

Lanczos process

Sea $A = A^T$ $n \times n$. Consideriamo la fattorizzazione
 parziale:

$$A Q_e = Q_{e+1} T_{e+1,e}, \quad e < n$$

dove $Q_e = [q_1 \ q_2 \ \dots \ q_e]$, $Q_{e+1} = [Q_e \ q_{e+1}]$ $q_i^T q_j = \delta_{ij}$

$$T_{e+1,e} = \begin{bmatrix} \alpha_1 & \beta_1 & & & & \\ & \beta_1 & \alpha_2 & & & \\ & & \ddots & \ddots & & \\ & & & \beta_{e-1} & \alpha_e & \\ & & & & \beta_{e-1} & \alpha_e \\ & & & & & \beta_e \end{bmatrix}, \quad \beta_i > 0$$

Costruzione: Sea b un vettore iniziale, $\beta_0 = \|b\|$, $q_1 = \frac{b}{\beta_0}$
 consideriamo la prima colonna del prodotto

$$A q_1 = \alpha_1 q_1 + \beta_1 q_2$$

$$A [q_1 \ q_2 \ \dots \ q_e] = [q_1 \ q_2 \ \dots \ q_e \ p_{e+1}] \begin{bmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & & & \\ & \beta_2 & \ddots & & \\ & & \ddots & \beta_{e-1} & \alpha_e \\ & & & \beta_{e-1} & \alpha_e \\ & & & & \beta_e \end{bmatrix}$$

①

$$A q_1 = \alpha_1 q_1 + \underbrace{\beta_1 q_2}_{\tilde{q}_2} \rightarrow \tilde{q}_2 = A q_1 - \alpha_1 q_1$$

$$q_1^T \tilde{q}_2 = 0 = q_1^T A q_1 - \alpha_1 \underbrace{q_1^T q_1}_{=1} \Rightarrow \alpha_1 = q_1^T A q_1$$

$$\beta_1 = \|\tilde{q}_2\|, \quad q_2 = \frac{\tilde{q}_2}{\beta_1}$$

② passiamo alla seconda colonna

$$A q_2 = \beta_1 q_1 + \alpha_2 q_2 + \underbrace{\beta_2 q_3}_{\tilde{q}_3}$$

$$\tilde{q}_3 = A q_2 - \alpha_2 q_2 - \beta_1 q_1 \quad \text{ricorriamo a 3 termini!}$$

$$\beta_1, q_1, q_2 \text{ noti} \quad q_2^T \tilde{q}_3 = 0 = q_2^T A q_2 - \alpha_2 \Rightarrow \underline{\alpha_2 = q_2^T A q_2}$$

$$\underline{\beta_2 = \|\tilde{q}_3\|} \quad \underline{q_3 = \frac{\tilde{q}_3}{\beta_2}}$$

② consideriamo il passo generico

$$A q_e = \beta_{e-1} q_{e-1} + \alpha_e q_e + \underbrace{\beta_e q_{e+1}}_{\tilde{q}_{e+1}}$$

$$\tilde{q}_{e+1} = A q_e - \alpha_e q_e - \beta_{e-1} q_{e-1} \quad \beta_{e-1}, q_{e-1}, q_e \text{ noti}$$

$$q_e^T \tilde{q}_{e+1} = 0 = q_e^T A q_e - \alpha_e \Rightarrow \underline{\alpha_e = q_e^T A q_e}$$

$$\underline{\beta_e = \|\tilde{q}_{e+1}\|} \quad \underline{q_{e+1} = \frac{\tilde{q}_{e+1}}{\beta_e}}$$

È facile vedere che i q_i sono tutti ortogonali.

1. $q_e^T \tilde{q}_{e+1} = 0$ per costruzione.

2. $\forall i \leq e-1, q_i^T \tilde{q}_{e+1} = q_i^T A q_e = q_i^T (\beta_{e-1} q_{e-1} + \alpha_e q_e + \beta_e q_{e+1}) = 0$

3. $q_{e-1}^T \tilde{q}_{e+1} = q_{e-1}^T A q_e - \beta_{e-1} = 0$ because

$$\beta_{e-1} q_e^T q_e = q_e^T A q_{e-1} - \underbrace{\alpha_{e-1} q_e^T q_{e-1}}_{=0} = \beta_{e-2} \underbrace{q_e^T q_{e-2}}_{=0} = q_{e-1}^T A q_e \quad (A = A^T)$$

Se non ci sono breakdown ($\beta_l \neq 0 \quad l=1, \dots, n$) si arriva a

$$A Q_n = Q_n T_{n,n}$$

infatti \tilde{q}_{n+1} non potrebbe essere ortogonale a $q_i, i=1, \dots, n$ e quindi $\tilde{q}_{n+1} = 0, \beta_n = 0$.

All'ultimo passo fornisce una trasformazione di similitudine che triangolarizza la matrice A :

$$Q_n^T A Q_n = T_{n,n} \quad \text{e} \quad A = Q_n T_{n,n} Q_n^T$$

Se il processo viene arrestato a $l < n$

$$A Q_l = Q_{l+1} T_{l+1,l} \Rightarrow Q_{l+1}^T A Q_l = T_{l+1,l}$$

e anche $T_{l+1,l} = Q_l^T A Q_l$.

