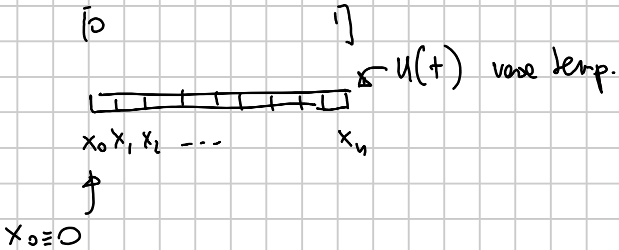
Problemi di controllo di larga scala

Note Title 2020-05-28

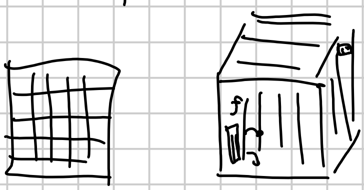
<u>Es</u>: eq. del calore su una barretta solida



$$\dot{x} = Ax + Bu$$

$A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times 1}$

Anche 2D, 3D:



Tipicamente, A grande e sparsa, B alta e stretta ($m \ll n$)

$Q \in \mathbb{R}^{n\times n}$ in $\int x^T Q x + u^T R u$ viene scelta di rango basso (target = temperatura solo in pochi punti della stanza).

$$A^T X + XA + Q - XGX = 0 \quad \text{(tutte } n\times n\text{, } n \text{ grande)}$$

A sparsa; rango basso; $G = BR^{-1}B^T$ rango basso, m

Andremo in realtà a risolvere solo Lyap: (con $\Lambda(A) \subset LHP$)

$$AX + XA^* + BB^* = 0$$

$A \in \mathbb{R}^{n\times n}$ sparsa, $B \in \mathbb{R}^{n\times m}$, $m \ll n$

(poi posso risolvere Riccati con Newton)

Possiamo supporre $m=1$: difatti $B = [b_1 \cdots b_m]$

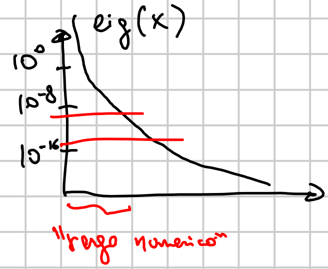
$$AX + XA^* + \underbrace{b_1b_1^* + b_2b_2^* + \dots + b_mb_m^*} = 0$$

ha come sol. la somma delle sol. di $\boxed{AX + XA^* + b_ib_i^* = 0}$ (L), $i = 1,\dots,m$

(I metodi che vedremo funzionano bene se A "non troppo lontana" da $AA^* = A^*A$)

Problema: X in generale è densa, e positiva definita ((A, b) controllabile)

Soluzione: spesso, i val. sing. (autovalori) di X decadono in fretta, quindi posso approssimarla bene con $X \approx ZZ^*$, Z di rango basso



Questo permette di "scendere" sotto $O(n^2)$ come complessità.

ADI ("alternating direction implicit iteration")

Idea: scegliamo $\tau > 0$, così $\Lambda(A - \tau I) \subset LHP$

$$AX + XA^* + bb^* = 0$$

$$\Leftrightarrow$$

$$\cancel{AXA^*} - \tau XA^* - \tau AX + \cancel{\tau^2 X} - \left(\cancel{AXA^*} + \tau XA^* + \tau AX + \cancel{\tau^2 X}\right) - 2\tau bb^* = 0$$

$$\Updownarrow$$

$$\underbrace{(A-\tau I)X(A-\tau I)^*}_{} - \underbrace{(A+\tau I)X(A+\tau I)^*}_{} - 2\tau bb^* = 0$$

$$\Updownarrow$$

(S) $$X - \underbrace{(A-\tau I)^{-1}(A+\tau I)}_{c(A)} X \underbrace{(A+\tau I)^*(A-\tau I)^{-*}}_{c(A)^*} - \underbrace{2\tau (A-\tau I)^{-1} b b^* (A-\tau I)^{-*}}_{= b_0 b_0^*,\ \text{con}} = 0$$

$$b_0 = \sqrt{2\tau}\,(A-\tau I)^{-1} b$$

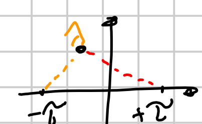
$$X - c(A)Xc(A)^* = b_0 b_0^* \quad \text{(eq. di Stein)}$$

$$c(x) = \frac{x+\tau}{x-\tau}$$ "parente" di $\frac{1+x}{1-x}$ usata nella sezione sulla funz. segno "Cayley transform".

