Costruzione di metodi di minimizzazione convergenti di tipo Broyden con algebre di matrici di bassa complessità

Candidato: **Isabella Iori**

Relatore: **Prof. Carmine Di Fiore**
Correlatore: **Dott. Stefano Cipolla**

Università degli Studi di Roma "Tor Vergata"
Corso di Laurea Magistrale in Matematica Pura ed Applicata

26 Novembre 2015

Minimizzazione non vincolata

Problemi di natura teorica e pratica

$$\Downarrow$$

Formulazione di un modello matematico:

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x})$$

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

$\mathbf{x} = (x_1, x_2, \cdots, x_n)$ per $n \geq 1$.

Obiettivo: trovare $\mathbf{x}^*$ tale che $f(\mathbf{x}^*) = \min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x})$.

$$\Downarrow$$

Tecniche di tipo algoritmico

Algoritmi di minimizzazione non vincolata di tipo line search

$$\mathbf{x}_0 \quad \Rightarrow \quad \{\mathbf{x}_k\}_{k=1}^{\infty}$$

Differenti strategie per passare da $\mathbf{x}_k$ a $\mathbf{x}_{k+1}$

$\Downarrow$

LINE SEARCH

Algoritmo di discesa

Scegli $\mathbf{x}_0 \in \mathbb{R}^n$; $\quad k := 0$;
While $\nabla f(\mathbf{x}_k) \neq 0$
{
  calcola $\mathbf{d}_k \in \mathbb{R}^n$; /* direzione di discesa */
  calcola $\alpha_k \in \mathbb{R}$; /* lunghezza del passo */
  $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$;
  $k := k + 1$;
}

$\Downarrow$

Aspetti critici:

- scelta del passo $\alpha_k \Longrightarrow$ STEP LENGTH
- scelta della direzione $\mathbf{d}_k$ $(\nabla f(\mathbf{x}_k))^T d_k < 0)$

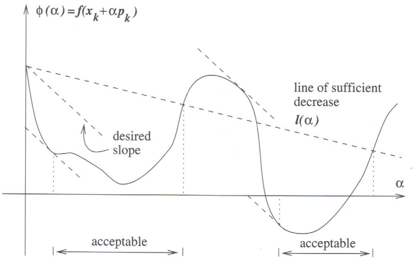
Step length e direzioni di discesa

Step length

Condizioni di Wolf:

$$\begin{cases} f(\mathbf{x}_k + \alpha_k \mathbf{d}_k) \le f(\mathbf{x}_k) + c_1 \alpha_k \nabla f(\mathbf{x}_k)^T \mathbf{d}_k \\ \nabla f(\mathbf{x}_k + \alpha_k \mathbf{d}_k)^T \mathbf{d}_k \ge c_2 \nabla f(\mathbf{x}_k)^T \mathbf{d}_k \end{cases}$$

con $0 < c_1 < 1/2$ e $0 < c_1 < c_2 < 1$.



Direzione Steepest Descent $\mathbf{d}_k^{SD}$

$$\mathbf{d}_k^{SD} = -\nabla f(\mathbf{x}_k)$$

- calcolo del gradiente
- convergenza lineare
- lenta in problemi “difficili”

Direzione di Newton $\mathbf{d}_k^{N}$

$$\mathbf{d}_k^{N} = -\nabla^2 f(\mathbf{x}_k)^{-1} \nabla f(\mathbf{x}_k)$$

- calcolo dell’Hessiano e del gradiente
- convergenza quadratica, ma....
- complessità $O(n^3) \rightarrow$ inadeguata per problemi di grandi dimensioni

Direzioni quasi-Newton

Obiettivo: ridurre il costo computazionale ad ogni iterazione $\longrightarrow$

- sostituire l'Hessiano con una sua approssimazione: $\nabla^2 f(\mathbf{x}_k) \approx B_k$
- aggiornare B_{k+1} usando B_k

