
BOLLETTINO UNIONE MATEMATICA ITALIANA

Sezione A – La Matematica nella Società e nella Cultura

MARCO LOCATELLI

Algoritmi di ottimizzazione globale

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. 189–192.

Unione Matematica Italiana

[<http://www.bdim.eu/item?id=BUMI_1998_8_1A_1S_189_0>](http://www.bdim.eu/item?id=BUMI_1998_8_1A_1S_189_0)

L'utilizzo e la stampa di questo documento digitale è consentito liberamente per motivi di ricerca e studio. Non è consentito l'utilizzo dello stesso per motivi commerciali. Tutte le copie di questo documento devono riportare questo avvertimento.

*Articolo digitalizzato nel quadro del programma
bdim (Biblioteca Digitale Italiana di Matematica)
SIMAI & UMI*

<http://www.bdim.eu/>

Algoritmi di ottimizzazione globale.

MARCO LOCATELLI

1. - Introduzione.

La tesi tratta il problema dell'ottimizzazione globale, cioè

quote dato un insieme $X \subseteq \mathbb{R}^n$, detto regione ammissibile, ed una funzione $f: X \rightarrow \mathbb{R}$, chiamata funzione obiettivo, trovare, se esiste, un punto $x^ \in X$ tale che*

$$\forall x \in X \quad f^* = f(x^*) \leq f(x).$$

Tra i riferimenti bibliografici sull'argomento si cita in particolare [6], dove viene fatta una dettagliata presentazione di diverse problematiche associate all'ottimizzazione globale e di diverse classi di problemi di ottimizzazione globale.

Il punto x^* viene detto minimo globale di f su X . L'esistenza di tale punto è garantita sotto condizioni piuttosto generali, ad esempio se f è semicontinua inferiormente e X è compatto. Un punto $y^* \in X$ si definisce minimo locale della funzione obiettivo f su X se esiste un $\delta > 0$ tale che

$$\forall x \in X \cap S(y^*, \delta) \quad f(y^*) \leq f(x),$$

dove $S(y^*, \delta)$ indica la sfera con centro y^* e raggio δ . I minimi globali sono casi particolari di minimi locali di una funzione. L'ottimizzazione globale si occupa di problemi con funzioni multimodali, in cui esistono minimi locali che non sono globali (il che esclude, per esempio, i problemi di programmazione convessa, in cui la f è una funzione convessa e la regione X è convessa).

I problemi di ottimizzazione globale risultano essere piuttosto complessi da risolvere: in [10] si è dimostrato che persino nel caso particolare in cui f sia quadratica e concava e X sia l'ipercubo unitario, il problema è NP-difficile. Esistono, comunque, diversi gradi di difficoltà, strettamente legati alla struttura del problema. La struttura è determinata dalle informazioni a priori circa la funzione obiettivo e la regione ammissibile. Nei problemi con una struttura forte, quali i problemi di ottimizzazione concava o di ottimizzazione di funzioni Lipschitz con costante di Lipschitz nota, è possibile definire algoritmi in grado di fornire dei limiti superiori sulla distanza tra la miglior soluzione individuata da essi ed il valore f^* dell'ottimo globale. Nei problemi con struttura più debole, quali quelli di ottimizzazione di una funzione continua su un insieme compatto, la povertà delle informazioni a priori impedisce di individuare algoritmi in grado di misurare la distanza della miglior soluzione individuata dall'ottimo globale.

Legata alla differenziazione dei problemi sulla base della loro struttura è la classificazione degli algoritmi di risoluzione in due grandi classi: algoritmi deter-

ministici ed algoritmi stocastici. Un algoritmo si definisce deterministico se, una volta fissati alcuni valori iniziali, quali possono essere un punto di partenza nella regione ammissibile od alcuni parametri da cui dipende l'algoritmo, diverse esecuzioni dell'algoritmo su un dato problema forniscono sempre gli stessi risultati ed in particolare risulta fissata la sequenza dei punti in cui la funzione viene osservata. Negli algoritmi stocastici si ha, invece, che ogni esecuzione, a parità di condizioni iniziali, può fornire risultati diversi ed in particolare ad ogni esecuzione può cambiare la sequenza dei punti in cui viene osservata la funzione.

