Equazioni di Sylvester

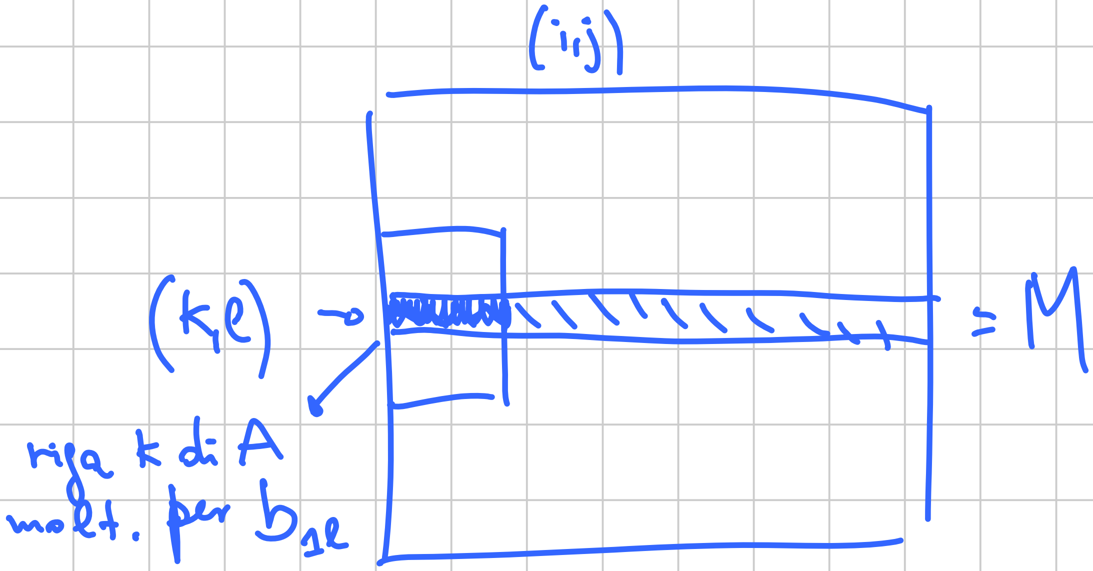
$$AX - XB = C \qquad A \in \mathbb{R}^{m\times m} \quad B \in \mathbb{R}^{n\times n} \quad X, C \in \mathbb{R}^{m\times n}$$

Entrata (i,j) $\quad \sum_k a_{ik} x_{kj} - \sum_k x_{ik} b_{kj} = c_{ij}$

$m \cdot n$ eq. lineari in $m \cdot n$ incognite

coppia di indici (i,j) $\longrightarrow$ singolo indice delle variabili vettorizzate

$$(AXB)_{k,l} = \sum_{i,j} a_{ki} x_{ij} b_{jl}$$



$$\underbrace{\sum_{i,j} a_{ki} x_{ij} b_{jl} = \left[a_{k1}b_{1l} \;\; a_{k2}b_{1l} \cdots a_{km}b_{1l} \,\middle|\, a_{k1}b_{2l} \cdots a_{km}b_{2l} \;\; \cdots a_{k1}b_{nl} \cdots a_{km}b_{nl} \right] \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{m1} \\ \hline x_{12} \\ x_{22} \\ \vdots \\ x_{m2} \\ \hline \vdots \\ \hline x_{1n} \\ \vdots \\ x_{mn} \end{bmatrix}^t}$$

La matrice M associata a $\operatorname{vec} X \mapsto \operatorname{vec}(AXB)$

è

$$M = \left[\begin{array}{c|c|c|c} b_{11}A & b_{21}A & \cdots & b_{n1}A \\ \hline b_{12}A & \ddots & A \quad \cdots & A \\ \hline b_{13}A & \ddots & A \quad \cdots & A \\ \hline \vdots & \vdots & & \ddots \\ b_{1q}A & A & & b_{nq}A \end{array}\right] = B^T \otimes A$$

