

Metodi di discesa

Calcolo Numerico CdS Informatica

November 9, 2022

Generalità

- Tutti gli algoritmi per l'ottimizzazione non vincolata richiedono che l'utente fornisca un punto iniziale x_0 . In generale, non esistono criteri generali per effettuare una buona scelta di x_0 e quindi si è interessati a definire algoritmi le cui proprietà di convergenza siano indipendenti dalla scelta del punto iniziale.
- A partire da x_0 , gli algoritmi generano una successione di iterati $\{x_k\}_{k=0}^{\infty}$ che termina o quando non è possibile fare alcun progresso verso la soluzione o quando la soluzione è stata approssimata con sufficiente accuratezza.
- Gli algoritmi utilizzano informazioni sulla funzione obiettivo f in x_k e, a volte, informazioni dagli iterati precedenti per determinare un nuovo punto x_{k+1} con un valore **più piccolo** della funzione obiettivo.
- Esistono due strategie fondamentali per muoversi dall'iterato corrente x_k verso il nuovo iterato x_{k+1} : **line search** e **trust region**.

Convergenza a punti stazionari

- Gli algoritmi di ottimizzazione garantiscono, in generale, la determinazione di punti che soddisfano condizioni necessarie di ottimo.
- Tutti i metodi considerati consentono di determinare **punti stazionari di f** .
- Nel caso non convesso, la determinazione di punti stazionari non fornisce una soluzione globale e non è neanche possibile, in generale, garantire che sia raggiunto un punto di minimo locale.

Criteri di arresto e fallimenti

- Negli algoritmi di ottimizzazione esistono diversi criteri per far terminare le iterazioni e che possono indicare il raggiungimento di una soluzione oppure il fallimento dell'algoritmo.
- Negli algoritmi di ottimizzazione non vincolata per criterio di arresto si intende il criterio che dovrebbe indicare il raggiungimento con successo di un punto stazionario con la tolleranza specificata dall'utente.
- Dal punto di vista teorico, il criterio di arresto di un algoritmo che genera la successione x_k dovrebbe essere la condizione

$$\|\nabla f(x_k)\| \leq \epsilon, \quad \epsilon > 0 \quad (1)$$

Altre condizioni di arresto:

- superamento di un massimo numero di iterazioni;
- fallimenti interni alle procedure usate dall'algoritmo per il calcolo di x_{k+1} (as esempio nella ricerca in linea);
- ...

Metodi di discesa

- Si consideri il problema della minimizzazione non vincolata di una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}$ differenziabile con continuità.
- I metodi con ricerca in linea sono metodi iterativi che, a partire da un iterato iniziale $x_0 \in \mathbb{R}^n$, generano una successione di vettori

$$x_0, x_1, x_2, \dots$$

definiti dall'iterazione

$$x_{k+1} = x_k + \alpha_k p_k \quad (2)$$

dove il vettore p_k è una **direzione di ricerca** e lo scalare α_k è un parametro positivo chiamato **lunghezza del passo** (step-length) che indica la distanza di cui ci si deve muovere lungo la direzione p_k .

- Il vettore p_k ed il parametro α_k sono scelti in modo da garantire la decrescita di $f(x)$ ad ogni iterazione:

$$f(x_{k+1}) < f(x_k), \quad k = 0, 1, \dots$$

Direzione di discesa

Definizione

Il vettore p è una direzione di discesa di f in x se esiste un $\bar{\alpha} > 0$ tale che

$$f(x + \alpha p) < f(x) \text{ per ogni } \alpha \in (0, \bar{\alpha}]$$

Lemma

Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}$ continuamente differenziabile nell'intorno di un punto $x \in \mathbb{R}^n$ e sia $p \in \mathbb{R}^n$ un vettore non nullo. Allora se risulta

$$p^T \nabla f(x) < 0 \tag{3}$$

la direzione p è una direzione di discesa per f in x .

Inversamente, se f è continuamente differenziabile e convessa in un intorno $B(x, \rho)$ di x e se p è una direzione di discesa in x , deve essere necessariamente soddisfatta la (3).

Esempio 1

$$f(x) = x_1^3 + x_2^2 - 3x_1$$

- Consideriamo il punto $\bar{x} = (-1, 0)^T$ e la direzione $d = (-1, 0)^T$
- Vale $\nabla f(x) = \begin{pmatrix} 3x_1^2 - 3 \\ 2x_2 \end{pmatrix}$ e $\nabla f(\bar{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
- La direzione d è di discesa. Infatti risulta $\bar{x} + td = (-1 - t, 0)^T$ e la funzione vale

$$f(\bar{x} + td) = (-1 - t)^3 - 3(-1 - t) = -t^3 - 3t^2 + 2$$

La condizione $-t^3 - 3t^2 + 2 < 2 = f(\bar{x})$ è verificata per ogni $t > 0$.

Esempio 2

- Consideriamo il punto $\tilde{x} = (1, 0)^T$ e la direzione $d = (-1, 0)^T$
- Vale $\nabla f(x) = \begin{pmatrix} 3x_1^2 - 3 \\ 2x_2 \end{pmatrix}$ e $\nabla f(\tilde{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
- La direzione d non è di discesa. Infatti risulta $\tilde{x} + td = (1 - t, 0)^T$ e la funzione vale

$$f(\tilde{x} + td) = (1 - t)^3 - 3(1 - t) = -t^3 + 3t^2 - 2$$

La condizione $-t^3 + 3t^2 - 2 < -2 = f(\bar{x})$ è verificata solo per ogni $t > 3$.