Notare che $\lambda \in LHP \Rightarrow |c(\lambda)| < 1$, difatti $\lambda \in LHP \Leftrightarrow \mathrm{dist}(\lambda, -\tau) < \mathrm{dist}(\lambda, \tau)$

$$\frac{d(\lambda,-\tau)}{d(\lambda,\tau)} = \frac{|\lambda+\tau|}{|\lambda-\tau|} < 1$$

$\Lambda(A) \subset LHP$
$\Downarrow$
$\Lambda(c(A)) \subset$ disco unitario



Possiamo usare un metodo di punto fisso:

$$X_{k+1} = c(A) X_k c(A)^* + b_0 b_0^*$$

$X_0 = 0 \quad X_1 = b_0 b_0^* \quad X_2 = b_0 b_0^* + c(A) b_0 b_0^* c(A)^*$

$$X_3 = b_0 b_0^* + c(A) b_0 b_0^* c(A)^* + c(A)^2 b_0 b_0^* c(A)^{2*}$$

$$X_{k+1} = b_0 b_0^* + c(A) b_0 b_0^* c(A)^* + \dots + c(A)^k b_0 b_0^* c(A)^{k*}$$

(una sorta di serie geometrica, che converge perché $\rho(c(A)) < 1$

<u>Modifica 1</u>: voglio poter cambiare τ ad ogni passo

$$c_k := \frac{x+\tau_k}{x-\tau_k} \qquad c_k(A) = (A-\tau_k I)^{-1}(A+\tau_k I)$$

$$\begin{cases} X_0 = 0 \\ X_k = \underbrace{c_k(A) X_{k-1} c_k(A)^*}_{Z_{k-1}Z_{k-1}^* \ rk\ k-1 \text{ per induzione}} + \underbrace{2\tau_k (A-\tau_k I)^{-1} b_0 b_0^* (A-\tau_k I)^{-*}}_{rk\ 1} \end{cases} \qquad X \approx ZZ^*$$

$$X_k = Z_k Z_k^*$$

$$Z_k = \left[c_k(A) Z_{k-1} \quad \sqrt{2\tau_k}(A-\tau_k I)^{-1} b_0 \right]$$

Posso riscrivere l'iterazione in termini degli Z_k!

$$\begin{cases} Z_1 = \sqrt{2\tau_1}(A-\tau_1 I)^{-1} b \\ Z_k = \left[c_k(A) Z_{k-1} \quad \underbrace{\sqrt{2\tau_k}(A-\tau_k I)^{-1}}_{} b_0 \right] \end{cases} \qquad \text{"Low-rank ADI"}$$

$$d_k(A) = \sqrt{2\tau_k}(A-\tau_k I)^{-1}$$

Ancora costosa, ma c'è un modo che la rende più comoda da calcolare:

k iterazioni producono

$$c_j(x) = \frac{x+\tau_j}{x-\tau_j} \qquad d_j(x) = \sqrt{2\tau_j}(x-\tau_j)^{-1}$$

funzioni razionali

$$Z_1 = d_1(A) b_0$$

$$Z_2 = \left[c_2(A) d_1(A) b_0 \quad d_2(A) b_0 \right]$$

$$Z_3 = \left[c_3(A) c_2(A) d_1(A) b_0 \quad c_3(A) d_2(A) b_0 \quad d_3(A) b_0 \right]$$

$$\vdots$$

$$Z_k = \left[c_k c_{k-1} c_{k-2} \cdots c_2 d_1(A) b_0 \quad c_k c_{k-1} \cdots c_3 d_2(A) b_0 \quad \cdots \quad c_k d_{k-1}(A) b_0 \quad d_k(A) b_0 \right]$$