<u>Def</u>: date X, Y matrici, si chiama $X \otimes Y$ (prodotto di Kronecker) la matrice a blocchi che ha nel blocco (i,j) $x_{ij}Y$

$$X \otimes Y \quad \left[\begin{array}{c|c|c} x_{11}Y & x_{12}Y & \cdots x_{1m}Y \\ \hline x_{21}Y & \ddots \quad Y & Y \\ \hline \vdots \quad Y & Y & \ddots \quad Y \\ \hline x_{m1}Y & Y & x_{mn}Y \end{array}\right]$$

Quindi, abbiamo dimostrato che

$$\mathrm{vec}(AXB) = (B^T \otimes A)\,\mathrm{vec}\,X$$

$$\underset{pq\times 1}{[\;]} = \underset{pq\times mn}{[\;\;\;]}\;\underset{mn\times 1}{[\;]}.$$

Proprietà del prodotto di Kronecker:

1) $(A \otimes B)(C \otimes D) = AC \otimes BD$

dim: $(A\otimes B)\underbrace{(C\otimes D)\,\mathrm{vec}(X)} = (A\otimes B)\,\mathrm{vec}(DXC^T) =$

$= \mathrm{vec}(BDXC^TA^T) = \mathrm{vec}((BD)X(AC)^T) = (AC\otimes BD)\,\mathrm{vec}\,X$

2) $(A \otimes B)^T = A^T \otimes B^T$

3) Q, Z ortogonali $\Rightarrow$ $Q \otimes Z$ ortogonale

$$(Q\otimes Z)^T (Q \otimes Z) = (Q^T \otimes Z^T)(Q \otimes Z) = Q^T Q \otimes Z^T Z = I \otimes I$$
$$= I$$

$$I \otimes I = \begin{pmatrix} I & & & 0 \\ & I & & \\ & & I & \\ 0 & & & \ddots \\ & & & & I \end{pmatrix}$$

3) T, U upper triangular (o diagonali) $\Rightarrow$ $T \otimes U$ upper tri. (o diag.)

4) Alcune decomposizioni "passano" ai fattori:

se ho $A = U_A S_A V_A^T$ $\quad B = U_B S_B V_B^T$ SVDs

$$A \otimes B = \underbrace{(U_A \otimes U_B)}_{\text{ortogonale}} \underbrace{(S_A \otimes S_B)}_{\text{diagonale}} \underbrace{(V_A \otimes V_B)^T}_{\text{ortogonale}}$$ è una SVD

I valori singolari di $A \otimes B$ sono i prodotti

$\{\sigma_i \mu_j : i=1,\dots m,\ j=1,\dots,n\}$, dove σ_i sono i v.s. di A $\quad S_A = \mathrm{diag}(\sigma_i)$

μ_j sono i v.s. di B $\quad S_B = \mathrm{diag}(\mu_j)$

Possiamo trasformare l'eq. di Sylvester

$$\underbrace{AX - XB}_{} = C$$

in un sist. lineare vettorizzato:

Matrice associata a $\quad AXI - IXB$:

$\mathrm{vec}(AXI) = (I \otimes A)\,\mathrm{vec}X \quad \mathrm{vec}(IXB) = (B^T \otimes I)\,\mathrm{vec}\,X$

$\mathrm{vec}(AXI - IXB) = (I \otimes A - B^T \otimes I)\,\mathrm{vec}\,X$

$(I \otimes A - B^T \otimes I)\,\mathrm{vec}\,X = \mathrm{vec}\,C$

sistema lineare $mn \times mn$

<u>Lemma</u>: $AX - XB = C$ ha una e una sola soluzione X

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

<u>dim</u>: $AX-XB=C$ ha una e una sola sol $\Leftrightarrow$

$I \otimes A - B^T \otimes I$ è invertibile.

Siano $A = Q_A T_A Q_A^*$, $B^T = Q_B T_B Q_B^*$ decomp. di Schur, allora

$$I\otimes A - B^T\otimes I = \underbrace{(Q_B\otimes Q_A)}_{\text{ortog.}}\underbrace{(1\otimes T_A - T_B\otimes 1)}_{\text{tr. sup.}}\underbrace{(Q_B\otimes Q_A)^*}_{\text{ortog}^*}$$