Interpretazione geometrica

Ricordando che il concetto di angolo tra due vettori $x \neq 0$ e $y \neq 0$ in \mathbb{R}^n si può introdurre attraverso la definizione di coseno, ponendo

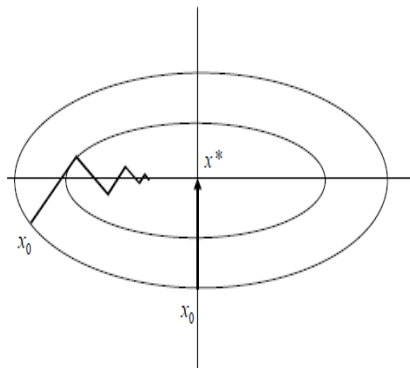
$$\cos \theta = \frac{x^T y}{\|x\| \|y\|}$$

si può dire che l'angolo tra p e $\nabla f(x)$ è

- *ottuso* se $p^T \nabla f(x) < 0$
- *acuto* se $p^T \nabla f(x) > 0$

I vettori p e $\nabla f(x)$ sono *ortogonali* se $p^T \nabla f(x) = 0$.

La retta $x = x_k + \alpha p_k$ deve formare un angolo ottuso con la direzione del gradiente (derivata direzionale negativa):



- Se $p^T \nabla f(x) < 0$ la direzione p è una direzione di discesa per f in x ;
- se $p^T \nabla f(x) > 0$ la direzione p è una direzione di salita per f in x ;
- se $p^T \nabla f(x) = 0$ non è possibile stabilire, in assenza di altre informazioni, se p sia o meno una direzione di discesa.

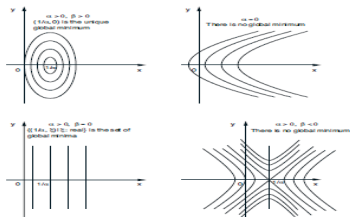


Illustration of the isocost surfaces of the quadratic cost function $f: \mathbb{R}^2 \mapsto \mathbb{R}$ given by

$$f(x, y) = \frac{1}{2}(\alpha x^2 + \beta y^2) - x$$

for various values of α and β .

- La direzione dell'antigradiente $p = -\nabla f(x)$ è sempre una direzione di discesa
- Nel caso generale (f non convessa), la condizione di discesa enunciata è solo sufficiente, in quanto possono esistere direzioni di discesa tali che $\nabla f(x)^T p = 0$.

- I metodi con ricerca in linea utilizzano direzioni p_k di discesa in quanto, se p_k è una direzione di discesa, allora è possibile garantire una decrescita di f nella direzione di p_k purchè il parametro α_k sia sufficientemente piccolo.
- A seconda della scelta della direzione p_k si hanno diversi metodi con ricerca in linea. La condizione $p_k^T \nabla f(x_k) < 0$ deve essere garantita dall'algoritmo per calcolare la direzione di ricerca.

Schema algoritmico

Un metodo con ricerca in linea è dunque un metodo iterativo che ha il seguente schema generale

Metodo con ricerca in linea

Dato x_0 , per $k = 0, 1, 2, \dots$

- i) determinare una direzione di discesa p_k ;
- ii) determinare una lunghezza del passo α_k ;
- iii) porre $x_{k+1} = x_k + \alpha_k p_k$ e $k + 1 = k$;

Scelta della lunghezza del passo

- Esistono diverse tecniche per determinare la lunghezza del passo in modo tale da garantire la convergenza del metodo verso punti stazionari di f . Esse sono, in generale, chiamate tecniche di *ricerca in linea* (line search) o di *ricerca unidimensionale* perchè la ricerca di un nuovo iterato x_{k+1} è fatta lungo la linea $y(\alpha) = x_k + \alpha p_k$.
- Osserviamo innanzitutto che la scelta di α_k tale che

$$f(x_{k+1}) < f(x_k)$$

non garantisce la convergenza del metodo.

- Gli algoritmi di ricerca unidimensionale hanno per obiettivo la determinazione del passo α_k lungo la direzione assegnata p_k e costituiscono usualmente la parte più complessa di un codice di calcolo per l'ottimizzazione non vincolata.
- In termini molto generali, si può dire che la ricerca unidimensionale è finalizzata, essenzialmente, ad assicurare la convergenza globale dell'algoritmo, mentre la rapidità di convergenza dipende in prevalenza dalla direzione di ricerca.
- Tuttavia, un valore di α_k poco appropriato può distruggere le proprietà di rapidità di convergenza associate alla scelta di p_k .
- Una tecnica di ricerca unidimensionale dovrebbe quindi anche assicurare che, almeno in intorno della soluzione, venga il più possibile accettato il passo che garantisce idealmente una buona rapidità di convergenza.
- Realizzare un compromesso tra le esigenze di convergenza globale e quelle di rapidità di convergenza può risultare difficile.

Ricerca in linea

Il passo α_k viene determinato con riferimento alla funzione di una sola variabile reale $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ definita (per k fissato) da:

$$\phi(\alpha) = f(x_k + \alpha p_k)$$

che caratterizza l'andamento di f lungo la direzione di ricerca p_k

Ricerca in linea esatta

- La scelta ideale è scegliere α_k come il punto di minimo globale della funzione

$$\phi(\alpha) = f(x_k + \alpha p), \quad \alpha \geq 0 \quad (4)$$

- La scelta di α_k impone che p_k e $\nabla f(x_{k+1})$ siano ortogonali, invero:

$$0 = \left. \frac{d}{d\alpha} \phi(\alpha) \right|_{\alpha=\alpha_k} = (\nabla f(x_k + \alpha_k p_k))^T p_k = (\nabla f(x_{k+1}))^T p_k$$

Ricerca in linea inesatta

In termini molto generali, i metodi di ricerca inesatta unidimensionale per cui è possibile stabilire risultati di convergenza si basano sull'individuazione di un *intervallo di valori accettabili* per α_k , definiti in modo tale da assicurare

- una sufficiente riduzione di f rispetto al valore corrente $f(x_k)$;
- un sufficiente spostamento rispetto al punto corrente x_k .

