Equazioni di Sylvester

Note Title 2023-03-02

Condizionamento

$$AX - XB = C \qquad A \in \mathbb{C}^{m\times m} \quad B \in \mathbb{C}^{n\times n} \quad X, C \in \mathbb{C}^{m\times n}$$

$$\Updownarrow$$

$$\underbrace{\left(I \otimes A - B^T \otimes I\right)}_{M \in \mathbb{C}^{mn \times mn}} \operatorname{vec} X = \operatorname{vec} C$$

Possiamo studiarlo come il condizionamento di un sist. lineare

ad es. se $\tilde{X}$ risolve $M \operatorname{vec} \tilde{X} = \operatorname{vec} \tilde{C}$

$$\frac{\|\tilde{X} - X\|_F}{\|X\|_F} \leq K(M) \cdot \frac{\|\tilde{C} - C\|_F}{\|C\|_F}$$

$$\uparrow$$

$$\|\operatorname{vec} X\|_2 = \|X\|_F \qquad K(M) = \|M\|_2 \, \|M^{-1}\|_2 = \frac{\sigma_{max}(M)}{\sigma_{min}(M)}.$$

$$\sigma_{max}(M) = \|M\|_2 = \| I \otimes A - B^T \otimes I \|_2 \leq \| I \otimes A\|_2 + \|B^T \otimes I\|_2 = \|A\|_2 + \|B\|_2$$

$$\sigma_{min}(M) = \min_{z \in \mathbb{C}^{mn}} \frac{\|Mz\|_2}{\|z\|_2} = \min_{Z \in \mathbb{C}^{m\times n}} \frac{\|AZ - ZB\|_F}{\|Z\|_F} =: \operatorname{sep}(A, B)$$

$$\sigma_{min}(M) \leq \left| \min_{\lambda \in \Lambda(M)} |\lambda| \right| = \min |\lambda_i - \mu_j| \qquad \begin{array}{l} \lambda_i \in \Lambda(A) \\ \mu_j \in \Lambda(B) \end{array}$$

Se A e B hanno un autovalore molto vicino, allora $|\lambda_i - \mu_j|$ è piccolo

$\sigma_{min}(M)$ piccolo $\Rightarrow$ mal condizionata

Se A, B sono matrici normali ($AA^* = A^*A$, $\Leftrightarrow$ A si diagonalizza con una matrice ortogonale $A = QDQ^*$), $\sigma_{min}(M) = \min_{\lambda \in \Lambda(M)} |\lambda|$ è un'uguaglianza.

Stabilità dell'algoritmo di Bartels-Stewart

B-S: trasformazioni ortogonali + sostituzione all'indietro su un sistema $mn \times mn$.

La sostituzione all'indietro è backward stable: il valore $\tilde{x}$ calcolato con sost. all'indietro su un sist. triangolare $Ux = c$ soddisfa

$$(U + \Delta_U)\tilde{x} = c + \Delta_c \qquad \frac{\|\Delta_U\|}{\|U\|} = O(u) \quad \frac{\|\Delta_c\|}{\|c\|} = O(u)$$

$\Rightarrow$ la matrice $\tilde{X}$ calcolata con B-S in aritmetica di macchina soddisfa

$$(M + \Delta_M)\,\mathrm{vec}\,\tilde{X} = \mathrm{vec}\,(C + \Delta_C)$$

Da questo, bound sul residuo:

$$\underbrace{\|M\,\mathrm{vec}\,\tilde{X} - \mathrm{vec}\,C\|_2}_{=\ \|\mathrm{vec}\,(A\tilde{X} - \tilde{X}B - C)\|_2} \leq \|\mathrm{vec}\,\Delta_C\|_2 + \|\Delta_M \cdot \mathrm{vec}\,\tilde{X}\|_2$$

$$\|A\tilde{X} - \tilde{X}B - C\|_F \lesssim \|\Delta_C\|_F + \|\Delta_M\| \cdot \|\tilde{X}\|_F \lesssim u\left(\|C\|_F + \|M\| \cdot \|\tilde{X}\|_F\right)$$

$$\lesssim u\left(\|C\|_F + (\|A\| + \|B\|)\|\tilde{X}\|_F\right)$$

⚠ $\|\tilde{X}\|_F$ può essere molto grande anche se $\|A\|, \|B\|, \|C\|$ sono moderate

⚠ L'algoritmo <u>non</u> è stabile all'indietro nel senso che la $\tilde{X}$ calcolata <u>non</u> risolve $(A + \Delta_A)\tilde{X} - \tilde{X}(B + \Delta_B) = C + \Delta_C$

$\Delta_M \neq 1 \otimes \Delta_A - \Delta_B^T \otimes 1$ <u>non</u> è vero in generale

$(mn)^2$ gradi di libertà $\qquad$ m^2+n^2 gradi di libertà.

