MANI-L04

Metodo Iterativi

$Ax = b$

$x^{(k+1)} = Bx^{(k)} + g$ → Schema dell'iterazione per $k \geq 0$

B e g non possono esser qualsiasi, serve che

$x = Bx + g$ → la consistenza.

Condizioni per garantire Convergenza dello schema

$e^{(k+1)} = Be^{(k)}$ → Legame che avevamo trovato

$e^{(k)} = Be^{(k-1)}$

$\|e^{k}\| = \|Be^{(k-1)}\|$

Norma di una matrice
↳ Ci sono vari metodi per normarle.

$A \in \mathbb{R}^{n \times n} \rightarrow \|A\|_2 \xrightarrow{Def.} \sqrt{\underbrace{\lambda_{max}(A^T A)}}$

matrice simmetrica positiva

L'autovalore massimo della matrice $A^T A$

→ Norma 2, norma spettrale
↳ Gli autovalori di una matrice sono il suo spettro

Se A è simmetrica $(A^T = A)$ $\quad \|A\|_2 = \sqrt{\lambda_{MAX}(A^2)} = \sqrt{[\lambda_{MAX}(A)]^2} = \lambda_{MAX}(A)$

<u>Proprietà di Compatibilità</u>

Per il livello dobbiamo accettare

Raggio spettrale di B

$$\|e^{(k)}\| = \|Be^{(k-1)}\| \le \|B\|_2 \|e^{(k-1)}\| \simeq \rho(B)\|e^{(k-1)}\|$$

In generale non è vero, solo se sono scelte opportunamente le norme, la norma di un vettore è compatibile con la norma 2 quindi si può fare.

Massimo degli autovettori in modulo

Matlab: max (abs (eig (B)))

Possiamo fare quando sono compatibili, che la norma 2 e quella euclidea lo sono

$$\|e^{(k)}\| \le \rho(B)\|e^{(k-1)}\| \le [\rho(B)]^2 \|e^{(k-2)}\| \le [\rho(B)]^3 e^{(k-3)} \le \dots$$

$$\|e^{(k-1)}\| \le \rho(B)\|e^{(k-2)}\|$$

$$\|e^{(k-2)}\| \le \rho(B)\|e^{(k-3)}\|$$

$$\le \dots \le [\rho(B)]^k \|e^{(0)}\|$$

$$\|e^{(k)}\| \leq [\rho(B)]^k \|e^{(0)}\|$$

Se non si ha consistenza tutto questo non si fa

Condizione per la convergenza sufficiente e necessaria : $\rho(B) < 1$

→ Anche un errore grande iniziale può migliorare.

Teorema

Si consideri lo schema $x^{(k+1)} = Bx^{(k)} + g$ e si supponga che sia consistente.

Allora lo schema è convergente $\underbrace{\forall x^{(0)} \in \mathbb{R}^n}_{\text{Per libertà possiamo prendere il vettore nullo}} \iff \rho(B) < 1$

Più $\rho(B)$ è piccolo, più la convergenza è rapida.

B_1, g_1	B_2, g_2	B_3, g_3
$\rho(B_1) = 0,7$	$\underbrace{\rho(B_2) = 0,01}_{\text{Si sceglie questo perché è più veloce.}}$	$\rho(B) = 1,00005$

$$\|e^{(k)}\| \leq [\rho(B)]^k \|e^{(0)}\|$$ → Anche utile come condizione di arresto in base a quanto è

abbattuto

? k_{MIN}

$$\boxed{\frac{\|e^{(k,min)}\|}{\|e^{(0)}\|} \le [\rho(B)]^{k_{min}} < TOL = 10^{-5}}$$

Per trovare k_{MIN} si può fare $k_{MIN} = \log_{\rho(B)} TOL$

Abbiamo visto l'errore ora guardiamo specifici schemi iterativi.

