# Bollettino Unione Matematica Italiana

*Sezione A – La Matematica nella Società e nella Cultura*



Cinzia Lazzari


## Nuovi algoritmi ammissibili e QP-free per la minimizzazione vincolata di funzioni $SC^1$


*Bollettino dell'Unione Matematica Italiana, Serie 8*, Vol. **6-A**—*La Matematica nella Società e nella Cultura* (2003), n.2, p. 279–282.

Unione Matematica Italiana

<http://www.bdim.eu/item?id=BUMI_2003_8_6A_2_279_0>

Bollettino dell'Unione Matematica Italiana, Unione Matematica Italiana, 2003.

Bollettino U. M. I.
La Matematica nella Società e nella Cultura
Serie VIII, Vol. VI-A, Agosto 2003, 279-282

# Nuovi algoritmi ammissibili e QP-free per la minimizzazione vincolata di funzioni $\mathrm{SC}^1$.

CINZIA LAZZARI

## 1. – Introduzione.

In questa tesi vengono sviluppati algoritmi locali e globali *ammissibili* e *QP-free* per la soluzione di problemi di ottimizzazione non lineare vincolata del tipo:

$$\min f(x) \qquad x \in \mathcal{F}, \tag{1}$$

dove $f : R^n \to R$, $\mathcal{F} := \{x \in R^n \quad : \quad g(x) \leq 0\}$ e $g : R^n \to R^m$. La funzione $f$ è supposta $SC^1$ cioè una funzione continuamente differenziabile con gradiente semismooth dove, parlando informalmente, una funzione semismooth è una funzione Lipschitziana non necessariamente continuamente differenziabile; mentre la funzione $g$, che definisce la regione ammissibile $\mathcal{F}$, è due volte continuamente differenziabile con derivate seconde Lipschitziane. Per algoritmo ammissibile si intende un algoritmo che ad ogni iterazione genera punti ammissibili, cioè appartenenti ad $\mathcal{F}$; mentre un algoritmo è QP-free se ad ogni iterazione risolve solo sottoproblemi lineari e non quadratici. L'interesse a sviluppare algoritmi ammissibili è chiarita nelle seguenti osservazioni:

– in molte applicazioni reali, la funzione obiettivo può non essere definita fuori dalla regione ammissibile,

– in molti contesti ingegneristici, si è interessati ad arrestare i processi di ottimizzazione dopo alcune iterazioni, richiedendo l'ammissibilità del punto,

– l'ammissibilità permette l'uso della funzione obiettivo come funzione di merito, evitando funzioni di merito artificiale che spesso sono complesse. L'interesse a sviluppare algoritmi per la minimizzazione vincolata di funzioni obiettivo $\mathrm{SC}^1$ è legata alla possibile applicazione dell'algoritmo alla soluzione di disequazioni variazionali a dimensini finite. Infatti, sotto opportune ipotesi, la soluzione di una disequazione variazionale può essere ottenuta risolvendo un problema di ottimizzazione non lineare vincolato in cui la funzione obiettivo è $\mathrm{SC}^1$. Algoritmi ammissibili esistenti per la soluzione del problema (1) si suddividono in metodi del primo ordine (che utilizzano solo le derivate prime) e metodi del secondo ordine (che usano anche le derivate seconde). Quest'ultimi sono i più significativi (vedi [3, 4]) poichè sotto opportune ipotesi, la stretta complementarità nella soluzione e la funzione obiettivo di classe $\mathrm{C}^2$, sono localmente veloci (convergenza locale superlineare/quadratica).

Nella tesi si sono sviluppati tre algoritmi locali ammissibili e QP-free per il problema (1) con funzione obiettivo $SC^1$; per tutti gli algoritmi proposti si è provata la convergenza locale superlineare/quadratica senza usare la stretta complementarità nella soluzione. Il primo algoritmo locale sviluppato è stato globalizzato. L'algoritmo globalmente convergente proposto migliora lo stato dell'arte nei seguenti punti:

(i) è il primo algoritmo ammissibile per il quale la convergenza superlineare è provata per funzioni obiettivo $SC^1$,

(ii) è il primo algoritmo ammissibile per il quale la convergenza superlineare è provata senza usare la stretta complementarità nella soluzione.

Inoltre, da un punto di vista pratico, gli algoritmi presentati in questa tesi hanno la caratteristica positiva di richiedere ad ogni iterazione la soluzione solo di sistemi lineari.