---

se X risolve Syl.

$$\begin{bmatrix} 1 & -X \\ 0 & 1 \end{bmatrix}\begin{bmatrix} A & -C \\ 0 & B \end{bmatrix}\begin{bmatrix} 1 & X \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A & -C-XB \\ 0 & B \end{bmatrix}\begin{bmatrix} 1 & X \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A & \overset{\overset{0}{\shortparallel}}{AX-XB-C} \\ 0 & B \end{bmatrix}$$

$\|X\|_F \sim K\left(\begin{bmatrix} 1 & X \\ 0 & 1 \end{bmatrix}\right)$ misura quanto mal condizionata è questa trasformazione da triangolare a blocchi a diagonale a blocchi

<u>ES</u> caso scalare:

$$\begin{bmatrix} \lambda & 1 \\ 0 & \mu \end{bmatrix} \sim \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix} \qquad \text{se } \lambda \neq \mu$$

$$\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \not\sim \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \qquad \text{(forma di Jordan)}$$

$$\begin{bmatrix} 1 & -x \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \lambda & 1 \\ 0 & \mu \end{bmatrix}\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix} \qquad \text{se } x \text{ risolve } \lambda x - x\mu = -1$$

$$x = \frac{1}{\mu - \lambda}$$

<u>Sottospazi invarianti</u>

Definizione: data $M \in \mathbb{C}^{n\times n}$, $\mathcal{U} \subseteq \mathbb{C}^n$ sottosp. vettoriale si dice <u>sottospazio invariante</u> per M se $M\mathcal{U} \subseteq \mathcal{U}$

Se $U_1 \in \mathbb{C}^{n\times k}$ le cui colonne formano una base per $\mathcal{U}$,

e $[U_1\ U_2] = U$ è il completamento a una matrice invertibile allora MU_1 ha colonne in $\mathcal{U}$ e quindi $MU_1 = U_1 A$, $A \in \mathbb{C}^{k\times k}$

e $\quad MU = U \begin{bmatrix} A & B \\ \boxed{0} & C \end{bmatrix} \qquad [\overbrace{MU_1}^{k} \ \ \overbrace{MU_2}^{n-k}] = \left[U\begin{bmatrix} A \\ 0 \end{bmatrix} \ \ U\begin{bmatrix} B \\ C \end{bmatrix} \right]$

cioè, se $\mathcal{U}$ è un sottosp. invariante allora $U^{-1}MU = \begin{bmatrix} * & * \\ 0 & * \end{bmatrix}$

Esempi di sott. invarianti:

- $\text{span}\{v\}$, dove v è un autovettore di M

  $M(\alpha v) = \alpha(\lambda v)$

- $\text{span}\{v_1, v_2, \dots v_k\}$, dove i v_i sono autovettori di M, anche con autovalori diversi

$$M(\underbrace{v_1\alpha_1 + v_2\alpha_2 + \dots + v_k\alpha_k}_{\text{generico elemento di } \mathcal{U}}) = v_1(\lambda_1\alpha_1) + v_2(\lambda_2\alpha_2) + \dots + v_k(\lambda_k\alpha_k) \in \mathcal{U}$$

- Data $M = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$ $\quad e_1$ autovettore $\Rightarrow \text{span}\{e_1\}$ è invariante

Anche $\text{span}\{e_1, e_2\}$ è invariante: difatti, $Me_2 = \lambda e_2 + e_1$
$Me_3 = \lambda e_3 + e_2$

In generale, data una catena di Jordan $\{v_1, v_2 \dots v_m\}$,

$$Mv_1 = \lambda v_1 \qquad Mv_2 = \lambda v_2 + v_1 \qquad Mv_3 = \lambda v_3 + v_2 \ , \dots$$

i suoi primi k elementi (per ogni $k \in 1, \dots m$) formano un sottosp. invariante.

