle di $f_i = 0$) ed iterazioni «esterne» (quelle su f_{q+1}). Il primo lavoro in questo senso è di Rabbat *et al.* [3] nel campo dell'analisi di grandi circuiti non lineari. Zhang, Byrd e Schnabel [4] propongono algoritmi paralleli quasi-Newton globalmente convergenti per il caso particolare in cui il $(q+1)$-esimo insieme di equazioni è lineare.

I casi di malcondizionamento vengono trattati con un approccio classico di tipo Levenberg–Marquardt. È tuttavia difficile rendere compatibile quest'approccio con il disaccoppiamento di variabili e funzioni e con la convergenza globale. In questa direzione D. Feng ed R. B. Schnabel [1] propongono un algoritmo parallelo robusto globalmente convergente noto come *algoritmo PICRN (Parallel Implicit Corrected Robust Newton algorithm)*: la direzione di ricerca da esso generata ad ogni iterazione è $d^k = -J(x^k)^{-1}\overline{F}(x^k)$ nel caso ben condizionato, mentre nel caso malcondizionato viene scelta fra

$$\Delta_R x^k = -\left(J(x^k)^T J(x^k) + D^{(k)}\right)^{-1} J(x^k)^T \overline{F}(x^k) \quad \text{e}$$

$$\Delta_N x^k = -\left(J(x^k)^T J(x^k) + D^{(k)}\right)^{-1} J(x^k)^T F(x^k),$$

dove $D^{(k)}$ è una matrice diagonale di perturbazione (non negativa) ed il contributo delle iterazioni interne è contenuto in $\overline{F}(x^k)$. Per i casi ben condizionati si prova la convergenza locale almeno superlineare:

TEOREMA 1. – *Se esiste una costante $\eta > 0$ indipendente da k tale che $\|J(x^k)\|_2 \leq \eta$ $\forall k$, $f_i(x)$ è differenziabile con continuità e $\nabla f_i(x)$ di Lipschitz $\forall i$, x^* è una soluzione del sistema $F(x) = 0$ tale che $J(x^*)$ e tutti i suoi blocchi diagonali sono non singolari, allora il passo generato dall'algoritmo PICRN si avvicina asintoticamente al passo di Newton standard qualsiasi sia il numero j_i di iterazioni interne compiute.*

3. – La stabilizzazione non monotona.

Negli ultimi anni è stata dimostrata l'efficacia della *stabilizzazione non monotona* dei metodi per la soluzione di equazioni non lineari. In generale un tecnica di stabilizzazione non monotona è descritta da una direzione di ricerca d, da una funzione obiettivo $g(x)$ e da una funzione ausiliaria $\mathcal{A}(x, d)$ [2]. La convergenza globale è assicurata senza richiedere ad ogni iterazione la discesa monotona della funzione obiettivo. La chiave che permette di ottenere metodi quasi–Newton globalmente convergenti è che le direzioni di ricerca generate nelle iterazioni siano *gradiente–connesse* (Bertsekas, 1982). Si dimostra che tutte le direzioni di ricerca generate dall'algortimo di Feng–Schnabel sono gradiente–connesse: in questo modo non solo si ottiene per altra via la convergenza globale dell'algoritmo, ma è anche possibile inserirlo in un ambito teorico più generale (Ortega-Rheinboldt, 1970), dal quale discendono immediatamente forti risultati di convergenza. Un'attenta analisi delle condizioni imposte per la convergenza globale suggerisce l'introduzione di una condizione del tipo

$$\zeta_1 \|J(x^k)^T F(x^k)\|_2^{u_1} \leq \|d^k\|_2 \leq \zeta_2 \|J(x^k)^T F(x^k)\|_2^{u_2},$$

con $0 < \zeta_1 < 1$, $\zeta_2 > 1$, $0 < u_1 \leq 1$, $u_2 \geq 1$, che può essere resa arbitrariamente poco restrittiva. Detto PICRMN l'algoritmo così modificato, si dimostra che

TEOREMA 2. – *Tutte le direzioni di ricerca generate dall'algoritmo PICRMN sono uniformemente gradiente-connesse.*

Una conseguenza importante è che, definite le seguenti tre *diverse* funzioni ausiliarie

$$\mathcal{A}_1(x^k, d^k) = \tau_1 \tau_2^{-2} \|J(x^k)\, d^k\|_2^2,$$

$$\mathcal{A}_2(x^k, d^k) = \|J(x^k) d^k\|_2^2 + \|(D^{(k)})^{1/2} d^k\|_2^2,$$

$$\mathcal{A}_3(x^k, d^k) = \sigma \|J(x^k)^T F(x^k)\|_2 \|d^k\|_2,$$

dove τ_1, τ_2 e σ sono costanti, l'algoritmo PICRMN non monotono che ne deriva risulta globalmente convergente:

TEOREMA 3. – *Sia $f_i(x)$ differenziabile con continuità e $\nabla f_i(x)$ di Lipschitz per $i = 1, \ldots, q+1$ e sia $\{w(k)\}$ una successione di interi tali che $w(k) \in \{1, 2, 3\}$ $\forall k$. Consideriamo l'algoritmo di stabilizzazione non monotona con funzione obiettivo $g(x) = 1/2\|F(x)\|^2$, direzione di ricerca d^k generata dall'algoritmo PICRMN e funzione ausiliaria alla k-esima iterazione tale che*

$$\mathcal{A}_{w(k)}(x^k, d^k) = \begin{cases} \mathcal{A}_1(x^k, d^k) & se\ d^k = -J(x^k)^{-1}\overline{F}(x^k) \\ \mathcal{A}_2(x^k, d^k) & se\ d^k = \Delta_N x^k \\ \mathcal{A}_3(x^k, d^k) & se\ d^k = \Delta_R x^k. \end{cases}$$

Supponiamo che l'insieme di livello $\Omega_0 = \{x \in \mathbb{R}^n \mid g(x) \leq g(x^0)\}$ sia limitato e che esista una costante η indipendente da k tale che $\|J(x^k)\|_2 \leq \eta$: allora la successione generata $\{x^k\}$ è tale che

i) *ogni punto di accumulazione x^* di $\{x^k\}$ è un punto critico di $g(x)$;*

ii) *se x^* è un punto di accumulazione di $\{x^k\}$ dove vale anche che $\nabla g(x) = 0 \Rightarrow g(x) = 0$, allora x^* è anche soluzione del sistema non lineare $F(x) = 0$;*

iii) *se x^* è un punto di accumulazione di $\{x^k\}$ dove la matrice Jacobiana $J(x^*)$ è non singolare allora la successione $\{x^k\}$ converge superlinearmente ad x^*.*

Si dimostra inoltre un ulteriore risultato di convergenza in presenza di precondizionamento diagonale dei sistemi.

4. – Esperienza numerica.

Un'iniziale sperimentazione numerica dell'algoritmo PICRN su significativi problemi test di piccole dimensioni, con diverso grado e distribuzione della non linearità, ha permesso di valutare l'influenza dei molti parametri che caratterizzano il metodo. Una successiva sperimentazione parallela dell'algoritmo sull'architettura MIMD del Cray T3D, con un sistema non lineare proveniente dalla simulazione di circuiti VLSI, ha mostrato che l'efficienza maggiore si ha per un basso numero di iterazioni interne (3 o 4) e che, nei casi mal condizionati, la necessità di trattare un sistema «perturbato» porta ad un degrado delle prestazioni teoriche, rendendo dominante la parte sequenziale dell'algoritmo. Recenti ulteriori investigazioni numeriche evidenziano come la stabilizzazione non monotona dell'algoritmo PICRMN permette di ovviare alle difficoltà dovute al malcondizionamento e di migliorare la convergenza anche nei casi ben condizionati, mantenendo inalterato il grado di parallelismo.

5. – Conclusioni e sviluppi.

I risultati ottenuti e le estensioni teoriche presentate, oltre confermare che l'approccio parallelo implicito non monotono è promettente per i sistemi bordati a blocchi, suggeriscono alcune direzioni per ulteriori approfondimenti, fra cui: l'analisi teorica della convergenza del caso generale in presenza di precondizionamento dei sistemi e nel caso di impiego di solutori iterativi arrestati dopo poche iterazioni; l'impiego di altre note tecniche di globalizzazione, quali la *trust region*, sia nell'algoritmo monotono che in quello non monotono; l'ampliamento della sperimentazione numerica, sia rispetto alle caratteristiche della funzione non lineare che rispetto alla dimensione del sistema.

BIBLIOGRAFIA

[1] FENG D. and SCHNABEL R.B., *Globally convergent parallel algorithms for solving block bordered systems of nonlinear equations*, Optim. Meth. and Soft., **2** (1993), 269-295.

[2] FERRIS M.C. and LUCIDI S., *Nonmonotone stabilization methods for nonlinear equations*, J. Optim. Th. Appl., **81** (1994), 53-71.

[3] RABBAT N.B.G., SANGIOVANNI-VINCENTELLI A.L. and HSIEH H.Y., *A multilevel Newton algorithm with macromodeling and latency for the analysis of large-scale nonlinear circuits in the time domain*, IEEE Trans. Circuits and Systems, **26** (1979), 733-740.

[4] ZHANG X., BYRD R.H. and SCHNABEL R.B., *Parallel methods for solving nonlinear block bordered systems of equations*, SIAM J. Sci. Stat. Comput., **13** (1992), 841 859.

Dipartimento di Matematica, Università di Ferrara
e-mail: zng@dns.unife.it, jango208dm.unife.it
Dottorato in Matematica Computazionale e Informatica Matematica
(sede amministrativa: Padova) - Ciclo VIII
Direttore di ricerca: Prof. Ilio Galligani
Correlatore: Prof. Valeria Ruggiero