Metodi di tipo Newton interior point in ottimizzazione vincolata nonlineare di grandi dimensioni

Silvia Bonettini

13 aprile 2006

1. – Presentazione del problema e dei metodi di Newton interior point

Questa tesi riguarda l'analisi e lo sviluppo di metodi interior point per la soluzione numerica di problemi di programmazione nonlineare (NLP). Un problema NLP può essere formulato nel modo seguente:

$$\begin{array}{ll} \min & f(x) \\ \text{t.c.} & g_1(x) = 0 \\ & g_2(x) - s = 0 \\ & s \geq 0 \end{array} \tag{1}$$

dove $f : \mathbb{R}^n \to \mathbb{R}$ è la funzione obiettivo, $g_1 : \mathbb{R}^n \to \mathbb{R}^{neq}$ e $g_2 : \mathbb{R}^n \to \mathbb{R}^m$ rappresentano i vincoli di uguaglianza e disuguaglianza rispettivamente e il vettore $s \in \mathbb{R}^m$ contiene le *variabili slack* (la scrittura $y > 0$, dove $y \in \mathbb{R}^p$ qui e nel seguito, equivale a $y_i > 0, i = 1, ..., p$).

Dal teorema di Karush–Kuhn–Tucker, sotto opportune ipotesi di regolarità sulle funzioni f, g_1 e g_2, introducendo i moltiplicatori di Lagrange $\lambda \in \mathbb{R}^{neq}$ e $w \in \mathbb{R}^m$, si possono scrivere le condizioni di ottimalità per il problema (1), rappresentate dal seguente sistema di equazioni nonlineari:

$$\nabla f(x) - \lambda^t \nabla g_1(x) - w^t \nabla g_2(x) = 0 \tag{2}$$

$$-g_1(x) = 0 \tag{3}$$

$$-g_2(x) + s = 0 \tag{4}$$

$$SWe_m = 0 \tag{5}$$

$$s, w \geq 0 \tag{6}$$

dove S e W indicano le matrici diagonali con elementi uguali agli elementi di s e di w rispettivamente ed $e_m \in \mathbb{R}^m$ è il vettore con elementi unitari.

Le equazioni (5) sono dette *condizioni di complementarità*.

L'idea fondamentale dei metodi interior point è di generare una successione di punti $\{v_k\}$ ottenuta considerando una perturbazione del sistema (2)–(5) che interessi solo le condizioni di complementarità, e che può essere espresso nel modo seguente

$$H(v) = \begin{pmatrix} H_1(v) \\ SWe_m \end{pmatrix} = \rho_k \tilde{e} \qquad (7)$$
$$s, w > 0$$

dove $H_1(v) = 0$ rappresenta le equazioni (2)–(4), $\rho_k \in \mathbb{R}^+$ è il parametro di perturbazione e $\tilde{e}$ indica il vettore di $\mathbb{R}^{n+neq+2m}$ che ha tutti gli elementi nulli tranne gli ultimi m che sono uguali a 1.
Seguendo questo approccio, al variare del parametro di perturbazione ρ_k $k = 0, 1, ...$ si genera una successione di sistemi nonlineari, per i quali l'equazione di Newton diventa

$$H'(v_k)\Delta v_k + H(v_k) = \rho_k \tilde{e}. \qquad (8)$$

Lo schema generale che descrive la classe dei metodi di tipo Newton interior point con strategia di line–search è il seguente:

Algoritmo **1**

- *Dati un punto iniziale v_0 tale che $(s_0, w_0) > 0$, una funzione di merito $\Phi(v)$ e una tolleranza tol*
- *Per $k = 0, 1, ...$ finchè $\Phi(v_k) \geq tol$*
  - *Calcolare ρ_k;*
  - *Calcolare la soluzione Δv_k del sistema lineare (8);*
  - *Calcolare α_k e $v_{k+1} = v_k + \alpha_k \Delta v_k$ tali che siano verificate:*

    Ammissibilità: $(s_{k+1}, w_{k+1}) > 0$;
    Sufficiente decrescita: $\Phi(v_{k+1}) < \Phi(v_k)$.

2. – Metodi interior point come metodi di Newton inesatto

In questa tesi la teoria dei metodi interior point viene sviluppata nel contesto dei metodi di Newton inesatto [2].
Si può infatti mostrare che con un'opportuna scelta del parametro di perturbazione, il vettore Δv_k, soluzione del sistema (8), è un passo di Newton inesatto.
Se invece Δv_k è calcolato risolvendo in modo approssimato le prime $n + neq + m$ equazioni del sistema (8) si ha

$$H'(v_k)\Delta v_k + H(v_k) = \begin{pmatrix} r_k \\ \rho_k e_m \end{pmatrix}$$

dove il vettore r_k è il residuo delle prime $n + neq + m$ equazioni; anche in questo caso si possono dare condizioni su ρ_k e su $\|r_k\|$ in modo che Δv_k sia ancora un passo di Newton inesatto [3] (qui e nel seguito $\| \cdot \| = \| \cdot \|_2$).
Il vantaggio di questo approccio consiste nel fatto che si può derivare la teoria di convergenza per gli schemi del tipo 1 sfruttando i risultati di convergenza globale dei metodi di Newton inesatto, anche nel caso in cui la direzione Δv_k sia calcolata come soluzione approssimata del sistema (8).