Teo (caratterizzazione dei sottosp. invarianti)

Se U è un sottosp. invariante per M, cioè $MU \subseteq U$, $MU_1 = U_1 A$
allora posso scrivere $U =$ span di pezzi iniziali di un po' di catene di Jordan di M

In particolare, $\Lambda(A) \subseteq \Lambda(M)$ con molteplicità.

Dim: Prendiamo un insieme completo di catene di Jordan di A,

cioè $\underbrace{(A-\lambda_i I)v_{i1} = 0 \quad (A-\lambda_i I)v_{i2} = v_{i1}, \ \dots \ (A-\lambda_i I)v_{i,\ell_i} = v_{i,\ell_i-1}}_{\text{(catena i-esima con autoval. } \lambda_i)}$

$i = 1, 2, \dots \ell$

Possiamo costruire catene di Jordan per M così:

$w_{ij} = U_1 \cdot v_{ij}$ per ogni i, j

$$(M-\lambda_i I)w_{i1} = \underbrace{MU_1}_{= U_1 A} v_{i1} - \lambda_i U_1 v_{i1} = U_1 (A - \lambda_i I) v_{i1} = 0$$

$$(M-\lambda_i I)w_{ij} = MU_1 v_{ij} - \lambda_i U_1 v_{ij} = U_1 (A - \lambda_i I) v_{ij} = U_1 v_{i,j-1} = w_{i,j-1}$$

Quindi w_{ij} è una catena di Jordan di M. Potrebbe però non essere finita! Quindi in generale ho solo un segmento iniziale di catena. $\square$

⚠ Questioni di non-unicità nella forma di Jordan di M

$M = I = V \cdot I \cdot V^{-1}$ è una dec. di Jordan $\forall\, V$ invertibile

Se prendo Ve_1, prima colonna di V, è un vettore di una catena di Jordan, e quindi span$\{Ve_1\}$ è invariante.

Caso in cui non abbiamo problemi di unicità: se prendo

tutte le catene di Jordan da ogni autovalore scelto:
scelgo un insieme di autovalori di M, e da ognuno prendo tutte le catene con quell'autovalore.

Es: sottospazio invariante stabile di M: span di tutti i vettori di tutte le catene di Jordan con autovalori $|\lambda| < 1$ (dentro il disco unitario):
cioè, tutti i vettori v tali che la successione

$$v, Mv, M^2v, M^3v, \dots \text{ converge a } 0.$$

Detto in altro modo: se scrivo (in una base di Jordan)

$$M = U\left[\begin{array}{c|c} J_1 & 0 \\ \hline 0 & J_2 \end{array}\right]U^{-1} \quad \text{con} \quad \begin{array}{l}\Lambda(J_1) \subset \text{disco unitario} \\ \Lambda(J_2) \text{ tutto fuori dell'interno del disco}\end{array}$$

allora se prendo $v = U\begin{bmatrix} * \\ 0 \end{bmatrix}$ ho

$$M^k v = U\left[\begin{array}{c|c} J_1^k & \\ \hline & J_2^k \end{array}\right]U^{-1} \cdot U\begin{bmatrix} * \\ 0 \end{bmatrix} = \begin{bmatrix} J_1^k * \\ 0 \end{bmatrix} \to 0$$

e in generale se $v = U\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$

$$M^k v = U\begin{bmatrix} J_1^k & \\ & J_2^k \end{bmatrix}U^{-1}U\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = U\begin{bmatrix} J_1^k w_1 \\ J_2^k w_2 \end{bmatrix}$$

$J_1^k w_1 \to 0$

$J_2^k w_2 \to$ non converge mai a 0 se $w_2 \neq 0$

Diciamo che un sottosp. inv. è <u>ben separato</u> se ha preso <u>tutte</u> le catene di <u>tutti</u> i blocchi da un certo sottoinsieme di autovalori.

Domanda: dato un insieme di autovalori di M (ad es.

quelli nel cerchio unitario), come faccio a calcolare il sott. inv. ben separato associato ad essi?