$\Longrightarrow$ mantenere il modello nei pressi della soluzione

$\Longrightarrow$ aggiornamento di B_{k+1}^{-1} a partire da B_k^{-1}

$\Longrightarrow$ basso costo computazionale

$\Downarrow$

Direzione quasi Newton $\mathbf{d}_k^{QN}$

$$\mathbf{d}_k^{QN} = -B_k^{-1} \nabla f(\mathbf{x}_k)$$

B_k approssimazione definita positiva dell'Hessiano,

- $B_{k+1}\mathbf{s}_k = \mathbf{y}_k$, (equazione secante), $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$ e $\mathbf{y}_k = \nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k)$
- calcolo del gradiente
- il migliore ha convergenza superlineare (es.: classe di Broyden)
- complessità $O(n^2) \rightarrow$ valida alternativa per risolvere problemi di grandi dimensioni

Una classe di metodi quasi-Newton: Broyden ristretta

$$B_{k+1} = B_k - \frac{B_k \mathbf{s}_k \mathbf{s}_k^T B_k}{\mathbf{s}_k^T B_k \mathbf{s}_k} + \frac{\mathbf{y}_k \mathbf{y}_k^T}{\mathbf{y}_k^T \mathbf{s}_k} + \phi(\mathbf{s}_k^T B_k \mathbf{s}_k)\mathbf{v}_k \mathbf{v}_k^T \qquad \text{dove} \qquad \begin{cases} \phi \in [0,1] \\ \mathbf{y}_k = \mathbf{g}_{k+1} - \mathbf{g}_k, \\ \mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k, \\ \mathbf{v}_k = \left[\frac{\mathbf{y}_k}{\mathbf{y}_k^T \mathbf{s}_k} - \frac{B_k \mathbf{s}_k}{\mathbf{s}_k^T B_k \mathbf{s}_k} \right] \end{cases}$$

Algoritmo: Broyden Class

Scegli $\mathbf{x}_0 \in \mathbb{R}^n$; B_0 r. s. d.p., $k := 0$;
While $\nabla f(\mathbf{x}_k) \neq 0$
{
- $\mathbf{d}_k := -B_k^{-1} \nabla f(\mathbf{x}_k)$;
- Calcola $\alpha_k \in \mathbb{R}$ che soddidfa le condizioni di Wolf;
- $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$;
- $B_{k+1} = \psi_\phi(B_k, \mathbf{s}_k, \mathbf{y}_k)$;
- $k := k + 1$;

}

Proprietà della classe di Broyden ristretta

- $\nabla f(\mathbf{x}_k + \alpha_k \mathbf{d}_k)^T \mathbf{d}_k \geq c_2 \nabla f(\mathbf{x}_k)^T \mathbf{d}_k \Longrightarrow \mathbf{y}_k^T \mathbf{s}_k > 0 \Longrightarrow B_{k+1}$ d.p.
- $B_{k+1} \mathbf{s}_k = \mathbf{y}_k$
- complessità computazionale $O(n^2)$
- $B_{k+1} = (1 - \phi) B_{k+1}^{BFGS} + \phi B_{k+1}^{DFP}$.

Sia $D = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leq f(\mathbf{x}_0)\}$ insieme di livello.

Assunzione 1: f due volte differenziabile in modo continuo, convessa e limitata inferiormente. Inoltre D è convesso ed esiste esiste una costante M per cui

$$\|H(\mathbf{x})\| \leq M$$

per ogni $\mathbf{x} \in D$.

- convergenza globale della classe (per problemi convessi) per $\phi \in [0, 1)$ (Byrd, 1987)
- convergenza superlineare per funzioni uniformemente convesse: $m\|\mathbf{z}\|^2 \leq \mathbf{z}^T H(\mathbf{x})\mathbf{z} \leq M\|\mathbf{z}\|^2$ per ogni $\mathbf{x} \in N(\mathbf{x}^*)$, con $N(\mathbf{x}^*)$ intorno di $\mathbf{x}^*$ e m, M costanti positive.