## 2. – Nuovi algoritmi ammissibili QP-free.

Tutti gli algoritmi locali proposti producono una sequenza $\{x^k\}$ del tipo:

$$x^{k+1} = x^k + d^k + \tilde{d}^k.$$

Le due direzioni hanno ruoli differenti: informalmente possiamo dire che $d^k$ è una direzione tipo Newton che assicura una rapidità di convergenza superlineare/quadratica ma non garantisce l'ammissibilità; mentre $\tilde{d}^k$ è una direzione di «correzione» che assicura l'ammissibilità, cioè garantisce che il nuovo punto $x^{k+1}$ appartenga alla regione ammissibile, senza alterare la rapidità di convergenza ottenuta con la prima direzione. Per il calcolo di entrambe le direzioni, $d^k$ e $\tilde{d}^k$, si utilizza una stima $I_0$ dell'insieme dei vincoli attivi nel punto ottimo $x^\star$, ovvero dei vincoli soddisfatti all'ugualianza in $x^\star$, ottenuta con la tecnica proposta da Facchinei, Fisher e Kanzow in [1]. L'algoritmo base proposto nella tesi (LFQP-free Algorithm) e le sue modifiche, Low-Cost LFQP-free Algorithm e Quasi-Newton LFQP-free Algorithm, differiscono per il calcolo della prima direzione mentre la seconda direzione $\tilde{d}^k$ è la stessa e si ottiene risolvendo il seguente problema ai minimi quadrati:

$$\begin{aligned} &\min_{\tilde{d}} \ \frac{1}{2}\|\tilde{d}\|^2 \\ &g_i(x^k + d^k) + \nabla g_i(x^k)^\top \tilde{d} = -\|d^k\|^2, \qquad i \in I_0. \end{aligned} \tag{2}$$

La prima direzione di ricerca $d^k$ per l'algoritmo base, LFQP-free Algorithm, è ottenuta risolvendo il sistema:

$$\begin{bmatrix} H^k & \nabla g_{I_0}(x^k) \\ \nabla g_{I_0}(x^k)^\top & 0 \end{bmatrix} \begin{bmatrix} d^k \\ z^k \end{bmatrix} = -\begin{bmatrix} \nabla f(x^k) \\ g_{I_0}(x^k) \end{bmatrix}. \tag{3}$$

con $H^k$ elemento di $\partial^2 L(x^k, \lambda(x^k))$, dove $\partial^2 L(x^k, \lambda(x^k))$ è l'Hessiano generalizzato della funzione Lagrangiana $L(x^k, \lambda(x^k))$, e $\lambda(x^k)$ è la funzione moltiplicatrice la cui espressione si ottiene risolvendo un sistema lineare $m \times m$. Osserviamo che $\nabla L(x^k, \lambda(x^k))$ e $\lambda(x^k)$ sono funzioni semismooth, essendo la funzione obiettivo $SC^1$; quindi per esse non si può parlare di Jacobiano ma di Jacobiano generalizzato. Nella tesi si è provato che sotto opportune ipotesi (l'indipendenza lineare dei vincoli attivi all'ottimo e la condizione sufficiente generalizzata del secondo ordine) la matrice dei coefficienti del sistema (3) è nonsingolare per ogni $H^k$ elemento di $\partial^2 L(x^k, \lambda(x^k))$, la sequenza $x^k + d^k$ è ben definita e converge superlinearmente a $x^\star$. Introducendo ulteriori ipotesi sulla funzione obiettivo si è provata la convergenza quadratica della sequenza $x^k + d^k$ al punto $x^\star$. Sotto le opportune ipotesi sopra menzionate e sfruttando l'informazione che la direzione $d^k$ è superlinearmente convergente (cioè $x^k + d^k$ converge superlinearmente), senza usare l'espressione analitica della $d^k$, si sono provate le seguenti proposizioni:

PROPOSIZIONE 1. – $\|\tilde{d}^k\| = o(\|d^k\|)$.

PROPOSIZIONE 2. – *La sequenza* $x^{k+1} = x^k + d^k + \tilde{d}^k$ *è ammissibile.*

Le proposizioni precedenti possono essere applicate a una qualunque sequenza $x^{k+1} = x^k + d^k + \tilde{d}^k$ in cui la prima direzione $d^k$ è superlinearmente convergente. Questo ci ha permesso di sviluppare due modifiche dell'algoritmo base variando solo la prima direzione per la quale, in entrambe le modifiche, è stata provata la convergenza superlineare. Sotto opportune ipotesi, si è provata la convergenza superlineare dell'intera sequenza $x^{k+1} = x^k + d^k + \tilde{d}^k$ prodotta da LFQP-free Algorithm; inoltre aumentando le richieste sulla funzione obiettivo si è provata la convergenza quadratica della sequenza $\{x^{k+1}\}$. Il carico computazionale dell'algoritmo base (LFQP-free Algorithm) ad ogni iterazione è dato da due sistemi lineari (uno per calcolare $d^k$ e l'altro per $\lambda(x^k)$) e un problema ai minimi quadrati che fornisce la $\tilde{d}^k$. Il carico computazionale viene ridotto a un sistema lineare e un problema ai minimi quadrati nel secondo algoritmo proposto nella tesi, Low-Cost LFQP-free Algorithm. In questo secondo algoritmo la prima direzione di ricerca, $d^k$, è ottenuta sostituendo nel sistema (3) la funzione moltiplicatrice $\lambda(x^k)$ con $\tilde{\lambda}^k$ definito da

