



---

# Approssimazione di Funzioni di Matrici di grandi dimensioni ed Applicazioni

V. Simoncini

Dipartimento di Matematica e CIRSA, Bologna

`valeria@dm.unibo.it`

## Il Problema

Dati  $v \in \mathbb{R}^n$  e  $A \in \mathbb{R}^{n \times n}$ , approssimare

$$x = f(A)v$$

con  $f$  funzione regolare tale che  $f(A)$  sia definita

### Enfasi:

- $A$  grandi dimensioni
- $A$  simmetrica (semi)def. positiva, oppure  $A$  reale positiva

## Contesto

- $A$  di piccole dimensioni:

$$A \text{ simmetrica, } A = X\Lambda X^{\top} \Rightarrow f(A) = Xf(\Lambda)X^{\top}$$

Analogo, ma più complesso, il calcolo per  $A$  non simmetrica

- $A$  di medie e grandi dimensioni:

$$f(A) \quad \text{vs.} \quad f(A)v$$

## le Applicazioni

Tra queste:

- Risoluzione numerica di EDP evolutive  
(es.  $\exp(\lambda)$ ,  $\sqrt{\lambda^{-1}}$ ,  $\cos(\lambda)$ ,  $\varphi_k(\lambda)$ ...)
- Risoluzione numerica di alcuni Problemi Inversi ( $\exp(\lambda)$ ,  $\cosh(\lambda)$ , ...)
- Flussi su varietà
- Problemi di calcolo scientifico (es. QCD,  $\text{sign}(\lambda)$ )
- (Analisi di) Modelli ridotti di Sistemi Dinamici  
(mediante matrici Grammiane)

⇒ Alcuni esempi in seguito.

## le Applicazioni

Tra queste:

- Risoluzione numerica di EDP evolutive  
(es.  $\exp(\lambda)$ ,  $\sqrt{\lambda^{-1}}$ ,  $\cos(\lambda)$ ,  $\varphi_k(\lambda)$ ...)
- Risoluzione numerica di alcuni Problemi Inversi ( $\exp(\lambda)$ ,  $\cosh(\lambda)$ , ...)
- Flussi su varietà
- Problemi di calcolo scientifico (es. QCD,  $\text{sign}(\lambda)$ )
- (Analisi di) Modelli ridotti di Sistemi Dinamici  
(mediante matrici Grammiane)

⇒ Alcuni esempi in seguito. L'idea:

$$\begin{cases} y' = -Ay \\ y(0) = y_0 \end{cases} \Rightarrow y(t) = \exp(-tA)y_0$$

## Approssimazione Numerica. I

$$f(A)v \approx \tilde{x} \quad \tilde{x} = ???$$

Varie alternative. Tra queste:

- Sostituire  $f$  con funzione più “semplice”,  $f \approx \mathcal{R}$ , per esempio,  $\mathcal{R}$  funzione razionale:

$$\|f(A)v - \tilde{x}\| \leq \|f(A)v - \mathcal{R}(A)v\| + \|\mathcal{R}(A)v - \tilde{x}\|$$

$$\text{e} \quad \Rightarrow \quad \tilde{x} \approx \mathcal{R}(A)v$$

## Approssimazione Numerica. I

$$f(A)v \approx \tilde{x} \quad \tilde{x} = ???$$

Varie alternative. Tra queste:

- Sostituire  $f$  con funzione più “semplice”,  $f \approx \mathcal{R}$ , per esempio,  $\mathcal{R}$  funzione razionale:

$$\|f(A)v - \tilde{x}\| \leq \|f(A)v - \mathcal{R}(A)v\| + \|\mathcal{R}(A)v - \tilde{x}\|$$

e  $\Rightarrow \tilde{x} \approx \mathcal{R}(A)v$

- Approssimazione per proiezione: trova  $V$  e

$$\tilde{x} \in \text{range}(V), \quad \dim \ll n$$

## Approssimazione Numerica. II

$$f(A)v \approx \tilde{x}$$

Le questioni importanti:

- ★ Misure della bontà dell'accuratezza?
- ★ Relazione tra  $f$  e la bontà dell'approssimazione
- ★ Relazione tra  $A$  e la bontà dell'approssimazione
- ★ Efficienza ?

## Approssimazione Razionale

$$x = f(A)v \approx \mathcal{R}_{\mu,\nu}(A)v$$

$$\mathcal{R}_{\mu,\nu}(\lambda) = \frac{\Phi_{\mu}(\lambda)}{\Psi_{\nu}(\lambda)}, \quad \Phi_{\mu}(\lambda), \Psi_{\nu}(\lambda) \text{ polinomi}$$

- Appross. polinomiale,  $\nu = 0$   
(Druskin & Knizhnerman, '89, Bergamaschi & Vianello, '00)
  - Appross. razionale di Padé o Chebyshev, p.e.  $\mu = \nu$
  - Appross. razionale con un polo multiplo
  - Metodi di quadratura (Trefethen et al. )
-

## Approssimazione Razionale

$$x = f(A)v \approx \mathcal{R}_{\mu,\nu}(A)v$$

$$\mathcal{R}_{\mu,\nu}(\lambda) = \frac{\Phi_{\mu}(\lambda)}{\Psi_{\nu}(\lambda)}, \quad \Phi_{\mu}(\lambda), \Psi_{\nu}(\lambda) \text{ polinomi}$$

- Appross. polinomiale,  $\nu = 0$   
(Druskin & Knizhnerman, '89, Bergamaschi & Vianello, '00)
- Appross. razionale di Padé o Chebyshev, p.e.  $\mu = \nu$
- Appross. razionale con un polo multiplo
- Metodi di quadratura (Trefethen et al. )

---

Consideriamo il caso di sviluppo in fratti semplici:

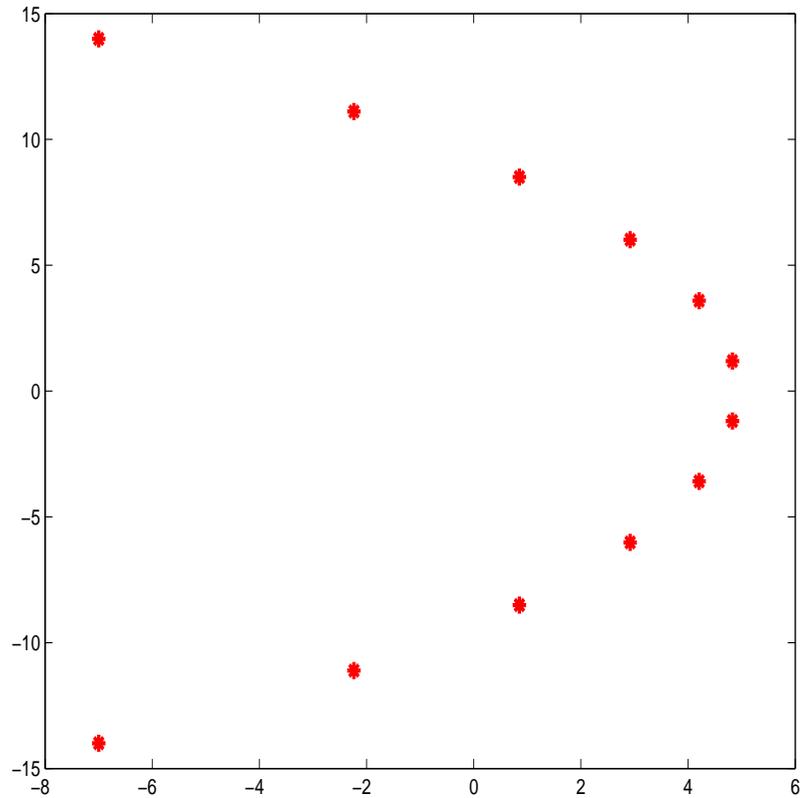
$$\mathcal{R}_{\mu,\nu}(\lambda) = q(\lambda) + \sum_{k=1}^{\nu} \frac{\omega_k}{\lambda - \xi_k} \quad (\mathcal{R}_{\nu} = \mathcal{R}_{\nu,\nu})$$

## Approssimazione Razionale: poli

$$f(\lambda) = \exp(-\lambda)$$

$\mathcal{R}_\nu$ :  $\ell_\infty$  migliore approssimaz.  
in  $[0, \infty)$ , Chebyshev

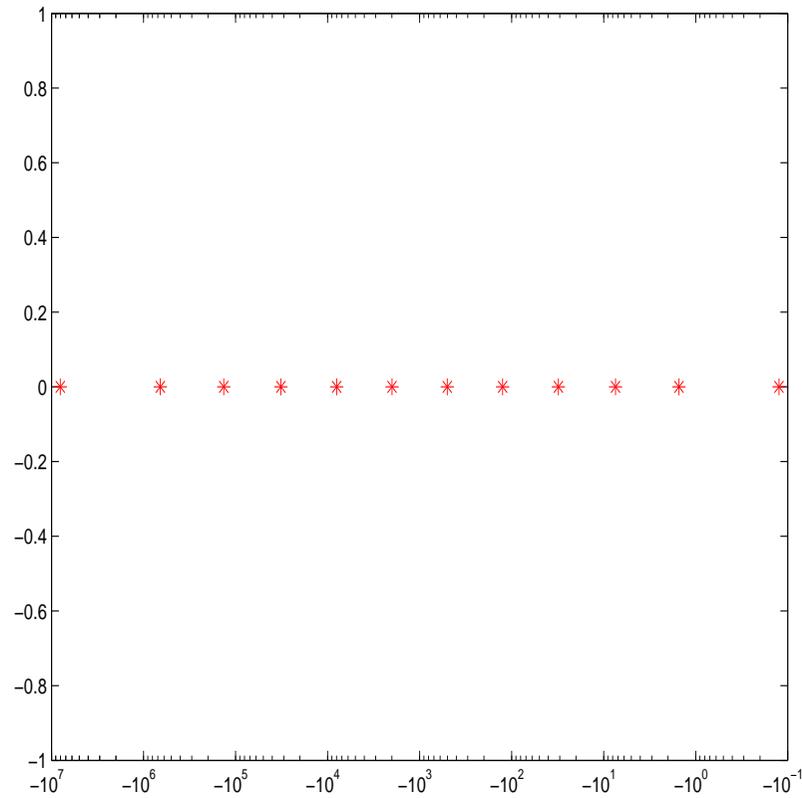
$$\|f - \mathcal{R}_\nu\|_\infty \approx 10^{-\nu}$$



$$f(\lambda) = \lambda^{-1/2}$$

$\mathcal{R}_\nu$ : approssimaz. Zolotarev  
in  $[a, b] \subseteq (0, \infty)$

$$\|f - \mathcal{R}_\nu\| \approx e^{-\pi\sqrt{2\nu}}$$



## Approssimazione Razionale Matriciale

$$\begin{aligned} f(A)v &\approx \mathcal{R}_\nu(A)v = \sum_{k=1}^{\nu} \omega_k (A - \xi_k I)^{-1} v \\ &\approx \sum_{k=1}^{\nu} \omega_k \tilde{x}_k \end{aligned}$$

- $\forall k, (A - \xi_k I)$  matrice “shiftata”,  $\xi_k \in \mathbb{C}$
- $\xi_{2j-1} = \bar{\xi}_{2j}, j = 1, \dots, \lfloor \nu/2 \rfloor$
- $\forall k, \tilde{x}_k$  soluzione approssimata

## Approssimazione Razionale Matriciale

$$\begin{aligned} f(A)v &\approx \mathcal{R}_\nu(A)v = \sum_{k=1}^{\nu} \omega_k (A - \xi_k I)^{-1} v \\ &\approx \sum_{k=1}^{\nu} \omega_k \tilde{x}_k \end{aligned}$$

- $\forall k, (A - \xi_k I)$  matrice “shiftata”,  $\xi_k \in \mathbb{C}$
- $\xi_{2j-1} = \bar{\xi}_{2j}, j = 1, \dots, \lfloor \nu/2 \rfloor$
- $\forall k, \tilde{x}_k$  soluzione approssimata

$\Rightarrow$  **Metodi iterativi per sistemi lineari shiftati**

## Stime dell'errore

$\tilde{x}_k$ : metodi del sottospazio di Krylov:  $\sum_{k=1}^{\nu} \omega_k (A - \xi_k I)^{-1} v \approx \sum_{k=1}^{\nu} \omega_k \tilde{x}_k$

$$\|\mathcal{R}_\nu(A)v - \sum_{k=1}^{\nu} \omega_k \tilde{x}_k\| = ??$$

Stima dell'errore durante l'iterazione : (Frommer & S., '08)

- Stima “dal basso” per  $A$  sdp e poli reali negativi
- Stima per  $A$  reale simmetrica e poli complessi