è una decomposizione di Schur

$I\otimes A - B^T\otimes I$ invertibile $\Leftrightarrow$ $1\otimes T_A - T_B\otimes 1$ $\Leftrightarrow$ elem. diag. sono $\neq 0$

elem. diagonali di $1\otimes T_A - T_B\otimes 1$:

$$\begin{bmatrix} \boxed{T_A} & & 0 \\ & \boxed{T_A} & \\ & \ddots & \\ 0 & & \boxed{T_A} \end{bmatrix} - \begin{bmatrix} t^B_{11} I & \ddots & t_{ij} I \\ & \ddots & \\ 0 & & t^B_{nn} I \end{bmatrix}$$

sulla diagonale, ho $t^A_{ii} - t^B_{jj}$

$I\otimes T_A - T_B\otimes I$ invertibile $\Leftrightarrow$ $t^A_{ii} - t^B_{jj} \neq 0$ per ogni i, j

$\iff \lambda_i^A \neq \lambda_j^B$ per ogni i,j. $\square$

Come si risolve numericamente $AX - XB = C$?

Vettorizzo + elim. Gauss: costo $O(m^3 n^3)$

Bartels-Stewart algorithm: $O(m^3 + n^3)$

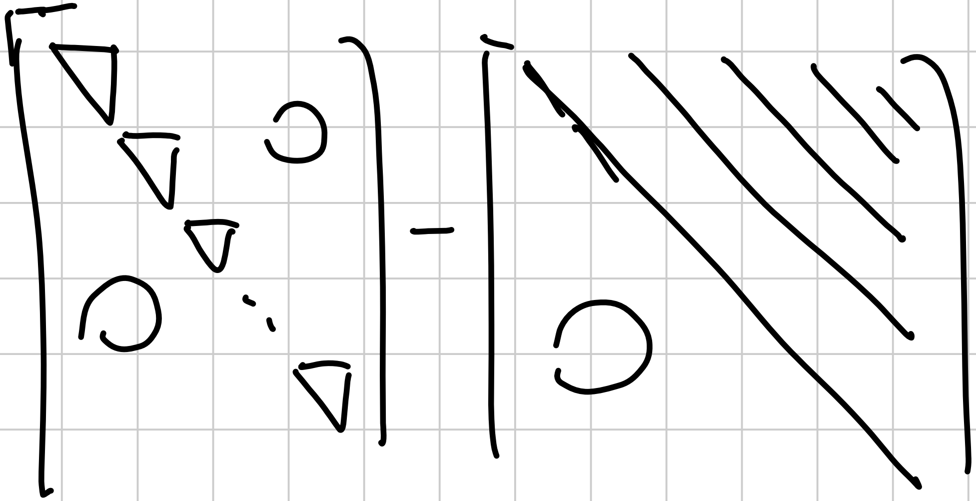
Idea: inverto pezzo per pezzo

$$I \otimes A - B^T \otimes I = (Q_A \otimes Q_B)(I \otimes T_A - T_B \otimes I)(Q_A \otimes Q_B)^*$$

(dec. di Schur costano $O(m^3)$ e $O(n^3)$ rispettivamente)

$$\operatorname{vec} X = \underbrace{(Q_A \otimes Q_B)}\,\underbrace{(I \otimes T_A - T_B \otimes I)^{-1}}\,\underbrace{(Q_A^* \otimes Q_B^*)\operatorname{vec} C}_{= \operatorname{vec}\left(Q_B^* C (Q_A^*)^T\right)}$$

$O(mn(m+n))$



$\rightsquigarrow O(m+n)$ nonzeri per riga

mn righe $\Rightarrow$ tempo $O(mn(m+n))$

Altro modo di vedere lo stesso algoritmo:

$$A = Q_A T_A Q_A^*, \quad B^T = Q_B T_B Q_B^*$$

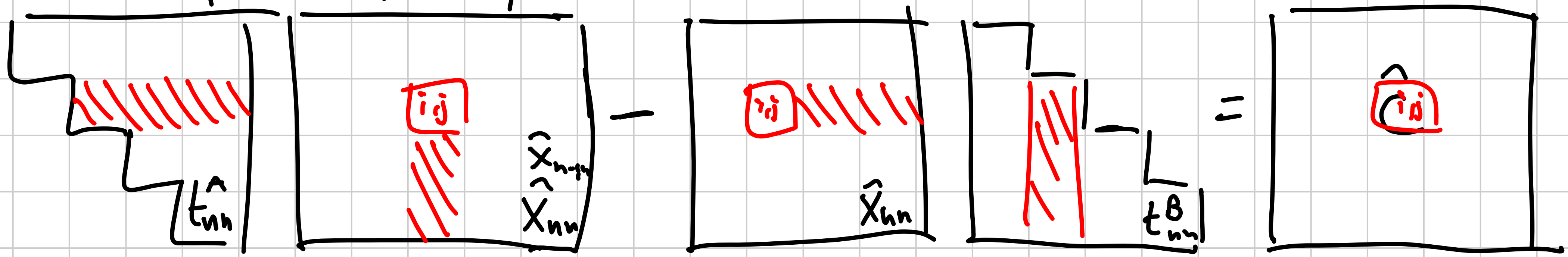
$$C = AX - XB = Q_A T_A Q_A^* X - X Q_B^{*T} T_B^T Q_B^T$$

$$\underbrace{Q_A^* C Q_B^{*T}}_{\hat{C}} = T_A \underbrace{Q_A^* X Q_B^{*T}}_{\hat{X}} - \underbrace{Q_A^* X Q_B^{*T}}_{\hat{X}} T_B^T$$

$$\hat{C} = T_A \hat{X} - \hat{X} T_B^T$$

eq. di Sylv. con coeff. triangolari

(m=n per semplicità)



Entrata (n,n): $t^A_{nn}\hat{X}_{nn} - \hat{X}_{nn}t^B_{nn} = \hat{C}_{nn}$

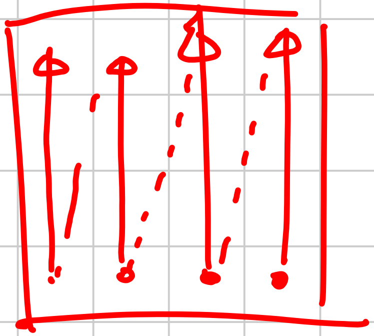
Entrata $(n-1,n)$: $t^A_{n-1,n-1}\hat{X}_{n-1,n} + t^A_{n-1,n}\hat{X}_{nn} - \hat{X}_{n-1,n}t^B_{nn} = \hat{C}_{n-1,n}$

che posso risolvere per $\hat{X}_{n-1,n}$ se so già $\hat{X}_{nn}$

In generale,

(i,j) $$\sum_{k\geq i} t^A_{ik}\hat{X}_{kj} + \sum_{k\geq j} \hat{X}_{ik}t^B_{jk} = \hat{C}_{ij}$$

contiene solo $\hat{X}_{ij}$ ed entrate sotto e a dx di $\hat{X}_{ij}$

$\Rightarrow$ posso risolvere per sost. all'indietro in un ordine opportuno ad es. 

Algoritmo di Bartels-Stewart:

1) Calcola $\hat{C} = Q_A^* C Q_B^{*T}$
2) Calcola $\hat{X}$ per sost. all'indietro
3) Calcola $X = Q_A X Q_B^T$

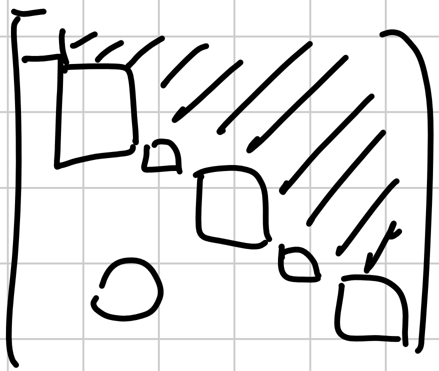
Commenti: la stessa idea funziona per

$X - AXB = C$

e per $AXB - CXD = E$ con una fattorizzazione QZ

$(A, C) \to (Q T_A Z, \; Q T_C Z)$ con Q, Z ortogonali, T_A, T_C tr. sup.

- La stessa cosa funziona con forma di Schur reale

$T_A =$ 

Basta farlo a blocchi: ogni equazione è

$$t^A_{ii}\hat{x}_{ij} - \hat{x}_{ij}t^B_{jj} = \hat{c}_{ij} - \text{roba}$$

che è una piccola equazione di sylvester con $(m,n) \in \{(1,1), (1,2), (2,1), (2,2)\}$

Si risolvono vettorizzando, in $O(1)$.