Forma di Jordan $\Rightarrow$ mal condizionata

Meglio utilizzare una forma di Schur

$$M = Q \cdot \begin{bmatrix} \lambda_1 & \times & \times & - & - & \times \\ & \lambda_2 & \ddots & & & \times \\ & & \ddots & \ddots & & \times \times \\ 0 & & & & \lambda_n & \times \end{bmatrix} Q^{-1} \qquad T = \begin{bmatrix} \overset{k \times k}{\boxed{T_{11}}} & T_{12} \\ 0 & T_{22} \end{bmatrix}$$

Oss: se gli autovalori che mi interessano sono nelle prime k posizioni, allora basta dividere $Q = [\underbrace{Q_1}_{k} \; Q_2]$

e $\operatorname{Im} Q_1$ è il sottosp. invariante che cerco.

$$MQ_1 = Q \begin{bmatrix} \lambda_1 & \times & \times \\ & \ddots & \times \\ 0 & & \lambda_k \end{bmatrix} = Q_1 \cdot \text{blocco } (1,1) \text{ della } T = Q_1 \begin{bmatrix} T_1 \\ 0 \end{bmatrix}$$

Riordinare forme di Schur: data $M = QTQ^{-1}$, sono in grado di trovare un'altra forma di Schur $M = \hat{Q}\hat{T}\hat{Q}^{-1}$ che ha elem. diagonali di mio interesse nelle prime k posizioni?

Siamo in grado di risolvere il problema con eq. di Sylvester

Oss: mi basta avere un modo di scambiare due blocchi di autovalori distinti, cioè, dato

$T = \begin{bmatrix} A & -C \\ 0 & B \end{bmatrix}$, trovare una fattorizzazione $T = Q \begin{bmatrix} \hat{B} & -\hat{C} \\ 0 & \hat{A} \end{bmatrix} Q^*$

dove $\operatorname{diag}(B) = \operatorname{diag}(\hat{B})$, $\operatorname{diag}(A) = \operatorname{diag}(\hat{A})$

$$\Lambda(A) \cap \Lambda(B) = \emptyset$$

Idea: uso un'equazione di Sylvester per passare a una diagonale a blocchi. Se X risolve eq. di Sylvester,

$$\overset{\begin{bmatrix}0&1\\1&0\end{bmatrix}}{\begin{bmatrix}1 & -X\\ 0 & 1\end{bmatrix}}\begin{bmatrix}A & -C\\ 0 & B\end{bmatrix}\overset{\begin{bmatrix}0&1\\1&0\end{bmatrix}}{\begin{bmatrix}1 & X\\ 0 & 1\end{bmatrix}} = \begin{bmatrix}A & 0\\ 0 & B\end{bmatrix}\cdot\begin{bmatrix}0&1\\1&0\end{bmatrix}$$

$$\begin{bmatrix}0&1\\1&0\end{bmatrix}$$

$$\begin{bmatrix}0 & 1\\ 1 & -X\end{bmatrix}\begin{bmatrix}A & -C\\ 0 & B\end{bmatrix}\underbrace{\begin{bmatrix}X & 1\\ 1 & 0\end{bmatrix}}_{P} = \begin{bmatrix}B & 0\\ 0 & A\end{bmatrix}$$

però questa non è ortogonale!

Idea: per ottenere lo stesso risultato, mi basta usare il fattore Q di $QR = \begin{bmatrix}X & 1\\ 1 & 0\end{bmatrix}$

$$R^{-1}Q^*\begin{bmatrix}A & -C\\ 0 & B\end{bmatrix}QR = \begin{bmatrix}B & 0\\ 0 & A\end{bmatrix}$$

$$Q^*\begin{bmatrix}A & -C\\ 0 & B\end{bmatrix}Q = R\begin{bmatrix}B & 0\\ 0 & A\end{bmatrix}R^{-1}$$

Se parto da una forma di Schur, A, B triangolari $\Rightarrow$

$$RHS = \begin{bmatrix}\hat{B} & \hat{C}\\ 0 & \hat{A}\end{bmatrix} \quad \text{con} \quad \mathrm{diag}(\hat{B}) = \mathrm{diag}(B),\ \mathrm{diag}(\hat{A}) = \mathrm{diag}(A).$$