## Stime dell'errore

$\tilde{x}_k$ : metodi del sottospazio di Krylov:  $\sum_{k=1}^{\nu} \omega_k (A - \xi_k I)^{-1} v \approx \sum_{k=1}^{\nu} \omega_k \tilde{x}_k$

$$\|\mathcal{R}_\nu(A)v - \sum_{k=1}^{\nu} \omega_k \tilde{x}_k\| = ??$$

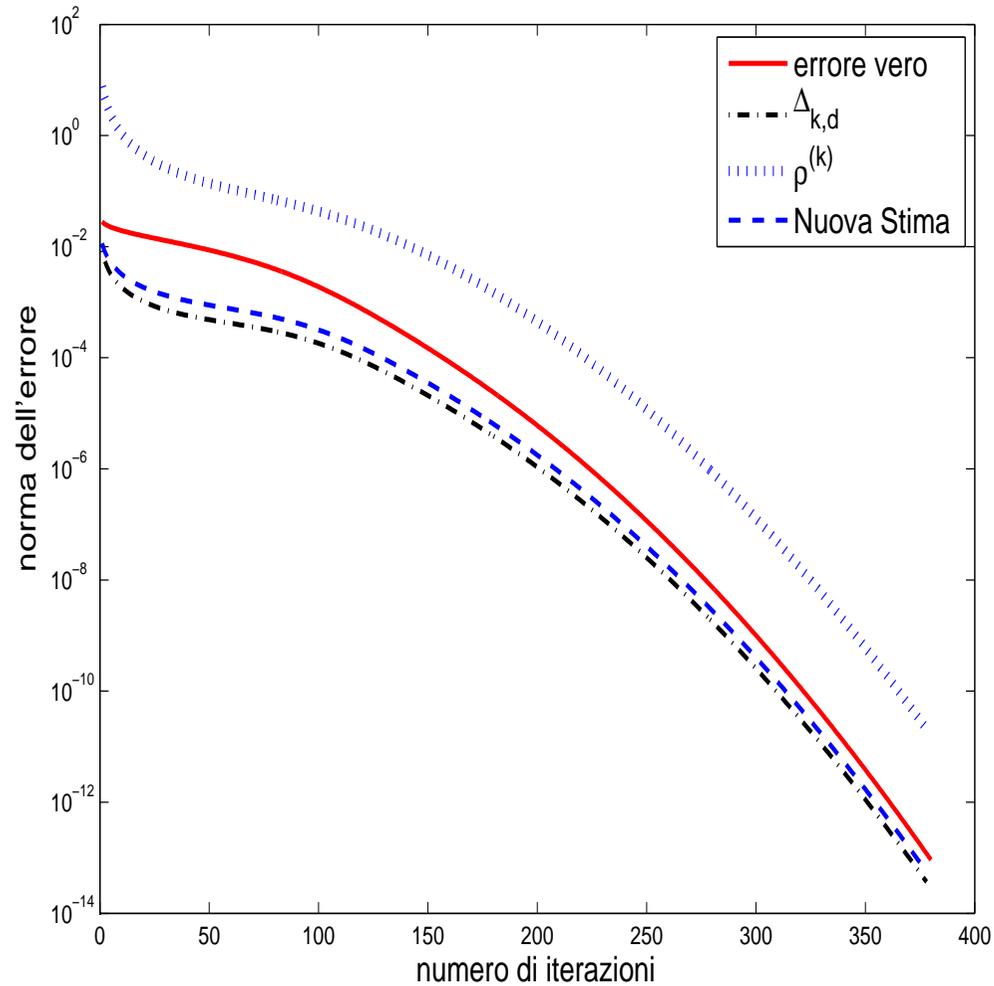
Stima dell'errore durante l'iterazione : (Frommer & S., '08)

- Stima “dal basso” per  $A$  sdp e poli reali negativi
- Stima per  $A$  reale simmetrica e poli complessi

La stima:

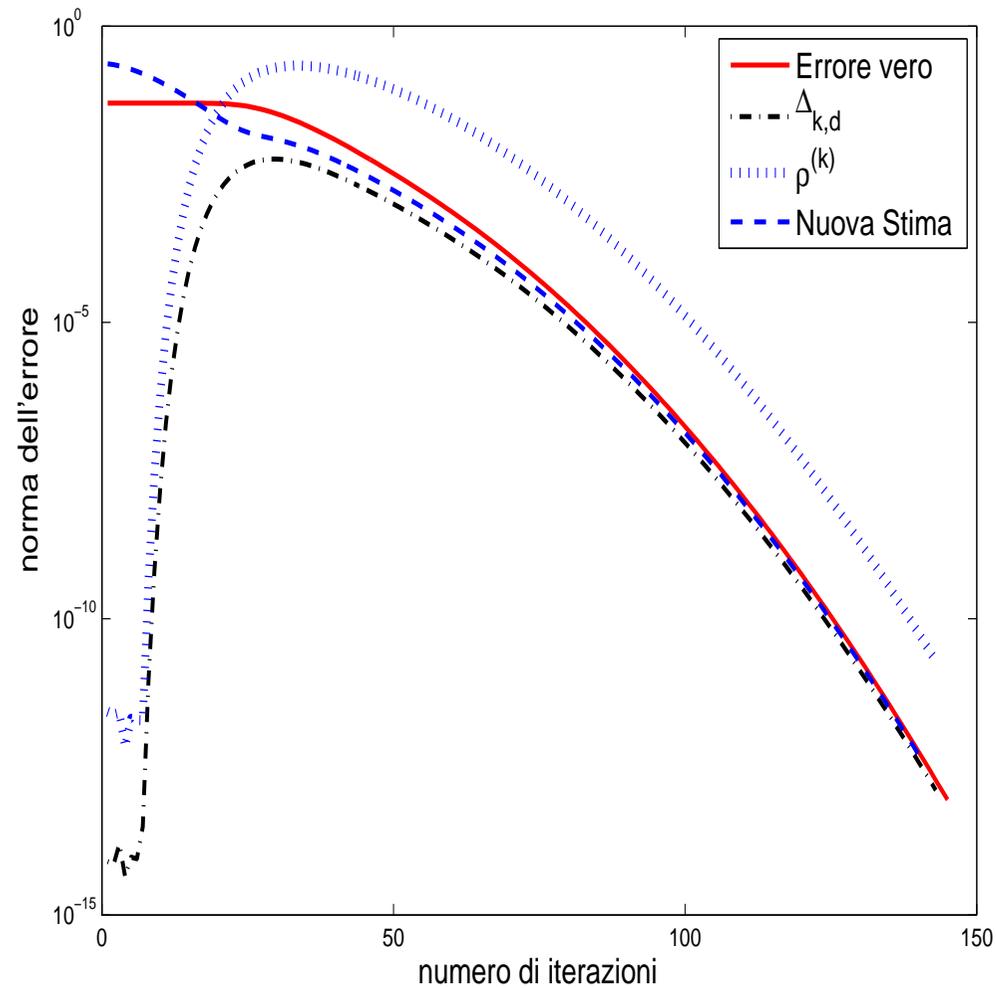
- ★ Non richiede informazioni spettrali
- ★ Il costo computazionale è di 3-5 iterazioni in più

