Metodi di accelerazione per il calcolo dell'esponenziale di matrici

Marina Popolizio

Dipartimento di Matematica, Università di Bari

popolizio@dm.uniba.it

In collaborazione con Valeria Simoncini, Università di Bologna

Problema

Approssimare

$$y = \exp(A)v$$

con

- $v \in \mathbb{R}^n, \quad ||v|| = 1$
- A simmetrica, semidefinita negativa, n grande
- $\sigma(A) \subset [\alpha, 0], \qquad |\alpha| \gg 1$

Metodo di Krylov

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

range$(V_m) = \mathcal{K}_m(A, v), \qquad V_m^* V_m = I$

H_m matrice tridiagonale

$$AV_m = V_m H_m + h_{m+1,m} v_{m+1} e_m^*$$

$$y = \exp(A)v \quad \approx \quad V_m \exp(H_m) e_1 \equiv y_m$$

Problemi con il metodo di Krylov:

La convergenza dipende da $\sigma(A)$ e può essere lenta [Hochbruck-Lubich SINUM'97] comportando

- problemi di occupazione di memoria
- costti computazionali elevati

I nostri obiettivi: *Tecniche di accelerazione*

- Tecnica dello *shift-and-invert*
  [Moret-Novati BIT'04, van den Eshof-Hochbruck SISC'06]

- *Real valued method* per risolvere i sistemi derivanti dallo sviluppo in fratti semplici
  [Axelsson-Kucherov, Numer. Linear Algebra Appl., 2000]

