Programmazione non-lineare: caso monodimensionale

Giovanni Righini

Ricerca Operativa

Università degli Studi di Milano

Scelta del passo

Una volta scelta la direzione $d^{(k)}$ nei metodi line search rimane un problema di minimizzazione ad una sola variabile

$$\text{minimize } f(x^{(k+1)}) = f(x^{(k)} + s_k d^{(k)})$$

dove la variabile è lo scalare $s_k \geq 0$.

L'ottimizzazione esatta di s_k non è indispensabile; una buona approssimazione è sufficiente per avviare l'iterazione successiva dopo aver migliorato $f(x)$.

Gli algoritmi per determinare il passo possono essere classificati in

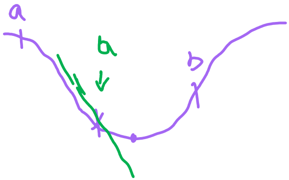
- algoritmi che richiedono il calcolo della derivata,
- algoritmi *derivative-free*.

Metodo di bisezione

Richiede il calcolo della derivata.

Dato un intervallo iniziale $r = [a, b]$ per s_k:

1. calcolare $\nabla f(x^{(k)} + \frac{a+b}{2} d^{(k)})$; derivata nel punto medio
2. se è positiva, porre $r := [a, \frac{a+b}{2}]$;
3. se è negativa, porre $r := [\frac{a+b}{2}, b]$;
4. ripetere finché r è abbastanza piccolo.



I numeri di Fibonacci

La sequenza dei numeri di Fibonacci inizia con $F_0 = 0$ e $F_1 = 1$ e si ricava applicando la ricorsione

$$F_k = F_{k-1} + F_{k-2}.$$

Si ottiene così:

k	0	1	2	3	4	5	6	7	8	9	10	11	...
F_k	0	1	1	2	3	5	8	13	21	34	55	89	...

Tabella: I primi numeri di Fibonacci.

Una proprietà

Proprietà. Dati 4 numeri di Fibonacci consecutivi a partire dal k-esimo, si ha

$$F_{k+1}F_{k+2} - F_kF_{k+3} = (-1)^k \quad \forall k \geq 0.$$

Esempio.

$$F_6F_7 - F_5F_8 = 8 \times 13 - 5 \times 21 = 104 - 105 = -1 \quad (k = 5, \text{dispari})$$

$$F_7F_8 - F_6F_9 = 13 \times 21 - 8 \times 34 = 273 - 272 = 1 \quad (k = 6, \text{pari}).$$

k 1 2 3

Saltata?

Dimostrazione (per induzione).

$$F_{k+1}F_{k+2} - F_kF_{k+3} = (-1)^k \quad \forall k \geq 0.$$

- Base dell'induzione (per $k = 0$):

$$F_1F_2 - F_0F_3 = 1 \times 1 - 0 \times 2 = 1^0$$

- Passo induttivo (da $k-1$ a k per ogni $k \geq 1$:

$$F_kF_{k+1} - F_{k-1}F_{k+2} = (-1)^{k-1} \ \Rightarrow F_{k+1}F_{k+2} - F_kF_{k+3} = (-1)^k.$$

Infatti:

$$F_{k+1}F_{k+2} - F_kF_{k+3} = [F_{k+1}(F_k + F_{k+1})] - [F_k(2F_{k+1} + F_k)] =$$
$$=F_kF_{k+1} + F_{k+1}^2 - 2F_kF_{k+1} - F_k^2 = (F_{k+1}^2 - F_k^2) - F_kF_{k+1} =$$
$$=(F_{k+1} + F_k)(F_{k+1} - F_k) - F_kF_{k+1} = F_{k+2}F_{k-1} - F_kF_{k+1} = \text{ (per ipotesi)}$$
$$= -(-1)^{k-1} = (-1)^k. \quad \text{[c.v.d.]}$$

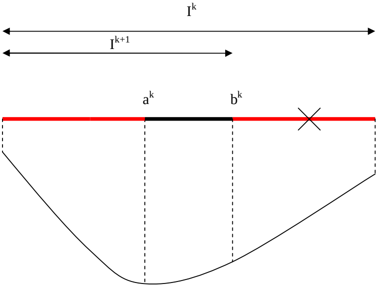
Il problema

- È data una funzione continua $f(x)$ di una sola variabile x.
- Si vuole cercare un punto di minimo di $f(x)$.
- È dato un intervallo di incertezza iniziale I^0.
- È richiesto un massimo intervallo di incertezza finale Δ.
- Si assume che la funzione sia *unimodale* nell'intervallo I^0, cioè abbia un solo punto di minimo nell'intervallo.
- Si suppone di non poter/voler calcolare la derivata prima di $f(x)$, come invece è richiesto dal metodo di bisezione.
- Si suppone che sia possibile valutare $f(x)$ in punti diversi, purché distanti tra loro almeno ϵ (risoluzione).