CG per  $A$  sdp e  $f(\lambda) = \lambda^{-\frac{1}{2}}$  :  $A^{-\frac{1}{2}}v$



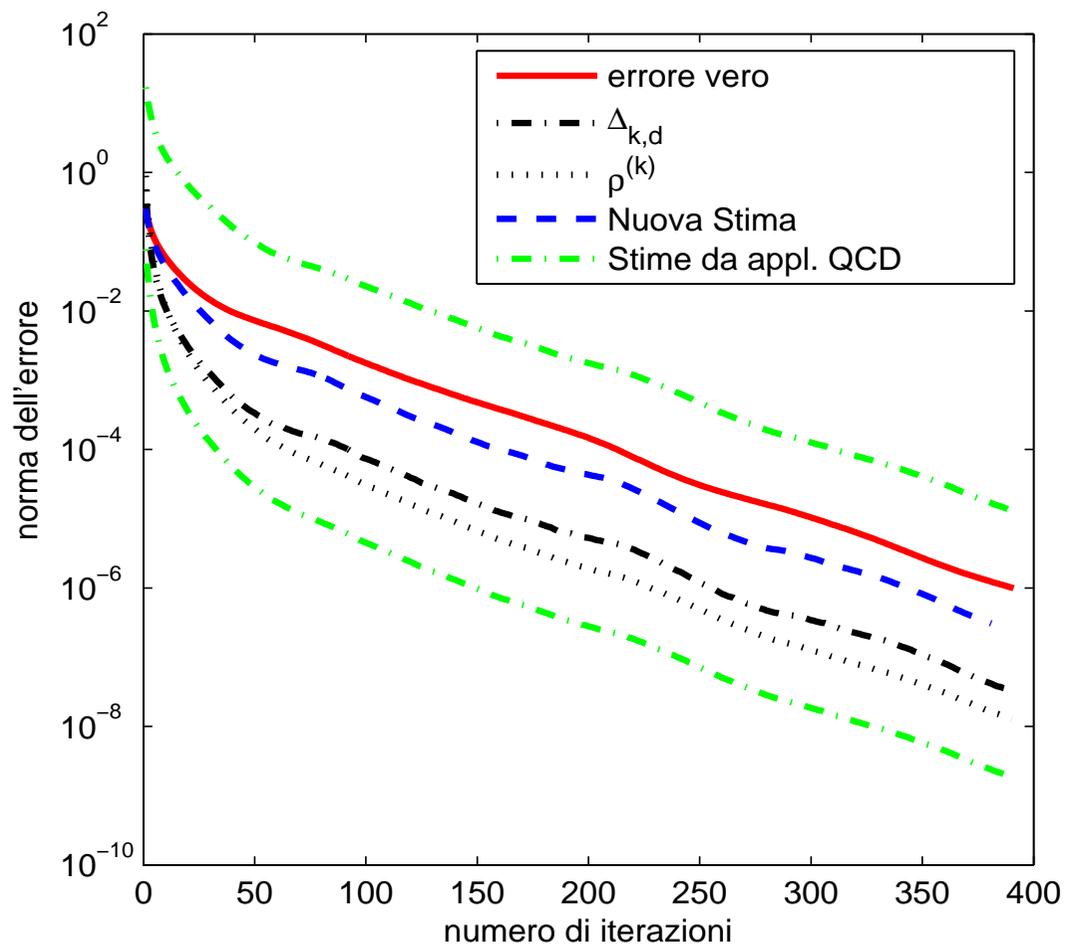
Nota: Convergenza superlineare.

CG “complesso” per  $A$  ssdp e  $f(\lambda) = \exp(-\lambda)$  :  $\exp(-A)v$



Nota: Convergenza superlineare.

CG per  $A$  sdp e  $f(\lambda) = \text{sign}(\lambda) = (\lambda^2)^{-1/2}\lambda$  :  $\text{sign}(A)v$



## Approssimazione mediante spazi di Krylov

$$\mathcal{K}_m \equiv \mathcal{K}_m(A, v) = \text{span}\{v, Av, \dots, A^{m-1}v\}$$

$$V_m \quad \text{t.c.} \quad \text{range}(V_m) = \mathcal{K}_m(A, v) \quad \text{e} \quad V_m^\top V_m = I$$

$\Rightarrow$  **Motivazione:**  $\exists p$  polinomio (interpolante):  $f(A) = p(A)$

## Approssimazione mediante spazi di Krylov

$$\mathcal{K}_m \equiv \mathcal{K}_m(A, v) = \text{span}\{v, Av, \dots, A^{m-1}v\}$$

$$V_m \quad \text{t.c.} \quad \text{range}(V_m) = \mathcal{K}_m(A, v) \quad \text{e} \quad V_m^\top V_m = I$$

$\Rightarrow$  **Motivazione:**  $\exists p$  polinomio (interpolante):  $f(A) = p(A)$

Approccio “classico” (per es. Saad '92):

Per  $H_m = V_m^\top A V_m$ ,  $v = V_m e_1$

$$f(A)v \approx x_m = V_m f(H_m) e_1 \quad \|v\| = 1$$

★  $x_m$  da pb. interpolazione nel senso di Hermite:  $V_m f(H_m) e_1 = p_{m-1}(A)v$

## Krylov vs. Approssimazione razionale

$f \rightarrow \mathcal{R}_\nu$

$$\text{Krylov: } \mathcal{R}_\nu(A)v \approx V_m \mathcal{R}_\nu(H_m)e_1$$

$$\begin{aligned} \mathcal{R}_\nu(A)v &= \omega_0 v + \sum_{j=1}^{\nu} \omega_j (A - \xi_j I)^{-1} v \\ &\approx \omega_0 v + \sum_{j=1}^{\nu} \omega_j V_m (H_m - \xi_j I)^{-1} e_1 \\ &= V_m \left( \omega_0 e_1 + \sum_{j=1}^{\nu} \omega_j (H_m - \xi_j I)^{-1} e_1 \right) \equiv V_m \mathcal{R}_\nu(H_m)e_1 \end{aligned}$$

$V_m \mathcal{R}_\nu(H_m)e_1$  è proiezione di Galerkin termine-a-termine

(van der Vorst, '87, Lopez & S., '06)

## Approssimazione di $\exp(-A)v$ in $\mathcal{K}_m$ . I

Stime di convergenza tipiche (Hochbruck & Lubich '97)

$A$  sim. semidef.

$$\|\exp(-A)v - V_m \exp(-H_m)e_1\| \leq 10e^{-m^2/(5\rho)}, \quad \sqrt{4\rho} \leq m \leq 2\rho$$

$$\|\exp(-A)v - V_m \exp(-H_m)e_1\| \leq \frac{10}{\rho} e^{-\rho} \left(\frac{e\rho}{m}\right)^m, \quad m \geq 2\rho$$