Condizioni di Wolfe

Le condizioni di Wolfe sono date da

$$f(x_k + \alpha p_k) \leq f(x_k) + c_1 \alpha \nabla f(x_k)^T p_k \quad (5)$$

$$\nabla f(x_k + \alpha_k p_k)^T p_k \geq c_2 \nabla f(x_k)^T p_k \quad (6)$$

con $0 < c_1 < c_2 < 1$.

- La prima disuguaglianza (5), detta **condizione di Armijo** (Armijo condition o sufficient decrease condition), assicura una *decrescita sufficiente* di f nella direzione p .
- La seconda disuguaglianza (6), detta **condizione della curvatura** (curvature condition), impedisce che α_k divenga troppo piccolo.

Condizione di Armijo e Backtracking

- La condizione di Armijo da sola non è sufficiente a garantire che la lunghezza del passo non sia troppo piccola. Tuttavia, se la lunghezza del passo è scelta con una **tecnica di backtracking**, allora è sufficiente usare la condizione di Armijo per garantire la convergenza.
- Nell'algoritmo di backtracking, inizialmente si pone $\alpha = \bar{\alpha}$ (di solito $\bar{\alpha} = 1$). Quindi si riduce questo valore finchè non si sia trovato un valore α_k per cui sia soddisfatta la condizione di Armijo (5).

Algoritmo di backtracing

- Si esamina se l'attuale valore di α non soddisfa la condizione di Armijo, cioè, se

$$f(x_k + \alpha p_k) > f(x_k) + c_1 \alpha \nabla f(x_k)^T p_k$$

- Se ciò avviene, si riduce α di un fattore $\rho < 1$, cioè si pone $\alpha = \rho \alpha$ e si esamina la condizione di Armijo con questo nuovo valore di α .
- Quando, finalmente, si trova un valore di α per cui la disuguaglianza di Armijo è soddisfatta, questo valore è il parametro α_k che si cercava.
- Nella pratica spesso si sceglie $\rho = \frac{1}{2}$ e quindi α_k è il primo numero della successione $\{1, \frac{1}{2}, \frac{1}{4}, \dots, \frac{1}{2^i}, \dots\}$ che verifica la condizione di Armijo.

Algoritmo di backtracing

Scegliere $\bar{\alpha} > 0$, $\rho, c_1 \in (0; 1)$; porre $\alpha = \bar{\alpha}$ e $j = 0$;
while $f(x_k + \alpha p_k) > f(x_k) + c_1 \alpha \nabla f(x_k)^T p_k$
 $\alpha = \rho \alpha$ e $j = j + 1$
end(repeat)
Porre $\alpha_k = \alpha$.

Metodi di discesa

- I metodi di discesa sono definiti dall'iterazione generale

$$x_{k+1} = x_k + \alpha_k p_k.$$

- Essi si differenziano per la scelta della direzione di discesa e della lunghezza del passo.

Il metodo di discesa ripida

- Il **metodo di discesa ripida** (steepest descent method) è uno dei metodi di ottimizzazione non lineare più semplici. Esso richiede, ad ogni iterazione, il calcolo delle derivate prime della funzione obiettivo e, solitamente, converge assai lentamente cosicché anche se il costo per iterazione è basso, il costo totale per risolvere un problema di ottimizzazione è solitamente alto.

- Il metodo di discesa ripida, ad ogni iterazione k , usa come direzione di discesa p^k la direzione dell'antigradiente:

$$p^k = -\nabla f(x_k). \quad (7)$$

Dunque, il costo computazionale per determinare p^k è quello per determinare il gradiente.

- La direzione (7) è una direzione di discesa di f in x_k se $\nabla f(x_k) \neq 0$.
- La lunghezza del passo α_k è determinata con una ricerca in linea esatta; cioè α_k è determinato come quel valore per cui $\phi(\alpha) = f(x_k - \alpha \nabla f(x_k))$ assume il valore minimo:

$$\alpha_k = \arg \min_{\alpha \geq 0} f(x_k - \alpha \nabla f(x_k)) \quad (8)$$

- Il metodo deve il suo nome all'interpretazione geometrica del vettore gradiente di f perchè la direzione p^k di più ripida discesa in x^k è $-\nabla f(x_k)$.
- Infatti, $(\nabla f(x^k))^T p^k$ ha il minimo negativo per il versore $-\nabla f(x_k)/\|\nabla f(x_k)\|^2$ tra tutti i possibili versori.
- Dunque:
 $-\nabla f(x_k)$ punta verso la direzione lungo cui la funzione $f(x)$ decresce più rapidamente.

Convergenza globale

- La convergenza globale è assicurata dal Teorema di Zoutendijk.
- Inoltre, si dimostra che la velocità di convergenza è *lineare*.