Costruiamo un'intera famiglia di schemi

Metodi di Richardson

$Ax = b$

Prendiamo arbitrariamente

Utilizzato per velocizzare convergenza

$\alpha_k Ax = \alpha_k b$ $\quad \alpha_k \in \mathbb{R}$ Parametro di accelerazione

Riscrivo/Manipolazioni algebriche

Lista di richieste su P:

$\alpha_k A = P - P + \alpha_k A$ $\quad P \in \mathbb{R}^{n \times n}$ - invertibile - facile $\rightsquigarrow$ spiegato dopo

Precondizionatore

$Px - (P - \alpha_k A)x = \alpha_k b$

$$\underset{k+1}{Px} = (P - \alpha_k A)\underset{k}{x} + \alpha_k b$$

$\longrightarrow$ Scelta molto arbitraria fatta da Richardson

Inserendo gli indici:

$$Px^{(k+1)} = (P - \alpha_k A)x^{(k)} + \alpha_k b$$

$$Px^{(k+1)} = Px^{(k)} + \alpha_k b - \alpha_k A x^{(k)}$$

Multiplicando per P^{-1} → Avevamo già chiesto che fosse invertibile

$$x^{(k+1)} = x^{(k)} + \alpha_k P^{-1}\underbrace{(b - Ax^{(k)})}_{r^{(k)}}$$

$r^{(k)}$ → Resto di iterazione

Quello che rimane quando togliamo la soluzione e la nostra approssimazione

$$\boxed{x^{(k+1)} = x^{(k)} + \alpha_k P^{-1} r^{(k)}}$$

→ Forma più nota del metodo di Richardson

$P^{-1} r^{(k)} = z^{(k)}$ → Residuo precondizionato

$$x^{(k+1)} = x^{(k)} + \alpha_k z^{(k)} \quad k \geq 0$$

→ Iterazione è iterazione prima più una correzione.

Schemi di Richardson

Stazionario se $\alpha_k = \text{cost} \ \forall k$

Dinamico se $\alpha_k \neq \text{cost} \ \forall k$

→ Dinamico si adatta ad ogni

Invece di trovare P^{-1} → alto costo computazionale, quindi risolviamo $Pz^{(k)} = r^{(k)}$

Dobbiamo lo stesso calcolare quindi prendiamo una matrice furba, prendiamo P allora facile e triangolare

$Pz^{(k)} = r^{(k)}$ deve esser risolto facilmente

Usiamo il metodo di fattorizzazione e poi calcolo

15:20 controlla se mancato qualcosa

Riscriviamo in forma:

$$x^{(k+1)} = Bx^{(k)} + g$$

Partendo da qui:

$$x^{(k+1)} = x^{(k)} - \alpha_k P^{-1}Ax^{(k)} + \alpha_k P^{-1}b$$

$$x^{(k+1)} = \underbrace{\left(I - \alpha_k P^{-1}A\right)}_{B_{\alpha_k}} x^{(k)} + \underbrace{\alpha_k P^{-1}b}_{g_{\alpha_k}}$$

Perché dipende da α_k

Ci sono 2 casi di schemi noti, guarderemo a lato o se si vuole nelle registrazioni dell'anno scorso

Richardson stazionario (Pe α definiscono lo schema)

Jacobi

Gauss-Seidel

!

$$P = D = \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix}$$

$\alpha = 1$

$$P = D - E = \begin{bmatrix} a_{11} & & & 0 \\ a_{21} & \ddots & & \\ \vdots & & \ddots & \\ a_{n1} & \dots & \dots & a_{nn} \end{bmatrix}$$

Porzione triangolare inferiore di A.