Metodi "di tipo" quasi-Newton: la Broyden class type

Obiettivo: $B_k \approx \tilde{B}_k$ per abbassare ulteriormente il costo computazionale e ridurre lo spazio di memoria scegliendo $\tilde{B}_k$ adeguatamente $\Longrightarrow$ metodi quasi-Newton type

Algoritmo: Broyden class type

Scegli $\mathbf{x}_0 \in \mathbb{R}^n$; $\mathbf{g}_0 = \nabla f(\mathbf{x}_0)$; $B_0, \tilde{B}_0$ d.p.; $\mathbf{d}_0 = -B_0^{-1}\mathbf{g}_0$; $k := 0$;

While $\nabla f(\mathbf{x}_k) \neq 0$

{

$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$;

$\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$;

$\mathbf{g}_{k+1} = \nabla f(\mathbf{x}_{k+1})$; $\mathbf{y}_k = \mathbf{g}_{k+1} + \mathbf{g}_k$;

$B_{k+1} = \psi_\phi(\tilde{B}_k, \mathbf{s}_k, \mathbf{y}_k)$;

- definisci $\tilde{B}_{k+1} \approx B_{k+1}$ d.p. , $\mathbf{d}_{k+1} = -\tilde{B}_{k+1}^{-1}\mathbf{g}_{k+1}$ $(\mathcal{NS})$;

oppure

- $\mathbf{d}_{k+1} = -B_{k+1}^{-1}\mathbf{g}_{k+1}$, definisci $\tilde{B}_{k+1} \approx B_{k+1}$ d.p. $(\mathcal{S})$;

$k := k + 1$;

}

Sotto quali ipotesi sulla matrice $\tilde{B}_k$ è garantita la convergenza globale del metodo secante e non secante?

Convergenza globale della classe di Broyden type secante

Convergenza globale del metodo secante

Sia $\mathbf{x}_0$ un punto iniziale per il quale f soddisfa l'Assunzione 1. Allora per ogni matrice definita positiva B_0, l'algoritmo generato con il metodo Broyden type ristretto $\mathcal{S}$ e α_k che soddisfa le condizioni di Wolf, produce una successione $\{\mathbf{x}_k\}$ tale che

$$\liminf_{k \to \infty} \|\mathbf{g}_k\| = 0$$

se per ogni k vale che

$$\operatorname{tr} \tilde{B}_k \leq \operatorname{tr} B_k$$

$$\det \tilde{B}_k \geq \det B_k$$

$$\frac{\|\tilde{B}_k \mathbf{s}_k\|^2}{(\mathbf{s}_k^T \tilde{B}_k \mathbf{s}_k)^2} \geq \frac{\|B_k \mathbf{s}_k\|^2}{(\mathbf{s}_k^T B_k \mathbf{s}_k)^2}$$

Convergenza globale del metodo non secante

Stesso enunciato del metodo secante, ma delle tre ipotesi sono necessarie solo quelle su traccia e determinante.

Convergenza globale dei metodi $\mathcal{L}^{(k)}QN$

Powell
Convergenza del metodo
BFGS
(1976)

$\Longrightarrow$

C.Di Fiore, S.Fanelli, F.Lepore, P.Zellini
Convergenza del metodo
$\mathcal{L}^{(k)}$**BFGS NON SECANTE**
(2003)

$\Downarrow$ $\searrow$ $\swarrow$

S.Cipolla, C.Di Fiore, F.Tudisco, P.Zellini
Convergenza del metodo
$\mathcal{L}^{(k)}$**BFGS SECANTE**
(2015)

$\Longrightarrow$

convergenza della
CLASSE DI
METODI
$\mathcal{L}^{(k)}$BROYDEN
SECANTE e
NON SECANTE

$\nwarrow$

R.H. Byrd, S.J.Nocedal, Y.X.Yuan
Convergenza della
CLASSE DI BROYDEN
(1987)

$\tilde{B}_k \in \mathcal{L}^{(k)}$=algebra di bassa complessità computazionale $\Rightarrow$Broyden type:$O(n)$=costo per passo
=memoria richiesta