Il caso delle forme quadratiche convesse

- Sia Q una matrice $n \times n$ simmetrica e $f(x)$ la forma quadratica

$$f(x) = \frac{1}{2}x^T Qx - x^T b.$$

- Vale

$$\nabla f(x) = Qx - b, \quad \nabla^2 f(x) = Q$$

- Condizioni necessarie:

$$Qx^* - b = 0, \quad Q \geq 0 \quad (Q \text{ semidefinita positiva})$$

- $Q \geq 0 \Rightarrow f$ convessa \Rightarrow le condizioni necessarie sono anche sufficienti e i punti di minimo locale sono anche punti di minimo globale
- Conclusioni:
 - se Q non è $\geq 0 \Rightarrow f$ non ha punti di minimo locale
 - se $Q > 0 \Rightarrow x^* = Q^{-1}b$ è l'unico punto di minimo locale
 - se $Q \geq 0$ ma non invertibile \Rightarrow o f non ha punti di minimo locale o ne esistono infiniti

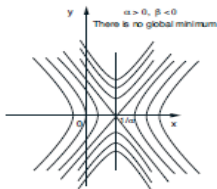
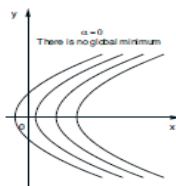
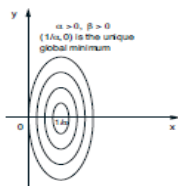


Illustration of the isocost surfaces of the quadratic cost function $f : \mathbb{R}^2 \mapsto \mathbb{R}$ given by

$$f(x, y) = \frac{1}{2}(\alpha x^2 + \beta y^2) - x$$

for various values of α and β .

- Sia Q definita positiva.
- Il punto di minimo di $f(x)$, ovvero la soluzione del sistema lineare $Qx = b$, può essere determinato con il metodo di discesa ripida.
- Il parametro α_k è ottenuto dalla equazione

$$0 = \left. \frac{d}{d\alpha} \phi(\alpha) \right|_{\alpha=\alpha_k}$$

$$\begin{aligned} 0 = \left. \frac{d}{d\alpha} \phi(\alpha) \right|_{\alpha=\alpha_k} &= -(\nabla f(x_k + \alpha_k \nabla f(x_k)))^T \nabla f(x_k) \\ &= -(\nabla f(x_{k+1}))^T \nabla f(x_k) = -(Qx_{k+1} - b)^T \nabla f(x_k) \\ &= -(Qx_k - b - \alpha_k Q \nabla f(x_k))^T \nabla f(x_k) \end{aligned}$$

da cui

$$\alpha_k = \frac{(\nabla f(x_k))^T \nabla f(x_k)}{(\nabla f(x_k))^T Q \nabla f(x_k)} > 0.$$

- Nel seguito, si indicherà con g_k il gradiente $\nabla f(x_k)$.

Algoritmo

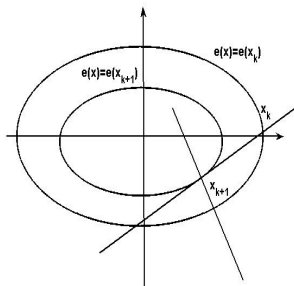
Metodo di discesa ripida per le forme quadratiche

Dato x_0 , per $k = 0, 1, 2, \dots$

- i) $g_k = Qx_k - b$;
- ii) $\alpha_k = \frac{g_k^T g_k}{g_k^T Q g_k}$
- iii) $x_{k+1} = x_k - \alpha_k g_k$ e $k = k + 1$;

Interpretazione geometrica nel caso $n = 2$

Nel caso $n = 2$ si ha la seguente interpretazione geometrica. L'equazione $e(x) = (x - x^*)^T Q(x - x^*) = \text{costante} > 0$ è l'equazione di una ellisse. Al variare di x_k si ottiene una famiglia di ellissi $e(x) = e(x_k)$ concentriche con centro x^* , punto di minimo di $e(x)$. Il vettore $-g_k$ è tangente all'ellisse $e(x) = e(x_{k+1})$ nel punto x_{k+1} .



Teorema

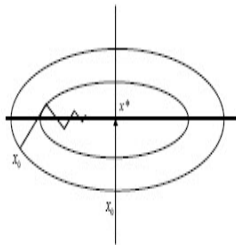
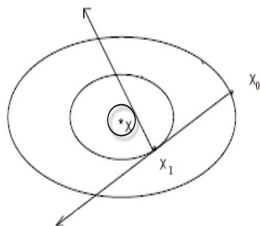
Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione differenziabile due volte con continuità. Si supponga che la successione generata dal metodo di discesa ripida converga a un punto x^* tale che la matrice Hessiana $\nabla^2(x^*)$ sia definita positiva. Allora:

$$f(x_{k+1}) - f(x^*) \leq \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} \right)^2 [f(x_k) - f(x^*)].$$

dove λ_n e λ_1 sono rispettivamente l'autovalore massimo e minimo della matrice Hessiana.

- La velocità di convergenza del metodo di discesa ripida nella norma energia è almeno lineare e dipende dalle eccentricità delle ellissi che rappresentano le curve di livello $f(x) = \text{costante}$. L'eccentricità è tanto maggiore quanto più la matrice Q è mal condizionata.
- La rapidità di convergenza del metodo del metodo di discesa ripida diminuisce all'aumentare del mal condizionamento della matrice Q .
- Le considerazioni svolte nel caso quadratico sono ovviamente indicative anche del comportamento locale del metodo nella minimizzazione di funzioni non quadratiche.
- Il metodo del metodo di discesa ripida (nelle realizzazioni usuali) si può considerare, complessivamente, un metodo di minimizzazione non vincolata scarsamente efficiente e ciò è confermato, del resto, dall'esperienza di calcolo.

Il metodo lavora bene nel caso della figura a sinistra per il quale $K(Q) \sim 1$ mentre nel caso della figura a destra, per il quale $K(Q) \approx 1$ il metodo converge molto lentamente. Le figure mostrano il tipico comportamento a zig-zag del metodo di discesa ripida, dovuto alla condizione di ortogonalità.