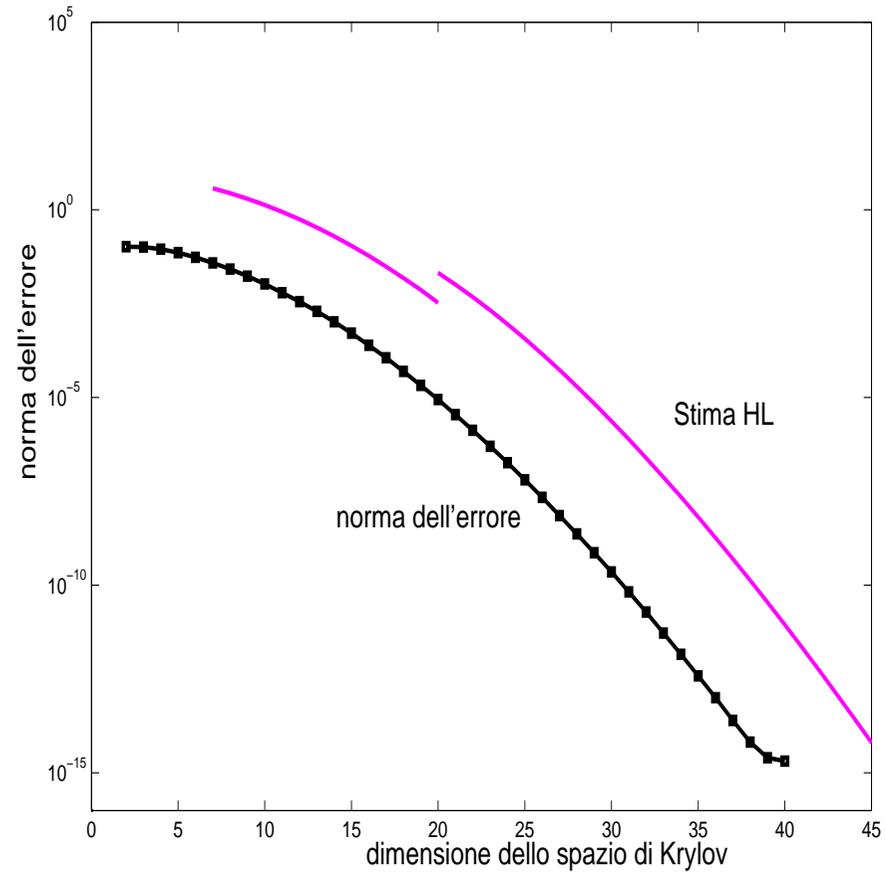
dove  $\sigma(A) \subseteq [0, 4\rho]$

si veda anche Tal-Ezer '89, Druskin & Knizhnerman '89, Stewart & Leyk '96

Altre stime simili per  $\lambda^{-1/2}$ , cosh, ecc.

## Tipico grafico per l'errore

$$\| \exp(-A)v - V_m \exp(-H_m)e_1 \|$$



Convergenza superlineare

## Applicazione. Problema evolutivo

$$\left\{ \begin{array}{l} \frac{\partial u(x,y,t)}{\partial t} = \Delta u, \quad (x,y) \in (0,1)^2 \quad t \in [0,0.1] \\ u(x,y,t) = 0, \quad (x,y) \in \partial([0,1]^2) \\ u(x,y,0) = 1, \quad (x,y) \in [0,1]^2 \end{array} \right.$$

**Eulero implicito:**  $u_{i+1} = (I + \delta t A)^{-1} u_i, \quad i = 0, 1, \dots$

**Integratore esponenziale:**  $u(t) = \exp(-tA)u_0 \quad t = 0.1$

## Applicazione. Problema evolutivo

$$\begin{cases} \frac{\partial u(x,y,t)}{\partial t} = \Delta u, & (x,y) \in (0,1)^2 \quad t \in [0,0.1] \\ u(x,y,t) = 0, & (x,y) \in \partial([0,1]^2) \\ u(x,y,0) = 1, & (x,y) \in [0,1]^2 \end{cases}$$

**Eulero implicito:**  $u_{i+1} = (I + \delta t A)^{-1} u_i, \quad i = 0, 1, \dots$

**Integratore esponenziale:**  $u(t) = \exp(-tA)u_0 \quad t = 0.1$

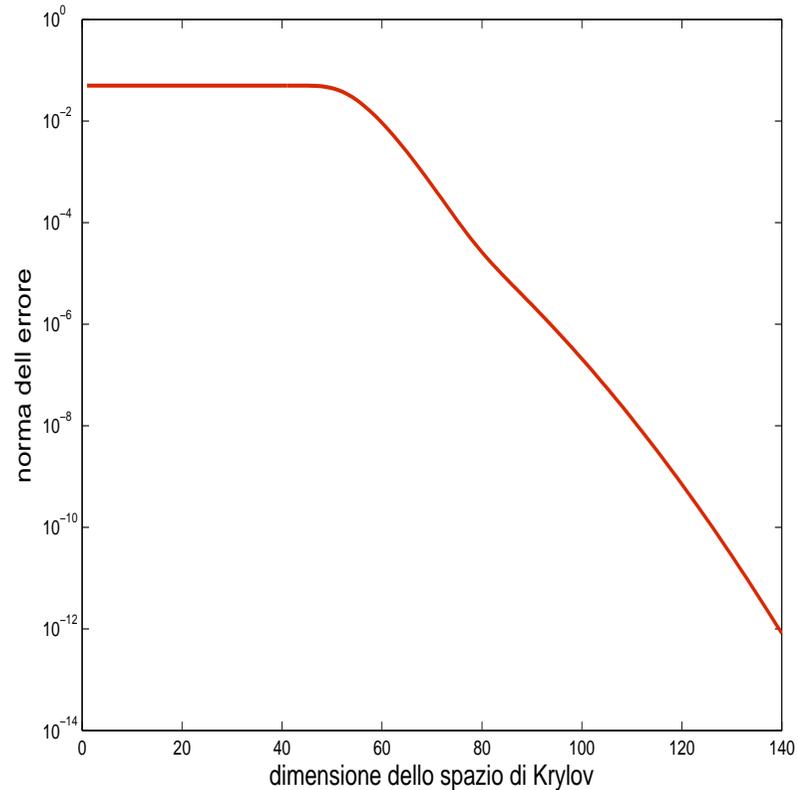
	Eulero		Exp	
passo $\delta t$	CPU	errore	CPU	errore (#its*)
0.001	1.9	$2 \cdot 10^{-3}$	0.09	$9 \cdot 10^{-4}(37)$
0.005	0.4	$1 \cdot 10^{-2}$	0.07	$4 \cdot 10^{-3}(28)$
0.01	0.2	$2 \cdot 10^{-2}$	0.05	$1 \cdot 10^{-2}(25)$