$$\tilde{\lambda}_i^k = z_i^{k-1}, \qquad \text{se } i \in I_0, \qquad \tilde{\lambda}_i^k = 0, \qquad \text{se } i \notin I_0.$$

Questo modo di aggiornare il moltiplicatore $\lambda$ evita di risolvere un sistema lineare, diminuendo il carico computazionale. Nel terzo algoritmo locale proposto, Quasi-Newton LFQP-free Algorithm, la prima direzione di ricerca è ottenuta sostituendo nel sistema (3) le matrici $H^k$ con le matrici $B^k$ che sono matrici $n \times n$ nonsingolari che soddisfano opportune assunzioni. L'uso delle matrici $B^k$ permette di non calcolare l'Hessiano generalizzato del Lagrangiano. Nella tesi i risultati di convergenza provati per Low-Cost LFQP-free Algorithm e Quasi-Newton LFQP-free Algorithm sono analoghi a quelli ottenuti per l'algoritmo base LFQP-free Algorithm.

I risultati relativi agli algoritmi locali sono contenuti in [2].

L'algoritmo locale base è stato globalizzato utilizzando una tecnica ibrida che combina il LFQP-free Algorithm con un algoritmo ammissibile globalmente convergente appartenente ad un'ampia classe di algoritmi esistenti in letteratura. La combinazione dei due algoritmi è ottenuta utilizzando opportuni test che ad ogni iterazione dicono quale dei due algoritmi utilizzare. Nel caso in cui si utilizzi l'algoritmo locale altri test sono stati introdotti per valutare efficientemente la «bontà» della direzione locale in termini di convergenza globale. Il ruolo svolto dai test è quello di preservare sia le proprietà di convergenza globale fornite dall'algoritmo globalmente convergente che le proprietà di convergenza locale «veloce» ottenute con il LFQP-free Algorithm. Il teorema finale di convergenza provato per questo nuovo algoritmo è il seguente:

TEOREMA 1. – *Supponendo che* $\mathcal{F}$ *sia un insieme convesso tale che* $\mathcal{F} \cap \{x \mid f(x) \leq f(x^0)\}$ *sia compatto, dove* $x^0$ *è un punto iniziale appartenente ad* $\mathcal{F}$, *allora la sequenza* $\{x^k\}$ *ammette almeno un punto di accumuazione, ogni punto di accumulazione di* $\{x^k\}$ *è un punto stazionario del problema* (1); *inoltre se* $x^\star$ *è un punto di accumulazione che soddisfa opportune ipotesi e* $f$ *è una funzione* $C^1$ *con* $\nabla f$ *semismooth di ordine* $p$ *in* $x^\star$ $(0.5 < p \leq 1)$, *allora l'intera sequenza* $\{x^k\}$ *converge a* $x^\star$ *e la velocità di convergenza è* $1 + p$.

## BIBLIOGRAFIA


[1] FACCHINEI F., FISHER A. e KANZOW C., *On the accurate identification of active constraints*, Siam Journal Optimization, **9** (1998), 14-32.

[2] FACCHINEI F. e LAZZARI C., *Local Feasible QP-free algorithms for the constrained minimization of* $SC^1$ *functions*, Technical Report, **TR 11-00** (2000), 1-20.

[3] LAWRENCE C.T. e TITS A.L., *A computationally efficient feasible sequential quadratic programming algorithm*, Siam Journal Optimization, **11(4)** (2001), 1092-1118.

[4] QI HOU-DUO e QI L., *A new QP-free, globally convergent, locally superlinearly convergent algorithm for inequality constrained optimization*, Siam Journal Optimization, **11(1)** (2000), 113-132.



Dipartimento di Informatica e Sistemistica «Antonio Ruberti»,
Università degli Studi di Roma «La Sapienza»
e-mail: lazzari@dis.uniroma1.it
Dottorato in Ricerca Operativa (sede amministrativa: Facoltà di Scienze Statistiche,
Università «La Sapienza» Roma) - Ciclo XIV
Direttore di ricerca: Prof. Francisco Facchinei, Università di Roma «La Sapienza»