Normalmente si ha che in presenza di problemi con una struttura forte vengono utilizzati algoritmi di tipo deterministico. La possibilità di combinare le informazioni raccolte attraverso le osservazioni fatte con le informazioni a priori su funzione obiettivo e regione ammissibile, consente di escludere parti della regione ammissibile, in cui si può garantire che non esistono punti migliori di quelli già osservati, e consente di elaborare regole deterministiche in grado di fornire buoni punti per le osservazioni successive. Nel momento in cui si indebolisce la struttura del problema, appare più difficile definire criteri deterministici in grado di individuare buoni punti per le osservazioni successive ed è opportuno in questi casi introdurre una componente aleatoria che rifletta l'incertezza derivante dalle poche informazioni sul problema di cui si dispone.

Nella tesi si sono analizzate diverse classi di problemi di ottimizzazione globale e diversi algoritmi per la loro risoluzione, sia di tipo stocastico che di tipo deterministico. In alcuni casi sono stati ripresi algoritmi già noti e si sono presentati nuovi risultati teorici su di essi. In altri sono stati proposti ed analizzati nuovi algoritmi. Nella sezione successiva si presenteranno in sintesi i principali risultati originali ottenuti nella tesi.

2. – Contributi originali della tesi

2.1. SIMULATED ANNEALING. – Dopo un breve capitolo dove sono stati presentati alcuni risultati teorici sulla convergenza in probabilità sia del miglior punto osservato che dell'intera sequenza di iterati per algoritmi di tipo stocastico, si sono applicati tali risultati di convergenza a particolari algoritmi stocastici, gli algoritmi di tipo Simulated Annealing, che si ispirano alla simulazione di un processo fisico, chiamato annealing, mediante il metodo Metropolis Monte Carlo (si veda [9]) ed introdotti nell'ambito dell'ottimizzazione combinatoria in [2] e [7]. In ambito combinatorio esistono condizioni necessarie e sufficienti per garantire la convergenza di algoritmi Simulated Annealing (si veda [4]). In ambito continuo esistono risultati di convergenza non completamente soddisfacenti per diverse ragioni. Alcuni risultati (si veda, ad esempio, [1]) sono ottenuti sotto ipotesi molto generali, ma non del tutto soddisfacenti dal punto di vista teorico. In particolare, in ogni algoritmo Simulated Annealing vengono generate due sequenze di punti, quella dei punti «accettati» e quella dei punti «generati», più importante della prima essendo quella costituita da tutti i punti in cui si valuta la funzione obietti-

vo. I risultati citati sono in grado di dimostrare la convergenza in probabilità all'ottimo globale della prima sequenza ma non della seconda. Altri risultati teorici sono molto più soddisfacenti da un punto di vista teorico (si veda, ad esempio, [3]) ma, d'altro canto, sono ottenuti sotto ipotesi piuttosto restrittive. Nella tesi sono stati ottenuti risultati di convergenza che, sotto ipotesi piuttosto generali, sono in grado di garantire la convergenza di entrambe le sequenze citate sopra, consentendo quindi di superare i limiti, sottolineati in precedenza, dei risultati presenti nella letteratura.

2.2. ALGORITMI MULTISTART. – Gli algoritmi di tipo multistart tentano di individuare l'ottimo globale di una funzione attraverso l'esecuzione multipla di ricerche locali. Ad una fase di campionamento uniforme sulla regione ammissibile segue una fase di decisione riguardante la scelta dei punti campionati dai quali far partire ricerche locali. Tra gli algoritmi di questo tipo è molto noto l'algoritmo denominato Multilevel Single Linkage (MLSL), introdotto in [11] e [12]. Nella tesi viene introdotto un nuovo algoritmo, Non Monotonic Multilevel Single Linkage, che, al costo di un maggior sforzo computazionale per iterazione, consente di rilassare alcune ipotesi sotto cui sono state dimostrate le proprietà teoriche di MLSL (per esempio la richiesta $f \in \mathcal{C}^2$ è sostituita da f lipschitziana). Successivamente è stata introdotta una nuova classe di algoritmi multistart, la classe Random Linkage, studiando le proprietà teoriche, assimilabili a quelle di MLSL, di tale classe ed in particolare di una sua sottoclasse, la classe Simple Linkage (SL). In SL il meccanismo di selezione dei punti da cui far partire ricerche locali è meno costoso rispetto a MLSL e tuttavia un dettagliato confronto teorico e pratico mostra un forte legame tra le scelte effettuate dai due algoritmi.