(k colonne della matrice Z_k)

Se c, d fossero la stessa funzione, potrei calcolare queste quantità "da destra" in modo semplice:

$$Z_k = \Big[\underbrace{d_k(A)b_0}_{z_k} \;\Big|\; \overbrace{\underbrace{d_{k-1}c_k(A)b_0}_{z_{k-1}} \quad \underbrace{d_{k-2}c_{k-1}c_k(A)b_0}_{z_{k-2}} \quad \cdots \quad \underbrace{d_1c_2c_3\cdots c_k(A)b_0}_{z_1}}^{*\,Z_{k-1}}\Big]$$

sto applicando al vettore di partenza b_0 una funzione razionale di A di grado 1 ad ogni passo:

se ho calcolato $d_j(A)v = \sqrt{2\tau_j}\,(A-\tau_j I)^{-1}v$, allora

$$c_j(A)v = (A+\tau_j I)(A-\tau_j I)^{-1}v = \frac{1}{\sqrt{2\tau_j}}(A+\tau_j I)\,d_j(A)v$$

Iterazione:

$$z_k = d_k(A)b_0 \qquad z_j = d_j\frac{1}{\sqrt{2\tau_{j+1}}}(A+\tau_{j+1}I)\,z_{j+1}$$

LR-ADI "versione finale"

$$z_j = \sqrt{2\tau_j}\,(A-\tau_j I)^{-1}\frac{1}{\sqrt{2\tau_{j+1}}}(A+\tau_{j+1}I)\,z_{j+1}$$

Iterazione "semplice" da calcolare: ad ogni passo, devo fare:

- 1 prodotto con A ↳ si può anche eliminare riscrivendo $\sqrt{\frac{2\tau_j}{2\tau_{j+1}}}\,\frac{x+\tau_{j+1}}{x-\tau_j} = \sqrt{\frac{2\tau_j}{2\tau_{j+1}}}\left(1+\frac{(\tau_{j+1}+\tau_j)}{x-\tau_j}\right)$
- 1 soluzione di un sist. lineare del tipo $(A-\tau_j I)^{-1}w$

ADI

$$\begin{cases} X_0 = 0 \\ X_k = \underbrace{c_k(A)X_{k-1}c_k(A)^*}_{} + \underbrace{d_k(A)b_0}_{}b_0^*d_k(A)^* \end{cases}$$

$$c_j(x) = \frac{x+\tau_j}{x-\tau_j} \qquad d_j(x) = \sqrt{2\tau_j}\,\frac{1}{x-\tau_j}$$

$X_k = [z_1 z_2 \cdots z_k][z_1 z_2 \cdots z_k]^*$ equivalentemente calcolabili come

$$\begin{cases} z_k = d_j(A)b_0 \\ z_j = \sqrt{\frac{2\tau_j}{2\tau_{j+1}}}\,(A-\tau_j I)^{-1}(A+\tau_{j+1}I)\,z_{j+1} \end{cases}$$

↳ versione implementabile, risolve un sist. lineare per passo.

Quanto velocemente converge? Chiamiamo X soluzione esatta

$$X_k = c_k(A)X_{k-1}c_k(A)^* + d_k(A)b_0b_0^*d_k(A)^*$$

$$X = c_k(A)X\,c_k(A)^* + d_k(A)b_0b_0^*d_k(A)^*$$

$$X_k - X = C_k(A)(X_{k-1}-X)C_k(A)^* = C_k(A)C_{k-1}(A)(X_{k-2}-X)C_{k-1}(A)^*C_k(A)^* =$$

$$\dots = g_k(A)(\underbrace{X_0}_{=0}-X_*)g_k(A)^*$$

$$g_k(x) = \prod_{j=1}^{k} \frac{x+\tau_j}{x-\tau_j}$$

residuo va a zero tanto più velocemente quanto $g_k(A)$ tende a zero.