3. – Il caso non monotono

Un contributo originale della tesi consiste nell'introduzione di una variante non monotona del metodo di Newton inesatto e di una classe di metodi di Newton interior point non monotoni.

DEFINIZIONE **1** *Sia N un intero positivo fissato e sia $\{v_{\ell(k)}\}$ l'elemento della successione $\{v_k\}$ tale che*

$$\|H(v_{\ell(k)})\| = \max_{0 \le j \le \min(N,k)} \|H(v_{k-j})\|$$

Si definisce metodo di Newton inesatto non monotono qualsiasi metodo che genera una successione $\{v_k\}$ tale che $v_{k+1} = v_k + \Delta v_k$ e

$$\|H'(v_k)\Delta v_k + H(v_k)\| \le \eta_k \|H(v_{\ell(k)}\| \tag{9}$$

$$\|H(v_k + \Delta v_k)\| \le \xi_k \|H(v_{\ell(k)})\| \tag{10}$$

dove $\eta_k \in (0,1)$ è il termine forzante e $\xi_k = (1 - \beta(1 - \eta_k))$ con $\beta \in (0,1)$.

Per la classe dei metodi di Newton inesatto non monotoni con strategia di line search è stato possibile provare il seguente risultato, analogo al teorema di convergenza dello schema monotono in [2].

TEOREMA **1** *Se v_* è un punto limite della successione $\{v_k\}$ che verifica le proprietà (9) e (10) e inoltre $H'(v_*)$ è non singolare e $\|\Delta v_k\|$ è limitata, allora $H(v_*) = 0$ e $\{v_k\}$ converge a v_*.*

Sul piano dei metodi interior point le scelte non monotone influiscono sul parametro di perturbazione, sulla tolleranza che la quantità $\|r_k\|$ deve soddisfare e sulla regola di backtracking che definisce la sufficiente decrescita della funzione di merito.
Per i metodi di Newton interior point non monotoni è stato possibile provare il teorema di convergenza globale [1] sotto le stesse ipotesi formulate in [3] per il caso monotono.

4. – Il sistema lineare

Un altro aspetto cruciale nel design di un algoritmo di tipo 1 è la scelta del risolutore per il sistema lineare (8) che deve essere risolto ad ogni passo e particolare attenzione è stata posta nei risolutori iterativi.
Applicando tecniche di eliminazione, la matrice $H'(v_k)$ può essere ridotta ad una forma a blocchi 2×2

$$\begin{pmatrix} A & B^t \\ B & 0 \end{pmatrix} \tag{11}$$

dove $A = \nabla^2 f(x_k) + \sum_{i=1}^{neq} (\lambda_k)_i \nabla^2 (g_1)_i(x_k) + \sum_{i=1}^{m} (w_k)_i \nabla^2 (g_2)_i(x_k) + C^t S^{-1} W C$, $B = -\nabla g_1(x_k)^t$, e $C = -\nabla g_2(x_k)^t$.
Una condizione sufficiente affinchè la matrice (11) sia non singolare è che B sia di rango massimo e A sia definita positiva sullo spazio nullo di B.
Sotto queste ipotesi si sono considerati vari risolutori iterativi per il sistema (8)

ridotto alla forma (11) ed in particolare si è considerato il metodo dei gradienti coniugati: come precondizionatore è stata scelta una permutazione simmetrica della matrice ottenuta da (11) approssimando il blocco A con una matrice diagonale con elementi positivi derivati dalla diagonale di A [4].
Tale precondizionatore, sotto le ipoetsi formulate, amette una fattorizzazione LDL^t: pertanto è stata implementata una routine di fattorizzazione di tipo Cholesky, che prevede anche una tecnica di regolarizzazione dinamica (scaricabile dalla pagina web *http://dm.unife.it/blkfclt/*), che è stata utilizzata per la fattorizzazione del precondizionatore.
I risultati numerici mostrano che l'Algoritmo 1 con il metodo dei gradienti coniugati precondizionati come solutore interno è particolarmente efficiente su problemi test di grandi dimensioni provenienti da problemi di controllo ottimo di tipo ellittico.
Su questo tipo di problemi, i tempi di esecuzione dell'algoritmo sono risultati inferiori a quelli del software commerciale Knitro.
I risultati di questa parte della tesi sono contenuti in un lavoro in corso di pubblicazione su *Computational Optimization and Applications*.

Riferimenti bibliografici

[1] S. Bonettini, *A nonmonotone inexact Newton method*, Optim. Meth. and Software, **20**, (2005), 475–491.

[2] S.C. Eisenstat e H.F. Walker, *Globally convergent inexact Newton methods*, SIAM J. Optimization, **4**, (1994), 393–422.

[3] C. Durazzi e V. Ruggiero, *A Newton inexact interior point method for large scale nonlinear optimization problems*, Annali Univ. Ferrara, Sez. VII, Sc. Matem., **49**, (2003), 333–357.

[4] V. Lukšan e J. Vlček, *Indefinitely preconditioned inexact Newton method for large sparse equality constrained non-linear programming problems*, Num. Lin. Alg. Appl., **5**, (1998), 219–247.

Dipartimento di Matematica Pura ed Applicata
Università di Modena e Reggio Emilia, via Campi 213/b, 41100 Modena
e-mail: bonettini.silvia@unimo.it
Dottorato in Matematica (sede amministrativa: Modena) - Ciclo XVII
Direttore di ricerca: Prof. Emanuele Galligani, Università di Modena e Reggio Emilia