2.3. ALGORITMI BAYESIANI. – Gli algoritmi bayesiani vengono utilizzati per l'ottimizzazione globale di funzioni continue su intervalli dell'asse reale. L'idea di base di tali algoritmi è quella di modellare la funzione obiettivo attraverso un processo stocastico, il processo di Wiener (si veda [8]). Viene presentato un algoritmo già noto nella letteratura (si veda [13]) ma completato con una nuova regola di arresto. Inoltre, essendo stata osservata la forte dipendenza dell'algoritmo dalla scelta di un parametro, viene proposto ed analizzato sia da un punto di vista teorico che pratico, un nuovo algoritmo in cui il parametro non viene fissato a priori, ma viene definita su di esso una distribuzione a priori, aggiornata dopo ogni osservazione. Nonostante anche il nuovo algoritmo dipenda da un parametro che compare nella distribuzione a priori, esso ha, rispetto al precedente, una maggiore capacità di autocorreggersi sulla base delle informazioni raccolte durante la propria esecuzione e i test computazionali mostrano che è molto meno sensibile alla scelta del parametro iniziale.

2.4. ALGORITMI CONICI PER L'OTTIMIZZAZIONE CONCAVA. – Gli algoritmi conici per l'ottimizzazione concava su politopi sono algoritmi di tipo branch-and-bound.

Ogni nodo dell'albero di branch-and-bound corrisponde ad un cono poliedrale e la risoluzione di un problema di programmazione lineare avente come regione ammissibile l'intersezione tra il politopo ed il cono, permette di stabilire se si possa tagliare o meno il nodo. La creazione di nuovi nodi da un nodo dato avviene attraverso un meccanismo di suddivisione da definire. Tra i meccanismi di suddivisione si citano la bisezione e le ω -suddivisioni. Quest'ultime utilizzano come punto di suddivisione la soluzione ottima del problema di programmazione lineare citato sopra. Nonostante esista una vasta letteratura sull'argomento (si veda [5] ed i riferimenti in esso), mancava un risultato di convergenza all'ottimo per l'algoritmo conico basato esclusivamente sulle ω -suddivisioni. Nella tesi si presenta una soluzione a tale problema, dimostrandone per la prima volta la convergenza e la finitezza nel caso di ε -approssimazione.

BIBLIOGRAFIA

- [1] BELISLE C.J.P., *Convergence Theorems for a Class of Simulated Annealing Algorithms on $\mathbb{R}^d$* , J. Appl. Prob., **29** (1992), 885-892.
- [2] ČERNÝ V., *Thermodynamical Approach to the Travelling Salesman Problem: An Efficient Simulation Algorithm*, J. Optim. Theory Appl., **45** (1985), 41-51.
- [3] GELFAND S.B. e MITTER S.K., *Metropolis-type annealing algorithms for global optimization in $\mathbb{R}^d$* , SIAM J. of Control and Optimization, **31**, No. 1 (1993), 111-131.
- [4] HAJEK B., *Cooling schedules for optimal annealing*, Mathematics of Operations Research, **13** (1988), 311-329.
- [5] HORST R. e TUY H., *Global optimization: deterministic approaches*, (second edition), Springer-Verlag (1992).
- [6] HORST R. e PARDALOS P. (editori), *Handbook of global optimization*, Kluwer Academic Publishers (1995).
- [7] KIRKPATRICK S., GELATT C.D. e VECCHI M.P., *Optimization by Simulated Annealing*, Science, **220** (1983), 671-680.
- [8] KUSHNER H., *A versatile stochastic model of a function of unknown and time varying form*, Journal of Math. Anal. Appl., **5** (1962), 150-167.
- [9] METROPOLIS N., ROSENBLUTH A.W., ROSENBLUTH M.N. e TELLER A.H., *Equation of State Calculations by Fast Computer Machines*, J. Chem. Phys., **21** (1953), 1087.
- [10] PARDALOS P.M. e SCHNITGER G., *Checking local optimality in constrained quadratic programming is NP-hard*, Operations Research Letters, **7** (1988), 33-35.
- [11] RINNOOY KAN A.H.G. e TIMMER G., *Stochastic global optimization methods. Part i: clustering methods*, Mathematical Programming, **39** (1987), 27-56.
- [12] RINNOOY KAN A.H.G. e TIMMER G., *Stochastic global optimization methods. Part ii: multi level methods*, Mathematical Programming, **39** (1987), 57-78.
- [13] ZILINSKAS A., *One-step Bayesian method of the search for extremum of an one-dimensional function*, Cybernetics, **1** (1975), 139-144.

Dip. di Sistemi ed Informatica, Università di Firenze
 Dottorato in Matematica Computazionale e Ricerca Operativa
 (sede amministrativa: Milano) - Ciclo IX
 Direttore di ricerca: Prof. Fabio Schoen