\* : Tolleranza del criterio d'arresto legato al passo temporale

⇒ Integratori esponenziali

...Quando le cose non vanno così bene

$$\| \exp(-A)v - V_m \exp(-H_m)e_1 \| \quad A \in \mathbb{R}^{400 \times 400}, \|A\| = 10^5$$



$$\| \exp(-A)v - V_m \exp(-H_m)e_1 \| \leq 10e^{-m^2/(5\rho)}, \quad \sqrt{4\rho} \leq m \leq 2\rho$$

dove  $\sigma(A) \subseteq [0, 4\rho]$

## Tecniche di Accelerazione

### ★: Miglioramento dello spazio di approssimazione

- Approssimazione spettrale:  $\mathcal{K}_m((I + \gamma A)^{-1}, v)$ ,  $\gamma > 0$

$$f(A)v \approx V_m f\left(\frac{1}{\gamma}(H_m^{-1} - I)\right)e_1$$

(Moret & Novati '04, van den Eshof & Hochbruck, '06, Popolizio & S. '08)

## Tecniche di Accelerazione

### ★: Miglioramento dello spazio di approssimazione

- Approssimazione spettrale:  $\mathcal{K}_m((I + \gamma A)^{-1}, v)$ ,  $\gamma > 0$

$$f(A)v \approx V_m f\left(\frac{1}{\gamma}(H_m^{-1} - I)\right)e_1$$

(Moret & Novati '04, van den Eshof & Hochbruck, '06, Popolizio & S. '08)

- Spazio “esteso”:  $\mathcal{K}_m(A^{-1}, v) \cup \mathcal{K}_m(A, v)$

$$f(A)v \approx \mathcal{V}_m f(\mathcal{T}_m)e_1, \quad \mathcal{T}_m = \mathcal{V}_m^\top A \mathcal{V}_m$$

(Druskin & Knizhnerman, '98, Simoncini, '07)

## Tecniche di Accelerazione

### ★: Miglioramento dello spazio di approssimazione

- Approssimazione spettrale:  $\mathcal{K}_m((I + \gamma A)^{-1}, v)$ ,  $\gamma > 0$

$$f(A)v \approx V_m f\left(\frac{1}{\gamma}(H_m^{-1} - I)\right)e_1$$

(Moret & Novati '04, van den Eshof & Hochbruck, '06, Popolizio & S. '08)

- Spazio “esteso”:  $\mathcal{K}_m(A^{-1}, v) \cup \mathcal{K}_m(A, v)$

$$f(A)v \approx \mathcal{V}_m f(\mathcal{T}_m)e_1, \quad \mathcal{T}_m = \mathcal{V}_m^\top A \mathcal{V}_m$$

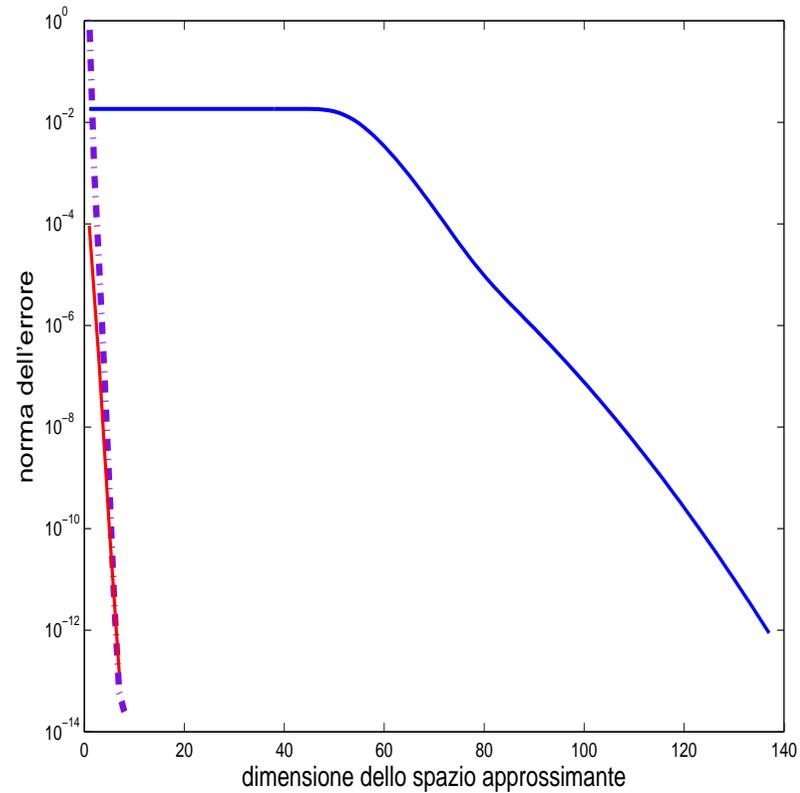
(Druskin & Knizhnerman, '98, Simoncini, '07)

### ★: Rilassamento delle proprietà di ottimalità

- Ortogonalità solo *locale* della base (Eiermann & Ernst '06)
- Limare costi nell'approssimazione di  $\mathcal{R}_\nu(A)v$  (Popolizio & S. '08)