$\alpha = 1$

Cosa facciamo a lab:

- Oltre usare $\rho(B)$ + consistenza
- Esistono condizioni sufficienti per garantire convergenza
  - Se A è parte di matrici speciali
- Jacobi vs. Gauss-Seidel

$\alpha = 1 \Rightarrow$ l'accelerazione è stata spenta

Cerchiamo se c'è una regola per scegliere P e α

__Proposizione__: siano A e P sdp. (simmetrica definita positiva)

Allora RS (Richardson Stazionario) converge $\forall x^{(0)} \in \mathbb{R}^n$

$$\Longleftrightarrow 0 < \alpha < \frac{2}{\lambda_{MAX}(P^{-1}A)}.$$

reale per sdp, non serve modulo

quella che massimizza la velocità di convergenza

La scelta ottimale per α è $$\alpha_{opt} = \frac{2}{\lambda_{MAX}(P^{-1}A) + \lambda_{min}(P^{-1}A)}$$

Inoltre $\|e^{(k)}\|_A \leq \underbrace{\left(\frac{k(P^{-1}A)-1}{k(P^{-1}A)+1}\right)^k}_{}\|e^{(0)}\|_A$

Rispetto una norma che non abbiamo ancora visto

Fattore di convergenza,

$k \to$ numero di condizionamento

Norma $A \to$ Per FEM è norma in energia

$\hookrightarrow$ se $w \in \mathbb{R}^n$ $\quad \|w\|_A = \sqrt{w^T A w}$

$\hookrightarrow A = I$ si ritorna alla norma euclidea,

perchè $w^T w$ è la somma dei prodotti.

mettendo A nel problema tiene traccia del problema che stiamo risolvendo.

Nota che si è dimenticato:

Quando abbiamo ricavato

$$P x = P x + \alpha_k b - \alpha_k A x$$

$$P x^{(k+1)} = P x^{(k)} + \alpha_k b - \alpha_k A x^{(k)}$$

$\to$ visto che siamo iniziali sopra implica che ogni sistema di richardson è consistente per costruzione.

$\to$ Poi abbiamo scelto arbitrariamente quali x fossero $x^{(k+1)}$ e $x^{(k)}$.

$\hookrightarrow$ Se ci viene chiesto possiamo dire che sono consistenti per costruzione

16:06

<u>Dimostrazione:</u> Prima parte

Dobbiamo verificare che $\rho(B_\alpha) < 1$

Per abbiamo detto che è convergente se è consistente con $\rho(B_\alpha) < 1$. È consistente per costruzione allora

$B_\alpha = I - \alpha P^{-1} A$

$\lambda_i \in \mathbb{R}^n$ → autovalori di $P^{-1}A$ → lo stesso reali e positivi

possiamo anche dire $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$

μ_i autovalori di B_α

Autovalore di I

$\mu_i = 1 - \alpha \lambda_i$

Dobbiamo andare a chiedere che $|1 - \alpha \lambda_i| < 1 \ \forall i$

$\Rightarrow \rho(B_\alpha) < 1$

Messi insieme si verifica che $\rho(B_\alpha) < 1$

$\Rightarrow -1 < 1 - \alpha \lambda_i < 1$

$\Rightarrow \alpha > 0$ perché sappiamo che $\lambda_i > 0$ per P e A sdp

$\alpha \lambda_i < 2$

$\alpha < \frac{2}{\lambda_i}$, affinché valga per ogni λ_i, prendo la più restrittivo, cioè λ_{max} $\alpha < \frac{2}{\lambda_{MAX}(P^{-1}A)}$

Seconda parte

Trovare tra

$\alpha_{opt} \Rightarrow 0 < \alpha < \frac{2}{\lambda_{max}(P^{-1}A)}$ α che minimizza il raggio spettra

$\Rightarrow$ minimizza il max autovalore

Supponiamo (esempio piccolo)

$\lambda_1, \lambda_2, \lambda_3$ di $P^{-1}A$ con $\lambda_1 \geq \lambda_2 \geq \lambda_3$