Questa è la proiezione di A in $\text{Ke}(A, b)$

Infatti $\text{span} \{q_1, q_2, \dots, q_l\} = \text{span} \{b, Ab, A^2b, \dots, A^{l-1}b\}$

e $x \in \text{Ke} \Rightarrow z = Ax = A Q_l y$ con $y \in \mathbb{R}^l$

Se proiettiamo z in Ke : $Q_l^T z = Q_l^T A Q_l y = T_{l+1,l} y$

Tuttavia $A \neq Q_{l+1} T_{l+1,l} Q_l^T$!

Il secondo membro fornisce una approssimazione della matrice A , utile per approssimare le soluzioni di problemi di grandi dimensioni.

Applicazioni

Sistemi lineari $Ax = b$

metodo iterativo $x^{(0)} = \arg \min_{x \in K_e} \|Ax - b\|_2$

$$x \in K_e \Leftrightarrow x = Qey, \quad y \in \mathbb{R}^e$$

$$\begin{aligned} \text{Quindi} \quad \|Ax - b\|^2 &= \|AQey - b\|^2 = \|Q_{e+1}^T e_{e+1} y - b\|^2 = \\ &= \|\nabla_{e+1} y - \|b\| e_1\|^2 \end{aligned} \quad (*)$$

Infatti:

$$b = \|b\| \cdot q_1 = \|b\| \cdot Q_{e+1} e_1$$

e

$$\|Q_{e+1} z\|^2 = z^T Q_{e+1}^T Q_{e+1} z = z^T z = \|z\|^2$$

Il problema (*) è piccolo e può essere risolto facilmente.

~~Algebra~~

Autovalori

Th. Fisher

$$\lambda_i = \min_{\substack{\dim(W) = \\ = n-i+1}} \max_{\substack{x \in W \\ \|x\|=1}} x^T A x$$

Consideriamo il Th. di Fisher in K_e :

Quoziente di Rayleigh: $y \in \mathbb{R}^e$

$$y^T T y = y^T Q_e^T A Q_e y = \underbrace{x^T A x}_{x \in K_e} \quad x \in K_e$$

$$y \in W \subset \mathbb{R}^e \Rightarrow x \in \tilde{W} \subset K_e \subset \mathbb{R}^n, \quad \dim(W) = \dim(\tilde{W}) = e - i + 1$$

Il fatto che $K_e \rightarrow \mathbb{R}^n$ (se non ci sono breakdown) implica che gli autovalori estremali di T_e convergono a quelli di A , perché il $\max_{y \in W} y^T T_e y$ viene calcolato in spazi di dimensione crescente.

Inoltre, se $T_e y = \lambda y$ allora

$$Q_e^T A Q_e y = \lambda y \rightarrow Q_e^T (A x) = Q_e^T (\lambda x)$$

x è un autovettore $\sin K_e$ (se proiettato in K_e)

Quindi x converge a un autovettore di A , al crescere di e .

In questo caso si inizializza con b vettore casuale.

Funzioni matriciali

$$F(A) = F(U\Lambda U^T) = U F(\Lambda) U^T$$

se $A = U\Lambda U^T$ è la fattorizzazione spettrale e $A = A^T$.

Allora

$$\begin{aligned} F(A) &\approx F(Q_e T_{e,e} Q_e^T) = F(Q_e U_e \Lambda_e U_e^T Q_e^T) = \\ &= (Q_e U_e) F(\Lambda_e) (Q_e U_e)^T \end{aligned}$$

Se $T_{e,e} = U_e \Lambda_e U_e^T$ è la fatt. spettrale.

$F(\Lambda_e)$ è facile da calcolare, così come la fatt. perché l è piccolo risp. a n ($l \ll n$)

Breakdown

Se al passo e $\beta_e = 0$ non si può proseguire.

Significa che \tilde{q}_{e+1} non può essere creato ortogonale ai precedenti q_i , e quindi che $\mathbb{R}q_e$

$$Aq_e \in \text{span}\{q_1, \dots, q_e\} = K_e$$

$$\text{cioè: } K_e = K_{e+1}$$

Se avviene:

1. $Ax=b$: $x \in K_e \rightarrow$ soluzione esatta

2. avel: ho trovato un sottospazio invariante

$\lambda_i(T_{e,e}) = \lambda_i(A) \rightarrow$ soluzione esatta, ma non posso approssimare gli altri autovalori.

Se serve ripartire, serve $b \perp K_e$ (Gram-Schmidt)
L \rightarrow RESTART

In realtà non si verifica $\beta_e = 0$, ma $|\beta_e| < \epsilon$.

Algoritmo

C'è solo 1 prodotto matrice-vettore per iterazione.

Se non serve memorizzare Q_k , l'occupazione di memoria è minima.

Si verifica spesso una perdita di ortogonalità nei q_i :
serve la RIORTOGONALIZZAZIONE.

Al passo k si riortogonalizza q_i rispetto ai precedenti.

Si usa CGS o NGS.

Per un RESTART serve Q_k !

Nei sistemi lineari, per risparmiare memoria
quando si usa la RIORTOGONALIZZAZIONE,

si fa un restart ogni \bar{k} iterazioni ponendo

$$b \rightarrow b - Ax^{(k)}$$