## Accelerazione

$$f(\lambda) = \exp(-\lambda)$$



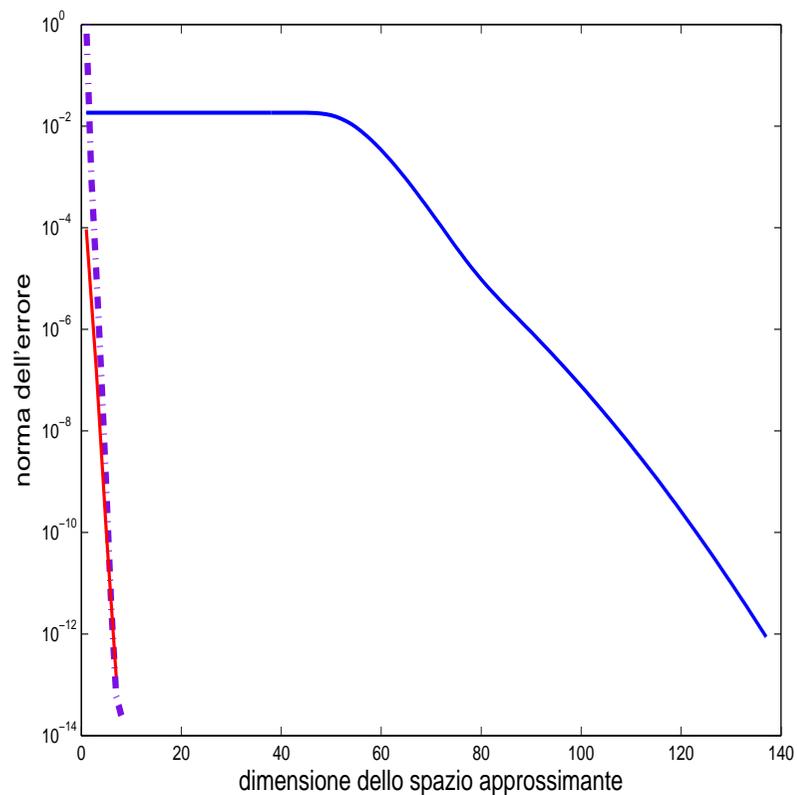
-: Krylov standard

-.: Acc. spettrale

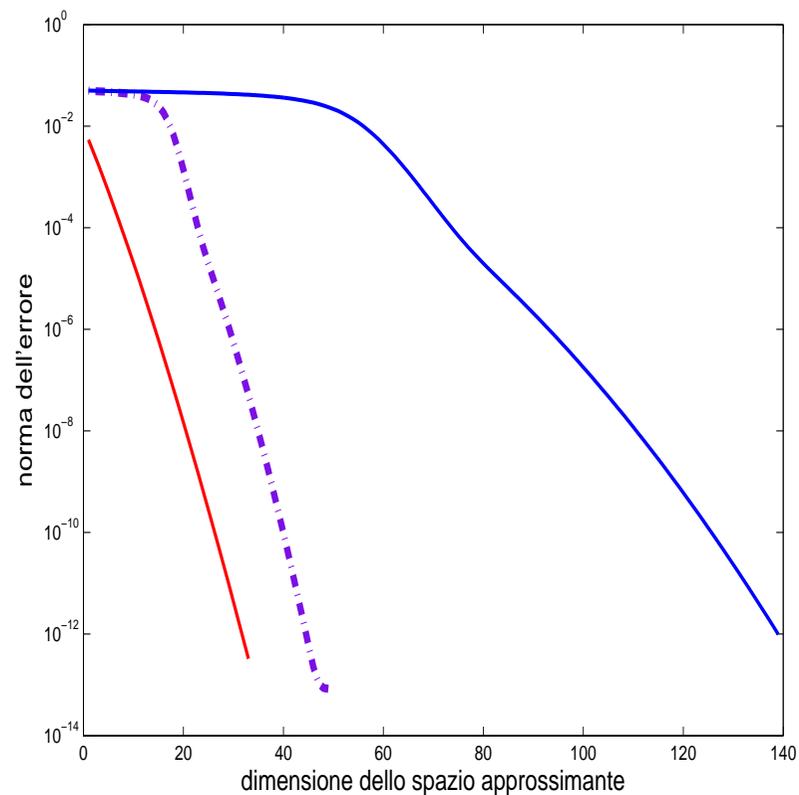
-: Spazio "esteso"

## Accelerazione

$$f(\lambda) = \exp(-\lambda)$$



$$f(\lambda) = \lambda^{-1/2}$$



-: Krylov standard

-.: Acc. spettrale

-: Spazio "esteso"

## Applicazioni. II

Equazione di Lyapunov:

$$AX + XA^T + Q = 0$$

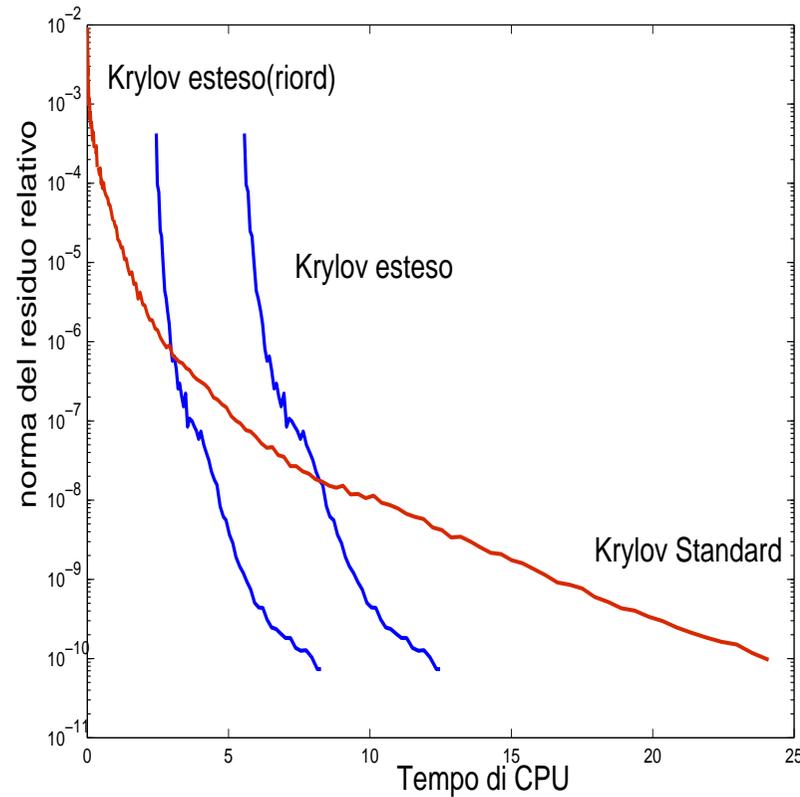
con  $-A$  dissipativa,  $Q = BB^T$  rango basso

$$X = \int_0^{\infty} e^{-tA} BB^T e^{-tA^T} dt = \int_0^{\infty} xx^T dt$$

con  $x = \exp(-tA)B$ .

## Un esempio. Sistemi lineari “Time-invariant”

$$\mathbf{x}' = \mathbf{x}_{xx} + \mathbf{x}_{yy} + \mathbf{x}_{zz} - 10x\mathbf{x}_x - 1000y\mathbf{x}_y - 10\mathbf{x}_z + \mathbf{b}(x, y)\mathbf{u}(t)$$



dim.spazio: 146 (Krylov Standard) 112 (Krylov “esteso”)  $A \in \mathbb{R}^{18^3 \times 18^3}$

Simoncini, '07

## Applicazioni. Problema mal posto. I

$$\left\{ \begin{array}{l} u_{zz} - Lu = 0, \quad (x, y, z) \in \Omega \times [0, z_1] \\ u(x, y, z) = 0, \quad (x, y, z) \in \partial\Omega \times [0, z_1] \\ u(x, y, 0) = g(x, y), \quad (x, y) \in \Omega \\ u_z(x, y, 0) = \mathbf{0}, \quad (x, y) \in \Omega \end{array} \right.$$

$L$  op. ellittico, lineare, autoaggiunto, definito positivo

Pb: determinare  $u$  per  $z = z_1$ :  $f(x, y) = u(x, y, z_1), (x, y) \in \Omega.$

## Applicazioni. Problema mal posto. I

$$\left\{ \begin{array}{l} u_{zz} - Lu = 0, \quad (x, y, z) \in \Omega \times [0, z_1] \\ u(x, y, z) = 0, \quad (x, y, z) \in \partial\Omega \times [0, z_1] \\ u(x, y, 0) = g(x, y), \quad (x, y) \in \Omega \\ u_z(x, y, 0) = \mathbf{0}, \quad (x, y) \in \Omega \end{array} \right.$$

$L$  op. ellittico, lineare, autoaggiunto, definito positivo

Pb: determinare  $u$  per  $z = z_1$ :  $f(x, y) = u(x, y, z_1)$ ,  $(x, y) \in \Omega$ .