Algebre di matrici $\mathcal{L}$ e $\mathcal{L}^{(k)}$ di bassa complessità

Obiettivo: scegliere $\tilde{B}_k$ in uno spazio opportuno di bassa complessità

$\Longrightarrow$ Idea: coinvolgere algebre di matrici di bassa complessità:

- $M_n(\mathbb{C})$ = spazio delle matrici $n \times n$ con elementi in $\mathbb{C}$;
- $\|A\|_F^2 = (A, A) = \sum_{i,j=1}^n \bar{a}_{ij} a_{ij} = \sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2$ = norma di Frobenius;
- $\mathcal{L} :=$ sd $U = \{U diag(\mathbf{z}) U^H : \mathbf{z} \in \mathbb{C}^n\}$, U=trasformata discreta veloce unitaria;
- $\min_{X \in \mathcal{L}} \|A - X\|_F = \|A - \mathcal{L}_A\|_F$.

- $\mathcal{L}_A = U diag((U^H A U)_{ii}) U^H$
- $[\mathbf{z}]_i = (U^H A U)_{ii}, i = 1, \dots, n$

- A r. s. d.p. $\Rightarrow \mathcal{L}_A$ r. s. d.p.
- $\mathrm{tr}(\mathcal{L}_A) = \mathrm{tr}(A)$
- $\det(A) \leq \det(\mathcal{L}_A)$

$$\mathcal{L}^{(k)} := \text{ sd } U_k = \{U_k diag(\mathbf{z}_k) U_k^H : \mathbf{z}_k \in \mathbb{C}^n\}$$

Una possibile scelta è:

$$\tilde{B}_k = \mathcal{L}^{(k)}_{B_k} = U_k diag(\mathbf{z}_k) U_k^H = U_k diag((U_k^H B_k U_k)_{ii}) U_k^H$$

$\Longrightarrow$ Con questa scelta le ipotesi su traccia e determinante sono immediatamente soddisfatte $\Longrightarrow$ il problema non secante risulta risolto

Convergenza della classe $\mathcal{L}^{(k)}$Broyden secante

Osservazione: $B_k \mathbf{s}_k = \sigma \tilde{B}_k \mathbf{s}_k$, con $\sigma \in \mathbb{R}^+ \Longrightarrow \frac{(\mathbf{s}_k^T \tilde{B}_k \mathbf{s}_k)^2}{\|\tilde{B}_k \mathbf{s}_k\|^2} \leq \frac{(\mathbf{s}_k^T B_k \mathbf{s}_k)^2}{\|B_k \mathbf{s}_k\|^2}$.

- è un'ipotesi più restrittiva, ma...
- permette di dimostrare l'esistenza di una $\tilde{B}_k$ di bassa complessità
- fornisce un modo pratico per costruire $\tilde{B}_k$
- $B_k \mathbf{s}_k = \sigma \tilde{B}_k \mathbf{s}_k \iff \mathbf{d}_{k+1}^{(\mathcal{NS})} = \sigma \mathbf{d}_{k+1}^{(\mathcal{S})}$

Problema 1: Total Non-Linear Problem

Dato $\mathbf{g}_{k+1} \in \mathbb{R}^n$, $B_{k+1} = \psi_\phi(\tilde{B}_k, \mathbf{s}_k, \mathbf{y}_k)$ definita positiva e $\mathbf{d}_{k+1}^{(\mathcal{S})} = -B_{k+1}^{-1} \mathbf{g}_{k+1}$ trovare una matrice unitaria $U_{k+1} \in \mathbb{R}^{n \times n}$ di bassa complessità, tale che, se $\mathcal{L}^{(k+1)} = sdU_{k+1}$, allora

$$\sigma \mathbf{d}_{k+1}^{(\mathcal{S})} = -(U_{k+1} diag((U_{k+1}^H B_{k+1} U_{k+1})_{ii}) U_{k+1}^H)^{-1} \mathbf{g}_{k+1} = -\left(\mathcal{L}_{B_{k+1}}^{(k+1)}\right)^{-1} \mathbf{g}_{k+1}.$$