Metodo del gradiente

- La direzione di ricerca è

$$p_k = -\nabla f(x_k) \quad \text{per ogni } k$$

La direzione dell'antigradiente $-\nabla f(x_k)$ è una direzione di discesa di f in x_k se $\nabla f(x_k) \neq 0$.

- La lunghezza del passo è determinata con una tecnica di ricerca in linea inesatta.
- La convergenza è assicurata dal Teorema di Zoutendijk.

Esempio

Esempio 16 *Sia dato il problema di minimizzazione*

$$\min x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2.$$

Si scrivano le prime iterazioni del metodo del metodo del gradiente con una ricerca di linea esatta a partire dal punto iniziale

$$x^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Esempio

Iterazione 0.

$$d^0 = -\nabla f(x^0) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Si determina il passo

$$\alpha^0 = \arg \min_{\alpha} f(x^0 + \alpha d^0) = \arg \min_{\alpha} \phi(\alpha) = \alpha(\alpha - 2).$$

Annullando la derivata $\dot{\phi}(\alpha) = 2\alpha - 2 = 2(\alpha - 1) = 0$, si ottiene $\alpha^0 = 1$ e quindi

$$x^1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Esempio

Iterazione 1. Verifica condizioni di arresto

$$\nabla f(x^1) = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \neq 0$$

quindi il punto non è ottimo. La nuova direzione è

$$d^1 = -\nabla q(x^1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Esempio

Determiniamo $\alpha^1 = \arg \min_{\alpha} f(x^1 + \alpha d^1) = \arg \min_{\alpha} (5\alpha^2 - 2\alpha - 1)$. Annullando la derivata

$$\dot{\phi}(\alpha) = 10\alpha - 2 = 0$$

si ottiene $\alpha^1 = \frac{1}{5}$. Quindi

$$\begin{aligned} x^2 &= x^1 - \alpha^1 \nabla f(x^1) = \\ &= \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1/5 \\ 1/5 \end{bmatrix} = \begin{bmatrix} -0.8 \\ 1.2 \end{bmatrix} \end{aligned}$$

Iterazione 2. Verifica condizione di arresto

$$\nabla f(x^2) = \begin{bmatrix} 1/5 \\ -1/5 \end{bmatrix} \neq 0,$$

Esempio

quindi il punto x^2 non è ottimo. Si calcola la direzione

$$d^2 = -\nabla f(x^2) = \begin{bmatrix} -\frac{1}{5} \\ \frac{1}{5} \end{bmatrix} = \begin{pmatrix} -0.2 \\ 0.2 \end{pmatrix}.$$

Determiniamo $\alpha^2 = \arg \min_{\alpha} f(x^2 + \alpha d^2) = \arg \min_{\alpha} (\alpha^2 - 2\alpha - 30)/25$. Annullando la derivata

$$\dot{\phi}(\alpha) = 2\alpha - 2 = 0$$

Esempio

si ottiene $\alpha^2 = 1$. Quindi il punto $x^3 = x^2 + \alpha^2 d^2$ è

$$x^3 = \begin{bmatrix} -1 \\ 7/5 \end{bmatrix} = \begin{bmatrix} -1 \\ 1,4 \end{bmatrix}.$$

Iterazione 3. Verifica condizione di arresto:

$$\nabla q(x^3) = \begin{bmatrix} -\frac{1}{5} \\ -\frac{1}{5} \end{bmatrix} \neq 0$$

e il punto non è ottimo. Si dovrebbe proseguire.

Noi terminiamo qui l'applicazione dell'algoritmo. La rapidità di convergenza è determinata dal rapporto

$$\frac{\lambda_{\max}(Q)}{\lambda_{\min}(Q)} = \left(\frac{3 + \sqrt{5}}{3 - \sqrt{5}} \right).$$

□

Il metodo di Newton puro (per l'ottimizzazione)

- Utilizza come direzione di ricerca la direzione di Newton

$$p_k = -H_f(x_k)^{-1} \nabla f(x_k)$$

dove $H_f(x_k)$ è la matrice Hessiana $\nabla^2 f(x_k)$.

- Questa direzione è ottenuta considerando lo sviluppo in serie di Taylor di $f(x)$ in un intorno di x_k

$$f(x_k + p) \approx f_k + p^T \nabla f_k + \frac{1}{2} p^T H_f p := m_k(p)$$

Supponendo che $H_f(x_k)$ sia definita positiva, la direzione di Newton è il vettore p che minimizza m_k .

- La lunghezza del passo è $\alpha_k = 1$ per ogni k .

Una interpretazione del metodo di Newton puro

- Una interpretazione del metodo di Newton equivalente alla precedente è ottenuta considerando il sistema di equazioni non lineari delle condizioni del primo ordine

$$\nabla f(x) = 0.$$

- La direzione di Newton p_k è ottenuta applicando il classico metodo di Newton per le equazioni non lineari alle condizioni del primo ordine.

- Dunque il metodo di Newton puro si scrive

$$x_{k+1} = x_k - H_f(x_k)^{-1} \nabla f(x_k)$$

- Come si comporta il metodo di Newton?
- La direzione di Newton è di discesa in x_k se l'Hessiana $H_f(x_k)$ è definita positiva.
- Poichè la matrice Hessiana $H_f(x_k)$ potrebbe non essere definita positiva, la direzione $p_k = -H_f(x_k)^{-1} \nabla f(x_k)$ potrebbe non essere di discesa.

Le proprietà di convergenza del metodo sono fortemente legate alla definita positività della matrice Hessiana!