Osservazione. Sarebbe ancora possibile usare il metodo di bisezione, valutando in ogni intervallo due punti distanti ϵ tra loro, posti al centro dell'intervallo stesso. L'informazione che se ne ricaverebbe sarebbe equivalente a quella fornita dal calcolo della derivata prima al centro dell'intervallo. In tal modo sarebbero necessarie *due* valutazioni della funzione ad ogni iterazione. Con il metodo dei numeri di Fibonacci, invece, basta *una* valutazione della funzione ad ogni iterazione.

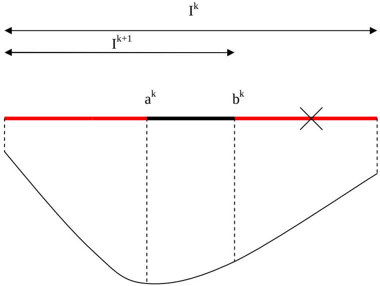
Iterazione generica

Alla generica iterazione k si ha un intervallo di incertezza I^k.
Si considerino due punti a^k e b^k interni all'intervallo, che lo dividono in tre parti.
Si conosca il valore della funzione $f(x)$ nei due punti interni.
Se $f(a^k) > f(b^k)$, allora il minimo di $f(x)$ non cade nella prima parte.
Se $f(a^k) < f(b^k)$, allora il minimo di $f(x)$ non cade nella terza parte.



Punti di valutazione

Alla successiva iterazione l'intervallo di incertezza risulta composto da due delle tre parti dell'intervallo di incertezza precedente.

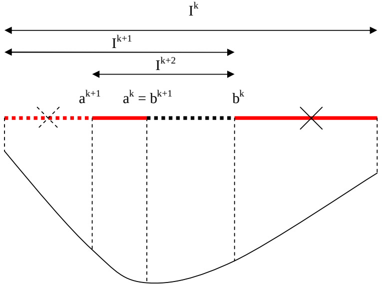


Uno dei due punti interni precedenti diventa un estremo dell'intervallo di incertezza.

Punti di valutazione

Per simmetria, ad ogni iterazione i punti in cui valutare $f(x)$ sono scelti in modo simmetrico nell'intervallo di incertezza corrente. Si ha quindi:

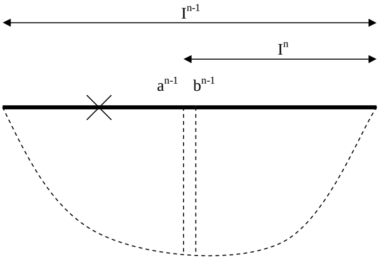
$$I^k = I^{k+1} + I^{k+2} \quad \forall k \geq 0.$$



In uno dei due punti interni la funzione è già stata valutata in precedenza.

Intervallo finale

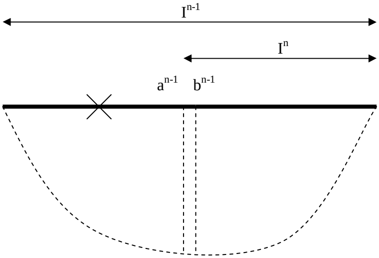
Osservazione. Più è grande l'intervallo "scartato" all'iterazione k e più risultano piccoli quelli "scartabili" all'iterazione $k+1$.
L'iterazione di massima efficacia è quella che consente di scartare metà dell'intervallo di incertezza, valutando due punti interni distinti vicinissimi tra loro (a distanza ϵ).



Tuttavia all'iterazione successiva il primo e il terzo intervallo sarebbero larghi solo ϵ e quindi si avrebbe un'iterazione di minima efficacia.

Intervallo finale

Si vuole quindi arrivare a compiere un'iterazione di massima efficacia *per ultima*.



Si ha perciò:

$$I^{n-1} = 2I^n - \epsilon.$$

Progressione delle iterazioni

Dalle due relazioni

$$I^k = I^{k+1} + I^{k+2} \;\; \forall k \geq 0$$

$$I^{n-1} = 2I^n - \epsilon$$

si ricava:

Fibonacci

$$I^{n-2} = I^{n-1} + I^n = 3I^n - \epsilon$$

$$I^{n-3} = I^{n-2} + I^{n-1} = 5I^n - 2\epsilon$$

$$I^{n-4} = I^{n-3} + I^{n-2} = 8I^n - 3\epsilon$$

$$\ldots$$

$$I^{n-k} = I^{n-k+1} + I^{n-k+2} = F_{k+2}I^n - F_k\epsilon \;\; \forall k \geq 1$$

$$\ldots$$

$$I^2 = I^3 + I^4 = F_nI^n - F_{n-2}\epsilon$$

$$I^1 = I^2 + I^3 = F_{n+1}I^n - F_{n-1}\epsilon$$

$$I^0 = I^1 + I^2 = F_{n+2}I^n - F_n\epsilon.$$

Perciò:

$$I^0 = F_{n+2}I^n - F_n\epsilon \;\; \text{ovvero} \;\; I^n = \frac{I^0}{F_{n+2}} + \frac{F_n}{F_{n+2}}\epsilon.$$