---

Separazione delle variabili:  $u(x, y, z) = \cosh(z\sqrt{L})g$

★:  $L$  illimitato  $\Rightarrow \cosh(z\sqrt{L})g$  instabile (rispetto a perturbazioni dei dati)

Eldén & S., in prep.

## Applicazioni. Problema mal posto. II

Regolarizzazione:  $\tilde{g}$  dato perturbato

$$u(x, y, z) = \sum_{k=1}^{\infty} \cosh(\lambda_k z) \langle s_k, g \rangle s_k(x, y)$$

$$\Rightarrow v(x, y, z) = \sum_{\lambda_k \leq \lambda_c} \cosh(\lambda_k z) \langle s_k, \tilde{g} \rangle s_k(x, y)$$

$(\lambda_k^2, s_k)$  autocopie di  $L$

## Applicazioni. Problema mal posto. II

Regolarizzazione:  $\tilde{g}$  dato perturbato

$$u(x, y, z) = \sum_{k=1}^{\infty} \cosh(\lambda_k z) \langle s_k, g \rangle s_k(x, y)$$
$$\Rightarrow v(x, y, z) = \sum_{\lambda_k \leq \lambda_c} \cosh(\lambda_k z) \langle s_k, \tilde{g} \rangle s_k(x, y)$$

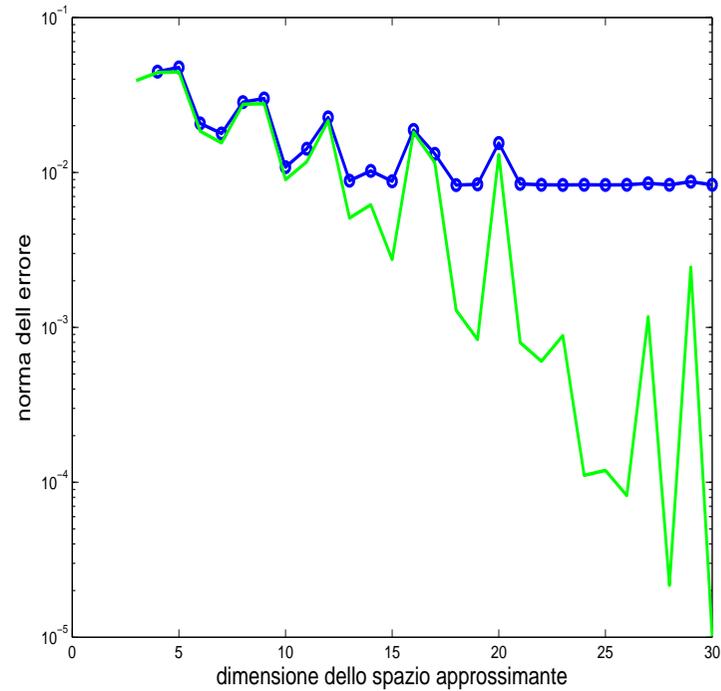
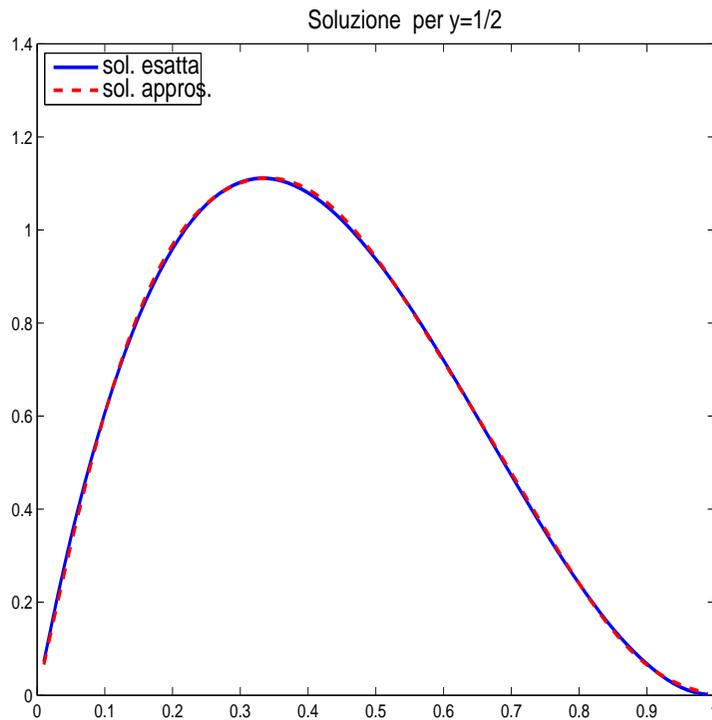
$(\lambda_k^2, s_k)$  autocopie di  $L$

Approssimazione, per esempio, nello spazio di Krylov  $\mathcal{K}_m(L, \tilde{g})$ :

$$u^{(m)}(z) = V_m \cosh(z \sqrt{H_m}) e_1 \|g\| \Rightarrow$$
$$\Rightarrow v^{(m)}(z) = V_m \sum_{\theta_j^{(m)} \leq \lambda_c} y_j^{(m)} \cosh(z \theta_j^{(m)}) (y_j^{(m)})^\top e_1 \|\tilde{g}\|$$

$((\theta_j^{(m)})^2, y_j^{(m)})$  autocopie di  $H_m$

## Applicazioni. Problema mal posto. III



errore funzionale: -  $\|v(z) - v^{(m)}(z)\| \quad z = 0.1$

errore perturb: -  $\|u(z) - v^{(m)}(z)\| \quad z = 0.1$

## Conclusioni

- Uso di  $f(A)v$  ha grandi potenzialità in problemi applicativi
- Si sfrutta il basso costo di  $A$ , rispetto a  $f(A)$
- Ulteriore sviluppo di metodi di accelerazione
- Caso di  $A$  nonsimmetrica