Se $A = V \cdot \Lambda \cdot V^{-1}$

$$\|g(A)\| = \left\| V \cdot \begin{bmatrix} g(\lambda_1) & & & \\ & g(\lambda_2) & & \\ & & \ddots & \\ & & & g(\lambda_n) \end{bmatrix} V^{-1} \right\| \le \kappa(V) \cdot \max_{\lambda \in \Lambda(A)} |g(\lambda)|$$

Possiamo scegliere $\tau_1, \tau_2, \dots \tau_k$ in modo che

$$\max_{\lambda \in \Lambda(A)} \frac{(x+\tau_1)(x+\tau_2)\cdots(x+\tau_k)}{(x-\tau_1)(x-\tau_2)\cdots(x-\tau_k)}$$

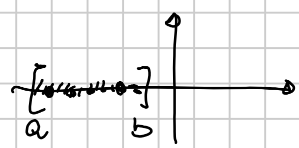
sia piccolo?

Idealmente vorrei prendere $-\tau_i$ = autovalori di A, se sono meno di $k \Rightarrow$ esattamente zero.

Detto in altro modo, stiamo cercando $p(x) = (x-\tau_1)(x-\tau_2)\cdots(x-\tau_k)$

tale che $p(-\lambda)$ piccolo, $p(\lambda)$ grande per ogni autovalore λ di A

Se sapete, per esempio, che $\Lambda(A) \subseteq [a,b]$



allora potete pensare di risolvere il problema

$$\min_{\tau_1,\tau_2,\dots,\tau_k} \max_{[a,b]} \left| \frac{(x+\tau_1)\cdots\cdots(x+\tau_k)}{(x-\tau_1)\cdots\cdots(x-\tau_k)} \right|$$

Si sa da risultati teorici che questo minimo va a zero come r^k per un certo $0<r<1$

Conseguenza.

$\|X_k - X\| \sim r^{2k}$ $\Rightarrow$ i valori sing. di X devono andare a 0 almeno come r^k.

matrice di rango k, ($Z_k Z_k^*$ dove Z_k ha k colonne)

Le formule per Z_k mettono in evidenza che

colonne di $Z_k \in K_q(A, b_0)$

spazio di Krylov razionale $\left\{ r(A) b_0 : r(x) \text{ funzione razionale } r(x) = \frac{p(x)}{q(x)}, \deg p \leq \deg q \right\}$

$$q(x) = (x - \tau_1)(x - \tau_2) \cdots (x - \tau_k)$$

$\Rightarrow$ stiamo costruendo un'approssimazione di Krylov razionale di X

Alternativa:

1) Calcolo $U_k \in \mathbb{R}^{n \times k}$ base ortonormale di $K_q(A, b_0)$ con Arnoldi razionale,

2) Cerco una soluzione $X_k = U_k Y_k U_k^*$ $\quad Y_k \in \mathbb{R}^{k \times k}$

risolvendo "proiezione di Galerkin"

$$U_k^* (A X_k + X_k A^* + b b^*) U_k = 0 \qquad \boxed{U_k^*}(\text{Lyap})\boxed{U_k} = 0$$

$$(U_k^* A U_k) Y_k \cancel{U_k^* U_k} + \cancel{U_k^* U_k} Y_k (U_k^* A^* U_k) + U_k^* b b^* U_k = 0$$

eq. di Lyapunov "proiettata" $k \times k$ nella Y_k.

<u>Difficoltà 1</u>: come per ADI, trovare poli τ_j "buoni"

<u>Difficoltà 2</u>: se A è nonsimmetrica, può succedere che $\Lambda(A) \subset LHP$

ma $\Lambda(U_k^* A U_k) \not\subset LHP$!