Teorema

Si supponga che la funzione obiettivo $f \in C^2$ e che l'Hessiano $\nabla^2 f(x)$ sia Lipschitziano in un intorno di una soluzione x^* che soddisfa le condizioni sufficienti del secondo ordine. Si consideri il procedimento iterativo $x_{k+1} = x_k + p_k$, dove $p_k = -H_f(x_k)^{-1} \nabla f(x_k)$.

Allora

- se il punto iniziale x_0 è sufficientemente vicino a x^* , la successione degli iterati x_k converge a x^* ;
- la velocità di convergenza è quadratica;
- la successione $\{\|\nabla f(x_k)\|\}$ converge quadraticamente a 0.

Per la dimostrazione si veda [Nocedal, Wright]

Come si calcola la direzione di Newton?

- Per trovare p_k **non** si inverte $H_f(x_k)$ ma si risolve il sistema $H_f(x_k)p = -\nabla f(x_k)$ ad esempio mediante fattorizzazione di $H_f(x_k)$ come matrice in forma di prodotto LDL^T , con L matrice triangolare inferiore e D matrice diagonale con elementi positivi sulla diagonale.

Vantaggi e svantaggi del metodo di Newton puro?

Vantaggi:

- converge rapidamente quando si avvicina alla soluzione x^* .

Svantaggi:

- può non convergere alla soluzione se l'iterato iniziale è “troppo distante” dalla soluzione;
- anche se converge, può avere un comportamento sbagliato se la funzione non è convessa;
- ha un elevato costo computazionale;
- il “sistema di Newton” può essere mal condizionato.

Al fine di ottenere un metodo robusto ed efficiente è necessario modificare opportunamente il metodo di Newton puro!

Possibili modifiche?

- metodi di Newton **modificati** nei quali la matrice Hessiana è perturbata in modo da renderla definita positiva (ad esempio $B_k = \nabla^2 f(x_k) + E_k$);
- tecniche in cui si alterna la direzione di Newton con la direzione opposta o con quella dell'antigradiente (test sulla condizione dell'angolo);
- metodi di Newton **inesatti** (ad esempio Newton-CG);
- metodi di Newton **rilassati** ($\alpha_k \leq 1$);
- metodi **quasi** Newton.

Metodo di Newton rilassato

- Il metodo di Newton rilassato è descritto dall'iterazione:

$$x_{k+1} = x_k - \alpha_k H_f(x_k)^{-1} \nabla f(x_k)$$

dove la lunghezza del passo $\alpha_k \leq 1$ è calcolata al solito con una tecnica di ricerca in linea (line search).

- La ricerca in linea garantisce la convergenza globale del metodo di Newton rilassato (Teorema di Zoutendijk) quando le matrici hessiane H_k (o le modifiche) sono definite positive e hanno numero di condizione uniformemente limitato:

$$\|H_k\| \|H_k^{-1}\| \leq M, \quad \text{per ogni } k.$$

Introduzione

- La funzione quadratica strettamente convessa

$$f(x) = \frac{1}{2}x^T A x - b^T x, \quad A \text{ definita positiva}$$

ammette un unico punto di minimo globale $x^* = A^{-1}b$ che è soluzione del sistema lineare

$$Ax = b$$

- I metodi delle direzioni coniugate sono stati originariamente introdotti come metodi iterativi per la risoluzione di sistemi di equazioni lineari e sono stati successivamente estesi alla minimizzazione di funzioni non quadratiche.

Direzioni A-coniugate

- Una delle proprietà più notevoli del metodo del gradiente coniugato è quella di generare, ad ogni iterazione, delle *direzioni coniugate*.
- L'importanza di tale proprietà sta nel fatto che è possibile minimizzare la funzione $f(x)$ in n passi minimizzandola successivamente lungo direzioni coniugate.

Definizione

Sia A una matrice $n \times n$ simmetrica e definita positiva. Due vettori u e v di \mathbb{R}^n si dicono coniugati rispetto ad A , oppure A -coniugati se

$$u^T A v = 0.$$

Osservazione

Esiste almeno un insieme di n vettori che sono A -coniugati: essi sono gli autovettori di A .

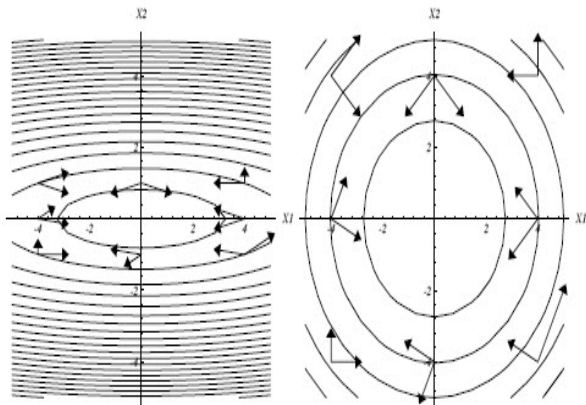


Figure: Sinistra: coppie di vettori A-coniugati; destra: gli stessi vettori in un piano deformato con la matrice A .

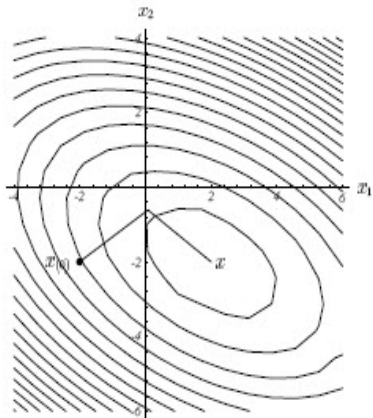


Figure: Interpretazione geometrica del metodo del gradiente coniugato in \mathbb{R}^2 : le due direzioni di ricerca sono A -coniugate, per cui la seconda necessariamente passa per il centro dell'ellissoide, e quindi per il punto soluzione x^* .