$\frac{1}{\lambda_1} < \frac{1}{\lambda_2} < \frac{1}{\lambda_3}$

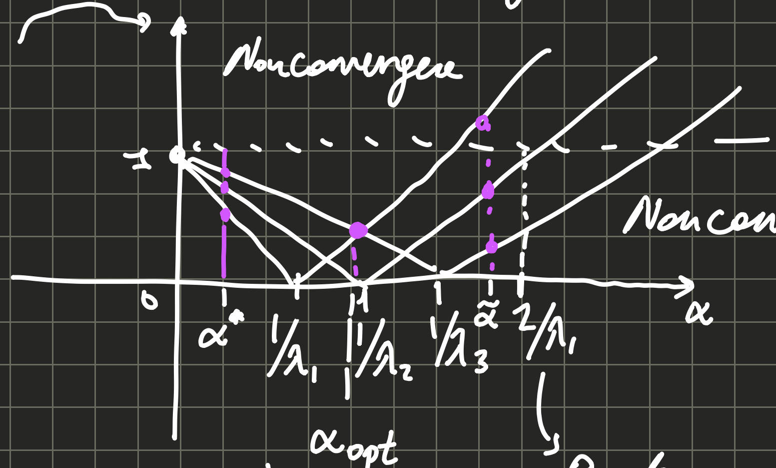
$P^{-1}A$

$|\mu_1| = |1-\alpha\lambda_1|$ $|\mu_2| = |1-\alpha\lambda_2|$ $|\mu_3| = |1-\alpha\lambda_3|$

B_α

Mettiamo in grafico i valori α-dipendenti

$|1-\alpha\lambda_i|$



Punto massimo per avere la convergenza

Per ogni α che scelgo quali sono gli autovalori

Dobbiamo trovare α dove il valore massimo è il più basso possibile.

↳ Dove il massimo passa di $\frac{1}{\lambda_3}$ a $\frac{1}{\lambda_1}$

↳ Cerchiamo μ_i ~~massimo~~ minimo tale per minimizzare $\rho(B)$

$$1 - \alpha \lambda_3 = \alpha \lambda_1 - 1$$

$$2 = \alpha(\lambda_1 + \lambda_3)$$

$$\alpha_{opt} = \frac{2}{\lambda_1 + \lambda_3}$$ → In generale λ_i sono λ_{MAX} e λ_3 un λ_{MIN}

Generalmente → $\alpha_{opt} = \frac{2}{\lambda_{MAX} + \lambda_{MIN}}$

$$\rho(B_{\alpha_{opt}}) = 1 - \alpha_{opt} \lambda_{MIN} = 1 - \frac{2}{\lambda_{MAX} + \lambda_{MIN}} \cdot \lambda_{MIN}$$

$$= \frac{\lambda_{MAX} + \lambda_{MIN} - 2\lambda_{MIN}}{\lambda_{MAX} + \lambda_{MIN}} = \frac{\lambda_{MAX} - \lambda_{MIN}}{\lambda_{MAX} + \lambda_{MIN}}$$

→ $\rho(B)$ è l'autovalore, μ_i, massimo di B, per avere μ_i massimo, sapendo che $\mu_{i,opt} = 1 - \alpha_{opt} \lambda_i$, dobbiamo ridurre 1 del minimo possibile, cioè avere $\lambda_i = \lambda_{MIN}$, di modo tale che μ_{max} sia il massimo possibile.

Ricerca di α_{opt} dinamico

Proposizione:

Siano A e P sdp. Allora RD (Richardson Dinamico)

<u>converge</u> $\forall x^{(0)} \in \mathbb{R}^n \xLeftrightarrow{se} \alpha_{k,opt} = \dfrac{[z^{(k)}]^T r^{(k)}}{[z^{(k)}]^T A z^{(k)}} \quad k \geq 0$

per la convergenza

↳ Un singolo valore invece di un intervallo

ricavato con il metodo del gradiente precondizionato

Inoltre $\|e^{(k)}\|_A \leq \left(\dfrac{k(P^{-1}A)-1}{k(P^{-1}A)+1}\right)^k \|e^{(0)}\|_A$