Tecnica dello *Shift-and-Invert*

Scelto $s > 0$ applichiamo il metodo di Lanczos a $(I - sA)^{-1}$

$$(I - sA)^{-1} V_m = V_m T_m + \beta_m v_{m+1} e_m^*$$

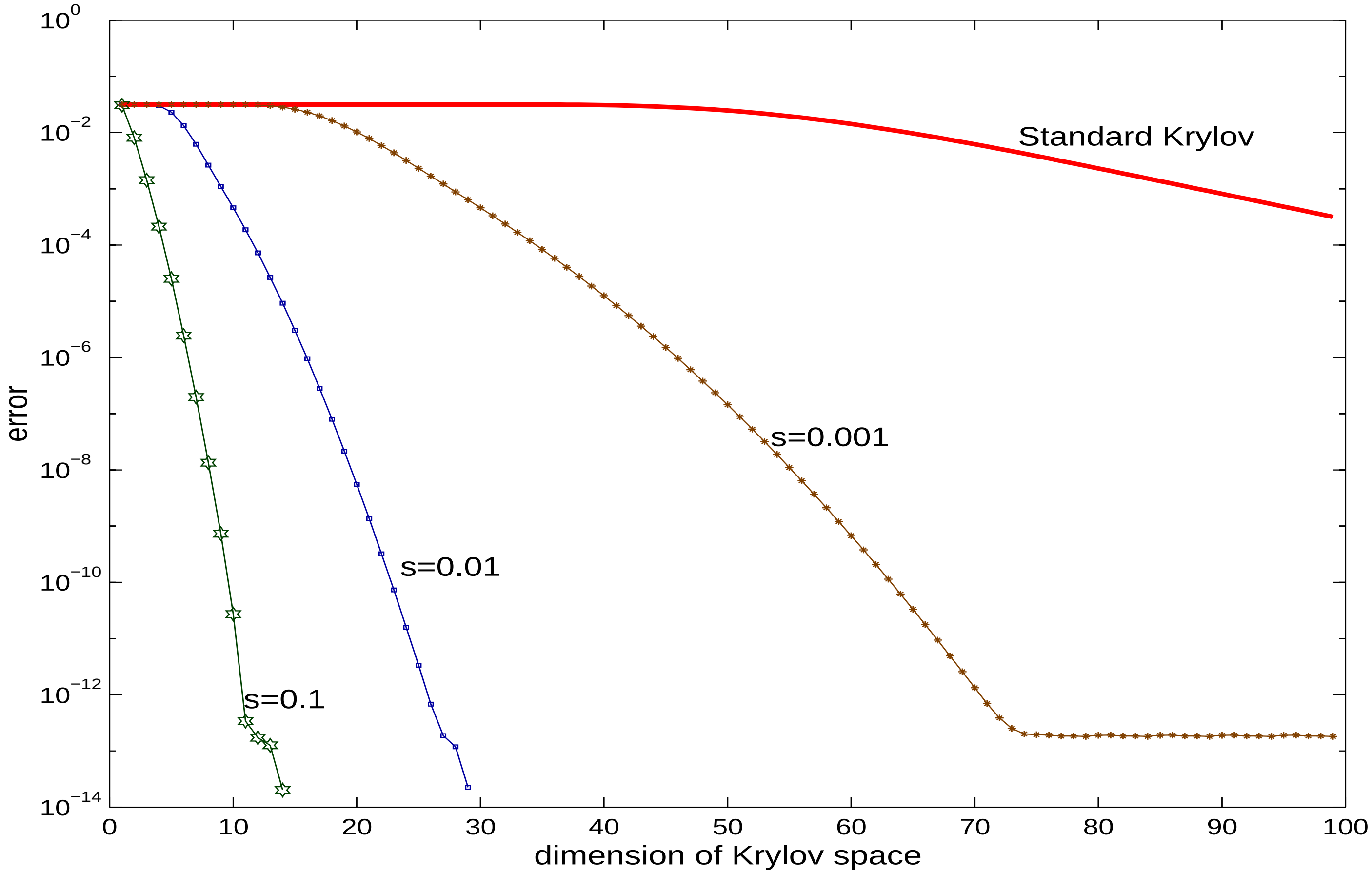
T_m tridiagonale, $V_m^* V_m = I$

$$y = \exp(A)v \quad \approx \quad y_{SI} \equiv V_m \exp\left(\frac{1}{s}(I - T_m^{-1})\right) e_1$$

Moret-Novati BIT'04, Hochbruck-van den Eshof SISC'06

Prove numeriche

A matrice diagonale 1000×1000, $\sigma(A) = \{-9000, -8999, \ldots, 0\}$;
$v \in \mathbb{R}^{1000}, v \propto \mathbf{1},\ ||v|| = 1$



Sviluppo in fratti semplici

Sia f una approssimazione razionale di $\exp$ con ν poli distinti

$$y \approx f(A)v = \tau_0 v + \sum_{j=1}^{\nu} \tau_j (A - \xi_j I)^{-1} v$$

$$(A - \xi_j I)x_j = v \qquad j = 1, \ldots, \nu$$

Equivalenza dei due approcci:

- f approssimazione razionale dell'esponenziale con ν poli distinti

- $y_{prec} = \tau_0 v + \sum_{j=1}^{\nu} \tau_j x_m^{(j)}$

  $x_m^{(j)}$ soluzione di $(A - \frac{1}{s}I)^{-1}(A - \xi_j I)x = (A - \frac{1}{s}I)^{-1}v$

  in $\quad \mathcal{K}_m\big((A - \frac{1}{s}I)^{-1}, v\big)$

- $y_{SI} = V_m f\left(\frac{1}{s}(I - T_m^{-1})\right)e_1$ [Moret-Novati, van den Eshof-Hochbruck]

allora

$$y_{SI} = y_{prec}$$

Scelta del parametro

$$err_m \equiv f(A)v - V_m f(T_m)e_1 = \sum_{j=1}^{\nu} \tau_j \underbrace{\left((A - \xi_j I)^{-1} v - V_m (T_m - \xi_j I)^{-1} e_1\right)}_{e_m(\xi_j)}$$

$$||e_m(\xi_j)|| \le g(\alpha, \xi_j) \frac{1}{\rho_j^m + \frac{1}{\rho_j^m}}$$ [Lopez-Simoncini SINUM'06]

con

$$\rho_j = \gamma_j + \sqrt{\gamma_j^2 - 1} \qquad \gamma_j = \frac{|\alpha - \xi_j| + |\xi_j|}{-\alpha}$$