Teorema

Sia $x_0 \in \mathbb{R}^n$ e sia $\{x_k\}$ la successione generata dal metodo delle direzioni coniugate. Allora

$$\nabla f(x_k)^T p_i = 0, \quad \text{per } i = 0, \dots, k-1 \quad (9)$$

e x_k minimizza $f(x) = \frac{1}{2}x^T A x - b^T x$ sull'insieme

$$\{x \mid x = x_0 + \text{span}\{p_0, \dots, p_{k-1}\}\} \quad (10)$$

Il metodo del gradiente coniugato

- Il metodo del gradiente coniugato non utilizza un insieme di n direzioni A -coniugate predefinite ma genera tali vettori mentre costruisce la successione dei punti $\{x_k\}$.
- Si dimostra che il metodo del gradiente coniugato è un metodo delle direzioni coniugate e quindi gode della proprietà di terminazione quadratica. Tuttavia a causa degli errori di arrotondamento nelle operazioni, non è detto che il metodo fornisca una soluzione accurata in n passi per matrici dense.
- Per matrici sparse di grosse dimensioni il numero di iterazioni si mantiene ragionevolmente basso. Da ciò l'uso del metodo come alternativa alla fattorizzazione di Cholesky.

Il metodo del gradiente coniugato

E' un procedimento iterativo descritto da

$$x_{k+1} = x_k + \alpha_k p_k$$

dove

- $p_k = -\nabla f(x_k) + \beta_{k-1} p_{k-1}$ direzione di discesa A -coniugata a tutte le direzioni precedenti p_0, p_1, \dots, p_{k-1}
- $\alpha_k = -\frac{(\nabla f(x_k))^T p_k}{p_k^T A p_k}$
- $\beta_{k-1} = \frac{(\nabla f(x_k))^T A p_{k-1}}{p_{k-1}^T A p_{k-1}}$ scelto in modo che p_k sia A -coniugato a p_{k-1}
- $p_0 = -\nabla f(x_0), \alpha_0 = -\frac{(\nabla f(x_0))^T p_0}{p_0^T A p_0}$

Algoritmo

Metodo del gradiente coniugato (versione preliminare)

Dato x_0 , calcolare $g_0 = \nabla f(x_0)$.

Se $g_0 = 0$, stop, altrimenti porre $p^0 = -g^0$. Per $k = 0, 1, 2, \dots$

$$\alpha_k = -\frac{g_k^T p_k}{p_k^T A p_k}$$

$$x_{k+1} = x_k + \alpha_k p_k$$

$$g_{k+1} = A x_{k+1} - b; \text{ se } g_{k+1} = 0, \text{ stop}$$

$$\beta_k = \frac{g_{k+1}^T A p_k}{p_k^T A p_k}$$

$$p_{k+1} = -g_{k+1} + \beta_k p_k$$

$$k = k + 1$$

- Un criterio di arresto del metodo del gradiente coniugato che viene comunemente usato è

$$\|r_k\|_2 < \epsilon \|r_0\|$$

p_k è una direzione di discesa

$$\begin{aligned}
 p_k^T \nabla f(x_k) &= (-g_k + \beta_{k-1} p_{k-1})^T g_k \\
 &= -g_k^T g_k + \beta_{k-1} p_{k-1}^T (g_{k-1} + \alpha_{k-1} A p_{k-1}) \\
 &= -g_k^T g_k + \beta_{k-1} \left(p_{k-1}^T g_{k-1} + \frac{-g_{k-1}^T p_{k-1}}{p_{k-1}^T A p_{k-1}} (p_{k-1}^T A p_{k-1}) \right) \\
 &= -g_k^T g_k < 0
 \end{aligned}$$

N.B: $g_k := \nabla f(x_k)$

Teorema

Sia A una matrice reale $n \times n$ simmetrica e definita positiva e sia x^ la soluzione del sistema $Ax = b$. Si supponga che l'iterato x_k generato dal metodo CG non sia la soluzione x^* . Allora, i vettori p_k generati dall'algoritmo CG sono A -coniugati, cioè soddisfano la relazione*

$$p_k^T A p_j = 0, \quad j = 0, \dots, k-1, \quad k = 1, 2, \dots, n-1. \quad (11)$$

Quindi, la successione $\{x_k\}$ converge in al più n passi, cioè $x_m = x^$ per un certo $m \leq n$.*

Inoltre, i residui $r_j = b - Ax_j$ soddisfano la condizione

$$r_k^T r_j = 0, \quad j = 0, \dots, k-1, \quad k = 1, 2, \dots, n-1. \quad (12)$$

Teorema

Se la matrice A possiede al più $s \leq n$ autovalori distinti, allora il metodo del gradiente coniugato converge in al più s iterazioni.

Una versione pratica dell'algoritmo del gradiente coniugato

- E' possibile fornire una versione del metodo del gradiente coniugato che ha un costo computazionale più ridotto nel quale i parametri α_k e β_k sono aggiornati con le formule

$$\alpha_k = \frac{r_k^T r_k}{p_k^T A p_k} \quad \beta_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$$

Metodo del gradiente coniugato

Dato x_0 , calcolare $r_0 = b - Ax_0$.