↳ se $P = I \rightarrow P$ è detto spento

↳ in questo caso per ricavare $\alpha_{k,opt}$ usiamo il metodo del gradiente:

che da: $\alpha_{k,opt} = \dfrac{[r^{(k)}]^T r^{(k)}}{[r^{(k)}]^T A r^{(k)}}$

<u>Dimostrazione</u> (metodo del gradiente non precondizionato)

$$x^{(k+1)} = x^{(k)} + \alpha_k r^{(k)}$$

$$\alpha_k = \frac{[r^{(k)}]^T r^{(k)}}{[r^{(k)}]^T A r^{(k)}}$$

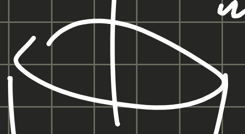
Ricordiamo che A è sdp

$x = [x_1, x_2]^T$

Risolvere $Ax = b \xLeftrightarrow{\text{se } A \text{ è sdp}}$ minimizzare $Q(x) = \overbrace{\underbrace{\frac{1}{2}x^T A x - x^T b}_{\text{Geometricamente è}}}^{\text{Forma Quadratica}}$

$$\nabla Q(x) = Ax - b$$

$$\nabla Q(x) = 0 \Rightarrow Ax - b = 0 \Rightarrow Ax = b$$



$\boxed{16:47}$

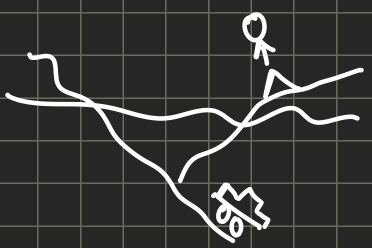
Scelgo la più ripida

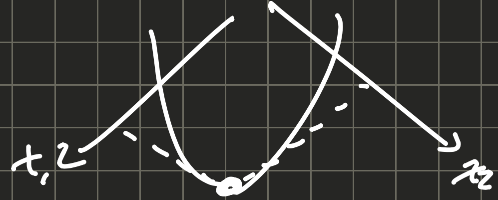
$$x^{(k+1)} = x^{(k)} + \gamma_k \, d^{(k)}$$

Direzione di spostamento ($d^{(k)}$)

Magnitudine di spostamento (γ_k)

ad ogni iterazioni ricalcoliamo il gradiente e la direzione per decidere il prossimo passo.





$$d^{(k)} = -\nabla Q\left(x^{(k)}\right) = b - Ax^{(k)} = r^{(k)}$$

↳ La direzione tattica è quella del residuo

$$Q\left(x^{(k)} + \gamma_k r^{(k)}\right) = \tilde{Q}(\gamma_k)$$ → dobbiamo γ_k dove Q inizia a tornare in su

$$\frac{d\tilde{Q}}{d\gamma_k} = 0$$

$$\underbrace{\frac{1}{2}\left(x^{(k)} + \gamma_k r^{(k)}\right)^T A \left(x^{(k)} + \gamma_k r^{(k)}\right) - \left(x^{(k)} + \gamma_k r^{(k)}\right)^T b}_{Q\left(x^{(k)} + \gamma_k r^{(k)}\right)} = \tilde{Q}(\gamma_k)$$

$$\frac{d\tilde{Q}}{d\gamma_k} = \left[r^{(k)}\right]^T A x^{(k)} + \gamma_k \left[r^{(k)}\right]^T A r^{(k)} - \left[r^{(k)}\right]^T b = 0$$

γ_k per $\frac{d\tilde{Q}}{d\gamma_k} = 0$

$$\gamma_k = \frac{\left[r^{(k)}\right]^T b - \left[r^{(k)}\right]^T A x^{(k)}}{\left[r^{(k)}\right]^T A r^{(k)}} = \frac{\left[r^{(k)}\right]^T \overbrace{\left(b - Ax^{(k)}\right)}^{r^{(k)}}}{\left[r^{(k)}\right]^T A r^{(k)}} = \frac{\left[r^{(k)}\right]^T r^{(k)}}{\left[r^{(k)}\right]^T A r^{(k)}}$$

Abbiamo trovato α_k usando il metodo di discesa più ripida,

↳ questo α_k è l'unico che garantisce la convergenza $\forall x^{(0)}$ nel caso dinamico

$Q(x)$ è quadratico perchè

$\underset{r \times c}{x^T} \underset{r \times c}{A x}$ → Gli elementi di x si multiplicano 2 volte con i moltiplicatori di A, rendendola una equazione quadratica, bx^T cambia la sua posizione nel piano.