$\Downarrow$

problema con alto grado di non linearità

$\Downarrow$

trovare strategie migliori

Scegliere:

$$\tilde{B}_{k+1} = U_{k+1} diag((V^H B_{k+1} V)_{ii}) U_{k+1}^H$$

$$= \mathcal{M}_{B_{k+1}}$$

$$\Downarrow$$

Metodo $\mathcal{L}^{(k)}QN$ ibrido

$\forall$ V unitaria:
$\mathcal{M} :=$ sd V;
$[\mathbf{z}_{k+1}]_i = (V^H B_{k+1} V)_{ii} = \lambda_i(\mathcal{M}_{B_{k+1}})$:

- $\det(B_{k+1}) \leq \det(\mathcal{M}_{B_{k+1}})$
- $\mathrm{tr}(B_{k+1}) \geq \mathrm{tr}(\mathcal{M}_{B_{k+1}})$

Problema 2: Partial Non-Linear Problem

Dato $\mathbf{g}_{k+1} \in \mathbb{R}^n$, $B_{k+1} = \psi_\phi(\tilde{B}_k, \mathbf{s}_k, \mathbf{y}_k)$ definita positiva, $\mathbf{d}_{k+1}^{(\mathcal{S})} = -B_{k+1}^{-1}\mathbf{g}_{k+1}$ e $\mathbf{z}_{k+1} \in \mathbb{R}^n$, $[\mathbf{z}_{k+1}]_i > 0$ per ogni $i = 1, \ldots, n$, trovare una matrice unitaria $U_{k+1} \in \mathbb{R}^{n \times n}$, tale che

$$\sigma \mathbf{d}_{k+1}^{(\mathcal{S})} = -(U_{k+1}(diag(\mathbf{z}_{k+1}))U_{k+1}^H)^{-1}\mathbf{g}_{k+1}.$$

Ad esempio: $(z_{k+1})_i = (U_k^H B_{k+1} U_k)_{ii}$

Convergenza della classe $\mathcal{L}^{(k)}$Broyden secante

Teorema di esistenza della soluzione

Dato $\mathbf{z}_{k+1} > 0$ esistono U_{k+1} e $\sigma > 0$ tali che

$$-[U_{k+1} diag(\mathbf{z}_{k+1}) U_{k+1}^H]^{-1}\mathbf{g}_{k+1} = \sigma \mathbf{d}_{k+1} \iff \frac{4z_m z_M}{(z_m+z_M)^2} \leq \frac{(-\mathbf{d}_{k+1}^T \mathbf{g}_{k+1})^2}{\|\mathbf{d}_{k+1}\|^2 \|\mathbf{g}_{k+1}\|^2} =: \gamma \qquad (c1)$$

dove $z_m := \min_{i=1,\dots,n}[\mathbf{z}_{k+1}]_i$; $\quad z_M := \max_{i=1,\dots,n}[\mathbf{z}_{k+1}]_i$

Cercare V tale che $[\mathbf{z}_{k+1}]_i = (V^H B_{k+1} V)_{ii}$ verifichi la condizione (c1) che equivale a

$$\frac{4\mu(\tilde{B}_{k+1})}{(1+\mu(\tilde{B}_{k+1}))^2} \leq \gamma \qquad (c2) \qquad \left(\mu(\tilde{B}_{k+1}) = \frac{z_M}{z_m}\right)$$