Se $r_0 = 0$, stop, altrimenti porre $p_0 = r_0$. Per $k = 0, 1, 2, \dots$

$$\alpha_k = \frac{r_k^T r_k}{p_k^T A p_k}$$

$$x_{k+1} = x_k + \alpha_k p_k$$

$$r_{k+1} = r_k - \alpha_k A p_k$$

se $\|r_{k+1}\| < \epsilon \|b\|$, stop

$$\beta_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$$

$$p_{k+1} = r_{k+1} + \beta_k p_k$$

$$k = k + 1$$

- Ad ogni iterazione del precedente algoritmo, l'operazione che richiede il maggior costo computazionale è il prodotto della matrice A per il vettore p_k , che richiede n^2 moltiplicazioni se la matrice A è densa. Se il numero di iterazioni è n si ha, allora, un costo complessivo dell'ordine di n^3 , superiore a quello richiesto dal metodo di Cholesky ($n^3/3$). Si intuisce, pertanto, che il metodo diventa interessante quando la matrice è sparsa e/o quando il numero di iterazioni necessarie è decisamente inferiore a n .
- Non è necessario conoscere esplicitamente la matrice A ma è sufficiente saper calcolare il prodotto di A per un qualsiasi vettore v di \mathbb{R}^n .
- A causa degli errori di arrotondamento nelle operazioni, non è detto che il metodo fornisca una soluzione accurata in n passi per matrici dense.
- Per matrici sparse di grosse dimensioni il numero di iterazioni si mantiene ragionevolmente basso. Da ciò l'uso del metodo come alternativa alla fattorizzazione di Cholesky.

Esempio

- In Figura 71 sono illustrati i risultati che si ottengono applicando, a partire dal punto $x_0 = (-1.5, 1)^T$, il metodo del gradiente coniugato al sistema $Ax = b$, con

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

- La direzione $p_0 = r_0$ è la direzione del gradiente nel punto x_0 e risulta ortogonale alla tangente in tale punto all'ellisse. La direzione successiva p_1 è A -coniugata a p_0 e passa per il centro dell'ellisse.
- La minimizzazione lungo tale direzione fornisce quindi il punto di minimo della $f(x)$, ossia la soluzione del sistema lineare dato.

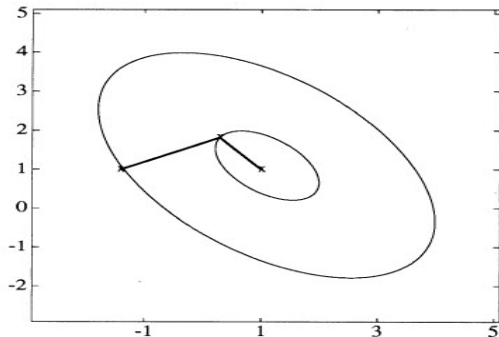


Figure: Illustrazione dei risultati ottenuti mediante l'algoritmo del gradiente coniugato nella risoluzione di un sistema lineare del secondo ordine. La soluzione viene ottenuta, a partire dal punto $x^{(0)} = (-1.5, 1)^T$ mediante due minimizzazioni lungo due direzioni coniugate.

- La (??) può essere riscritta in termini della norma euclidea $\| \cdot \|$ come segue

$$\|x^* - x_k\| \leq \sqrt{K(A)} \left(\frac{\sqrt{K(A)} - 1}{\sqrt{K(A)} + 1} \right)^k \|x^* - x_k\| \quad (13)$$

- Quando $K(A) = 1$, il fattore $\frac{\sqrt{K(A)}-1}{\sqrt{K(A)}+1}$ nella (13) si annulla e diventa uguale a 1 quando $K(A) \rightarrow \infty$. Pertanto, dalle relazioni precedenti, si deduce che, quanto più cresce il numero di condizione $K(A)$ tanto più lenta diventa la convergenza del metodo del gradiente coniugato.
- Dunque, il metodo del gradiente coniugato converge linearmente, ma la costante asintotica dell'errore è più piccola del caso del metodo di discesa ripida.

- Ad esempio, se $K(A) = 10^4$ e si richiede di ridurre la norma dell'errore iniziale $\|x^* - x_0\|$ di un fattore $\epsilon = 10^{-4}$ occorre eseguire $k \leq 496$ iterazioni con il metodo del gradiente coniugato e $k \leq 46052$ iterazioni con il metodo di discesa ripida.
- Dunque, il metodo del gradiente coniugato converge linearmente, ma la costante asintotica dell'errore è più piccola del caso del metodo di discesa ripida

Minimi quadrati lineari

- Si consideri il problema di *minimi quadrati lineari*

$$\min_x f(x) = \frac{1}{2} \|Qx - c\|^2 \quad Q \in \mathbb{R}^{m \times n}, m \geq n$$

- E' un problema di programmazione quadratica convessa del tipo:

$$\min_x f(x) = \frac{1}{2} x^T Q^T Q x - x^T Q^T c$$

- Il problema ammette soluzione e x^* è soluzione se e solo se

$$Q^T Q x = Q^T c$$

- Se la matrice dei coefficienti $Q^T Q$ è definita positiva è possibile usare il GC per risolvere il sistema delle equazioni normali $Q^T Q x = Q^T c$

Algoritmo CGLS

Dato x_0 , calcolare $r_0 = c - Qx_0$.

Se $r_0 = 0$, stop, altrimenti porre $z_0 = Q^T r_0$, $p_0 = z_0$.

Per $k = 0, 1, 2, \dots$

$$w_k = Qp_k$$

$$\alpha_k = \frac{z_k^T z_k}{w_k^T w_k}$$

$$x_{k+1} = x_k + \alpha_k p_k$$

$$r_{k+1} = r_k - \alpha_k w_k$$

se $\|r_{k+1}\| < \epsilon \|b\|$, stop

$$z_{k+1} = Q^T r_{k+1}$$

$$\beta_k = \frac{z_{k+1}^T z_{k+1}}{z_k^T z_k}$$

$$p_{k+1} = z_{k+1} + \beta_k p_k$$

$$k = k + 1$$