Numero di iterazioni

Dalla relazione

tende a scendere

$$I^n = \frac{I^0}{F_{n+2}} + \frac{F_n}{F_{n+2}}\epsilon$$

→ tende a rapp. aureo al quad (costante)

e dal requisito sull'incertezza finale

$$I^n \leq \Delta$$

si ricava il numero di iterazioni necessarie:

$$\overline{n} = \min\{n | \frac{I^0}{F_{n+2}} + \frac{F_n}{F_{n+2}}\epsilon \leq \Delta\}.$$

Formula di Binet:

$$F_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right)$$

Scelta del primo punto interno

Dalle relazioni

$$I^{\overline{n}} = \frac{I^0}{F_{\overline{n}+2}} + \frac{F_{\overline{n}}}{F_{\overline{n}+2}}\epsilon$$

e

$$I^1 = F_{\overline{n}+1} I^{\overline{n}} - F_{\overline{n}-1}\epsilon$$

si ricava I^1 in funzione di I^0:

$$\begin{aligned}
I^1 &= F_{\overline{n}+1}\left(\frac{I^0}{F_{\overline{n}+2}} + \frac{F_{\overline{n}}}{F_{\overline{n}+2}}\epsilon\right) - F_{\overline{n}-1}\epsilon = \\
&= \frac{F_{\overline{n}+1}}{F_{\overline{n}+2}} I^0 + \frac{\epsilon}{F_{\overline{n}+2}}(F_{\overline{n}+1}F_{\overline{n}} - F_{\overline{n}+2}F_{\overline{n}-1}) = \\
&= \frac{F_{\overline{n}+1}}{F_{\overline{n}+2}} I^0 + \frac{\epsilon}{F_{\overline{n}+2}}(-1)^{\overline{n}-1}.
\end{aligned}$$

Esempio

- È dato un intervallo di incertezza iniziale $I^0 = [0, 100]$.
- È richiesto un massimo intervallo di incertezza finale $\Delta = 2$.
- È data una risoluzione $\epsilon = 1$.

Esempio

Scelta del numero di iterazioni

$$\overline{n} = \min\{n | \frac{100}{F_{n+2}} + \frac{F_n}{F_{n+2}} \leq 2\} = 9.$$

Infatti si ha:
per $n = 8$: $\frac{100}{F_{10}} + \frac{F_8}{F_{10}} = 100/55 + 21/55 = 121/55 > 2$,
per $n = 9$: $\frac{100}{F_{11}} + \frac{F_9}{F_{11}} = 100/89 + 34/89 = 134/89 < 2$.

k	0	1	2	3	4	5	6	7	8	9	10	11	...
F_k	0	1	1	2	3	5	8	13	21	34	55	89	...

Scelta del primo punto interno

$$I^1 = \frac{F_{\overline{n}+1}}{F_{\overline{n}+2}} I^0 + \frac{\epsilon}{F_{\overline{n}+2}} (-1)^{\overline{n}-1} =$$

$$= \frac{55}{89} 100 + \frac{1}{89} (-1)^8 = 5500/89 + 1/89 = 5501/89 \approx 61,81.$$

Esempio

Iterazioni. Gli intervalli di incertezza successivi risultano avere le seguenti ampiezze:

$$
\begin{aligned}
I^2 &= I^0 - I^1 = \frac{8900 - 5501}{89} = \frac{3399}{89} \\
I^3 &= I^1 - I^2 = \frac{5501 - 3399}{89} = \frac{2102}{89} \\
I^4 &= I^2 - I^3 = \frac{3399 - 2102}{89} = \frac{1297}{89} \\
I^5 &= I^3 - I^4 = \frac{2102 - 1297}{89} = \frac{805}{89} \\
I^6 &= I^4 - I^5 = \frac{1297 - 805}{89} = \frac{492}{89} \\
I^7 &= I^5 - I^6 = \frac{805 - 492}{89} = \frac{313}{89} \\
I^8 &= I^6 - I^7 = \frac{492 - 313}{89} = \frac{179}{89} \\
I^9 &= I^7 - I^8 = \frac{313 - 179}{89} = \frac{134}{89}.
\end{aligned}
$$

Conclusioni

- Il metodo dei numeri di Fibonacci consente di approssimare il minimo di una funzione di una sola variabile continua.
- Deve essere noto un intervallo di incertezza iniziale e la funzione deve essere unimodale in esso.
- Il metodo non richiede il calcolo della derivata prima della funzione.
- Il metodo richiede di valutare la funzione in un numero di punti dello stesso ordine di grandezza del numero di iterazioni.