$[\mathbf{z}_{k+1}]_i = (U_k^H B_{k+1} U_k)_{ii} = \lambda_i(\mathcal{L}^{(k)}_{B_{k+1}})$, se (c1) OK $\quad \Rightarrow \quad V := U_k$;

altrimenti $\ldots \Longrightarrow$ CORREZIONE SPETTRALE

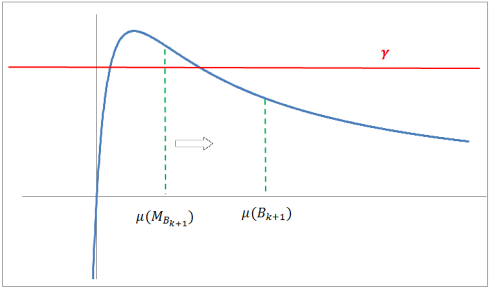
Correzione spettrale

Riassumendo: condizioni su $\mathbf{z}_{k+1}$:

$$\det(B_{k+1}) \leq \prod_{i=1}^{n} [\mathbf{z}_{k+1}]_i$$

$$\mathrm{tr}\,(B_{k+1}) \geq \sum_{i=1}^{n} [\mathbf{z}_{k+1}]_i$$

$$\frac{4 z_m z_M}{(z_m+z_M)^2} = \frac{4\,\frac{z_M}{z_m}}{\left(1+\frac{z_M}{z_m}\right)^2} \leq \gamma$$



Osservazioni

- noti $\mathbf{x}_m$= autovettore relativo all'autovalore minimo di B_{k+1}
  $\mathbf{x}_M$= autovettore relativo all'autovalore massimo di B_{k+1}
  $\rightarrow \exists V_{k+1}$ t. c. $V_{k+1}\mathbf{e}_m = \mathbf{x}_m$ e $V_{k+1}\mathbf{e}_M = \mathbf{x}_M \rightarrow \mathbf{z}_m = \lambda_m(B_{k+1})$ e $\mathbf{z}_M = \lambda_M(B_{k+1})$
  ma è troppo costoso calcolare $\mathbf{x}_m$ e $\mathbf{x}_M$
- se $\mathbf{z}_{k+1} = d(V_{k+1}^H B_{k+1} V_{k+1})$ e $\det B_{k+1} = \prod_{i=1}^{n}[\mathbf{z}_{k+1}]_i \Longrightarrow V_{k+1}$ sarebbe la matrice che diagonalizza $B_{k+1} \longrightarrow \det B_{k+1} < \prod_{i=1}^{n}[\mathbf{z}_{k+1}]_i$

Correzione spettrale

Sia $\mathbf{z} = d(V^H B V)$

Strategie per ottenere (c2):

Calolo di $\rho > 0$ tale che

$$\det(B) = (z_m - \rho) \prod_{i \neq m} [\mathbf{z}]_i$$
$$\Downarrow$$
$$\det(B) = \prod_{i=1}^{n} [\mathbf{z}']_i$$
$$\operatorname{tr}(B) > \sum_{i=1}^{n} [\mathbf{z}']_i$$
$$\text{Se } g\left(\frac{z_M}{z_m - \rho}\right) \leq \gamma$$
$$\Downarrow$$

Obiettivo raggiunto!!
Altrimenti....

$$\Longrightarrow$$

Calolo di $\delta > 0$ tale che

$$\det(B) = (z_m - \delta)(z_M + \delta) \prod_{i \neq m} [\mathbf{z}]_i$$
$$\Downarrow$$
$$\det(B) = \prod_{i=1}^{n} [\mathbf{z}']_i$$
$$\operatorname{tr}(B) = \sum_{i=1}^{n} [\mathbf{z}']_i$$
$$\text{Se } g\left(\frac{z_M + \delta}{z_m - \delta}\right) \leq \gamma$$
$$\Downarrow$$

Obiettivo raggiunto!!
Altrimenti....