- no divisioni per 0, se l'equazione era risolubile con sol. unica.
- È stabile all'indietro nel senso che la $\tilde{X}$ che calcolate è la soluzione esatta di $\tilde{M} \operatorname{vec} \tilde{X} = \operatorname{vec} \tilde{C}$ con $\|\tilde{C} - C\| = \|C\| \cdot O(u)$, $\|\tilde{M} - M\| = \|M\| \cdot O(u)$

$$M = I \otimes A - B^T \otimes I$$

- Esistono $\tilde{A}$, $\tilde{B}$ tali che la $\tilde{X}$ calcolata è sol. esatta di $\tilde{A}\tilde{X} - \tilde{X}\tilde{B} = \tilde{C}$, e $\frac{\|A-\tilde{A}\|}{\|A\|}$, $\frac{\|B-\tilde{B}\|}{\|B\|}$, $\frac{\|C-\tilde{C}\|}{\|C\|}$ sono $O(u)$?

Sketch dimostrazione: voglio trovare Δ_A, Δ_B, Δ_C t.c.

$$(A+\Delta_A)\tilde{X} - \tilde{X}(B+\Delta_B) = C + \Delta_C$$

$$\Delta_A\tilde{X} - \tilde{X}\Delta_B - \Delta_C = -\Big(\underbrace{A\tilde{X} - \tilde{X}B - C}_{\text{residuo della } \tilde{X}}\Big)$$

piccolo perché è lo stesso residuo del sist. lineare $M \operatorname{vec}\tilde{X} - \operatorname{vec} C$, che sto risolvendo in modo stabile all'indietro

Sto cercando di rendere più piccolo possibile

$$\left\| \begin{bmatrix} \tilde{X}^T \otimes I & -I \otimes \tilde{X} & -I \otimes I \end{bmatrix} \begin{bmatrix} \mathrm{vec}\,\Delta_A \\ \mathrm{vec}\,\Delta_B \\ \mathrm{vec}\,\Delta_C \end{bmatrix} - \mathrm{vec}\left(-(A\tilde{X} - \tilde{X}B - C)\right) \right\|$$

La soluzione di minima norma di questo prob. ai minimi quadrati

è

$$\underbrace{\begin{bmatrix} \tilde{X}^T \otimes I & -I \otimes \tilde{X} & -I \otimes I \end{bmatrix}^{\dagger}}_{} \cdot \mathrm{vec}\underbrace{\left(-(A\tilde{X} - \tilde{X}B - C)\right)}_{\text{piccolo}} = \begin{bmatrix} \mathrm{vec}\,\Delta_A \\ \mathrm{vec}\,\Delta_B \\ \mathrm{vec}\,\Delta_C \end{bmatrix}.$$

e questo può essere grande. (Ad esempio, prendendo $X = \mathrm{diag}\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ con $\lambda_n \ll 1$)

Numero di condizionamento: è $K(M)$ $\quad M = I \otimes A - B^T \otimes I$

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

$$= \| I \| \cdot \| A \| + \| B^T \| \cdot \| I \| = \| A \| + \| B \|$$

$\left(\| X \otimes Y \| = \| X \| \cdot \| Y \| \right.$, perché $SVD(X \otimes Y)$ = prodotti di $SVD(X)$ e $\left. SVD(Y) \right)$

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

Difetti,

$$\sigma_{min}(M) = \min_{z \in \mathbb{C}^{mn}} \frac{\| M z \|}{\| z \|} = \min_{Z \in \mathbb{C}^{m \times n}} \frac{\| AZ - ZB \|_F}{\| Z \|_F}$$

Se A, B sono diagonali, $M = \mathrm{diag}(\lambda_i - \mu_j)$

$$\mathrm{sep}(A,B) = \min_{\substack{\lambda \in \Lambda(A) \\ \mu \in \Lambda(B)}} |\lambda - \mu|$$

È anche se A, B sono normali (forma di Schur diagonale) perché posso cambiare base e renderle diagonali).