Osservazioni

- $\mathbf{z}'$ ha ancora tutte le componenti positive
- $\rho < \delta$

Clusterizzazione

Obiettivo: modificare $\mathbf{z}$ ottenendo $\mathbf{z}^{'}$ tale che

$$\prod_{i \neq m,M} [\mathbf{z}]_i < \prod_{i \neq m,M} [\mathbf{z}^{'}]_i$$

in modo da ottenere

$\swarrow$ $\searrow$

un $\rho^{'} > 0$ tale che

$$\det(B) = (z_m - \rho^{'}) \prod_{i \neq m} [\mathbf{z}^{'}]_i$$

e che verifichi

$$\rho < \rho^{'}$$

un $\delta^{'} > 0$ tale che

$$\det(B) = (z_m - \delta^{'})(z_M + \delta^{'}) \prod_{i \neq m,M} [\mathbf{z}^{'}]_i$$

e che verifichi

$$\delta < \delta^{'}$$

Ma come ottenere questa disuguaglianza?

$\Downarrow$

Lemma

Siano a e b due quantità positive tali che $a < b$. Sia inoltre $0 < \Delta < (b - a)$. Allora

$$ab < (a + \Delta)(b - \Delta)$$

$\Downarrow$

CLUSTERIZZAZIONE

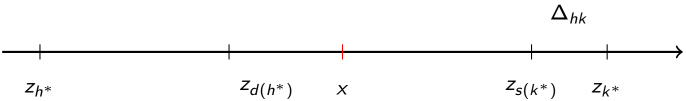
Cluster "ottimale"

$$\text{Sia } x := \frac{\sum_{i \neq m,M} [\mathbf{z}]_i}{n-2}$$

Cluster "ottimale"

$$\max_{\substack{z_h, z_k \\ z_h < x \\ z_k > x}} \left[\left(\prod_{i \neq m,M,h,k} [\mathbf{z}]_i \right) (z_h + \Delta_{hk})(z_k - \Delta_{hk}) - \prod_{i \neq m,M} [\mathbf{z}]_i \right]$$

- tale algoritmo fornisce la coppia di nodi ottimale
- viene mantenuto l'ordine rispetto al vettore originario
- potrebbe essere computazionalmente dispensioso



Cluster sulla media

Cluster dei nodi adiacenti alla media

Siano z_l e z_v i due nodi adiacenti alla media

$$\Delta := |x - z_v| \longrightarrow z_v \pm \Delta = x$$
$$\Downarrow$$
$$z_v z_l < (z_v \pm \Delta)(z_l \mp \Delta)$$

- all'ultima iterazione si ha $z_v \pm \Delta^{'} = x = z_l \mp \Delta^{'}$
- viene mantenuto l'ordine rispetto al vettore originario
- potrebbe essere meno efficace rispetto al cluster ottimale, ma meno dispensiosa

Cluster dei nodi più distanti dalla media

Siano z_l e z_v i due nodi più distanti dalla media

$$\Delta^{'} := |x - z_v| \longrightarrow z_v \pm \Delta^{'} = x$$
$$\Downarrow$$
$$z_v z_l < (z_v \pm \Delta^{'})(z_l \mp \Delta')$$

- $\Delta < \Delta^{'} \Rightarrow$ più veloce
- non viene mantenuto l'ordine rispetto al vettore originario

Altre strategie di correzione spettrale con cluster

Obiettivo: modificare il meno possibile il vettore $\mathbf{z}$ per discostarci il meno possibile dal modello dell'Hessiano che si sta costruendo.

$$\Downarrow$$

Oltre a variare il numero indispensabile di elementi di $\mathbf{z}$, possiamo mantenere la proprietà sul prodotto degli elementi interni:

$$\prod_{i=1}^{n}[\mathbf{z}]_i = (z_m - \rho)\prod_{i \neq m}[\mathbf{z}']_i$$

oppure

$$\prod_{i=1}^{n}[\mathbf{z}]_i = (z_m - \delta)(z_M + \delta)\prod_{i \neq m,M}[\mathbf{z}']_i$$

Infatti se $\det(B) \ll \prod_{i=1}^{n}[\mathbf{z}]_i \rightarrow \begin{cases} z_m - \rho \ll \lambda_m(B) \\ \text{oppure} \\ z_m - \delta \ll \lambda_m(B) \text{ e/o } z_M + \delta \gg \lambda_M(B) \end{cases}$