Per il sistema precondizionato con $A - pI, \quad p = 1/s$:

$$\gamma_j^{prec}(p, \xi_j) = \frac{(p - \alpha)|\xi_j| + p|\alpha - \xi_j|}{-\alpha|p - \xi_j|}$$

Cerchiamo dunque $p = \frac{1}{s}$ tale che

$$\gamma_j^{prec}(p, \xi_j) > \gamma_j$$

Poiché

$$\gamma_j^{prec}(p, \xi_j) > \frac{p+|\xi_j|}{|p-\xi_j|}$$

scegliamo p che risolve

$$\max_{p_1,\ldots,p_\nu} \quad \min_{\xi_1,\ldots,\xi_\nu} \quad \frac{p+|\xi|}{|p-\xi|}$$

$\therefore$ Per ogni grado ν

$$s_{opt} = \frac{1}{p_{opt}} = \frac{1}{|\xi_1|}$$

ξ_1 polo con la parte immaginaria più grande

Valori ottimali del parametro di shift

ν	3	4	5	6	7	8
s_{opt}	0.4134	0.2720	0.1988	0.1551	0.1264	0.1062
ν	9	10	11	12	13	14
s_{opt}	0.0914	0.0801	0.0711	0.0639	0.0580	0.0530
ν	15	16	17	18	19	20
s_{opt}	0.0488	0.0452	0.0421	0.0394	0.0369	0.0348

Approccio di Axelsson-Kucherov I

Il sistema complesso

$$Cu = v$$

con $\quad C = R + \imath S, \;\; u = u_R + \imath u_I, \;\; v = v_R + \imath v_I$

per ogni $p > 0$ è equivalente al sistema reale

$$\begin{pmatrix} C_p & 0 \\ \sqrt{1+p^2}S & -R - pS \end{pmatrix} \begin{pmatrix} u_R \\ \frac{pu_R - u_I}{\sqrt{1+p^2}} \end{pmatrix} = \begin{pmatrix} \tilde{f}_p \\ \frac{v_I - pv_R}{\sqrt{1+p^2}} \end{pmatrix}$$

con

$$C_p = R - pS + (p^2+1)S(R+pS)^{-1}S$$

e

$$\tilde{f}_p = v_R + S(R+pS)^{-1}(v_I - pv_R)$$

Approccio di Axelsson-Kucherov II

Allora

$$C_p u_R = \tilde{f}_p$$

e

$$R u_R - S u_I = b_R$$

Per u_R precondizioniamo con

$$B_p := R - pS$$

e risolviamo

$$M_p u_R = f_p$$

con $f_p = B^{-1} \tilde{f}_p$

Approccio di Axelsson-Kucherov nel nostro caso

1. $y \;\approx\; f(A)v \;=\; \tau_0 v + \sum_{j=1}^{\nu} \tau_j (A - \xi_j I)^{-1} v$

2. $(A - \xi I)u = v$ corrispondente a $\xi = \xi_R + i\xi_I$

Per calcolare $\qquad u_{AK} = u_R + i u_I \approx u$

i) **Fissiamo un parametro** $p > 0$

ii) Poniamo $B = pI - A$

iii) Calcoliamo $f = B^{-1}(v + (p - \xi_R)B^{-1}v)$

iv) Risolviamo $Mu_R = f$ **con** $M = I - 2(p - \xi_R)B^{-1} + |p - \xi|^2 B^{-2}$

v) Calcoliamo $u_I = \frac{1}{\xi_I}((A - \xi_R I)u_R - v)$

Equivalenza degli approcci

Se

- $\xi = \xi_R + \imath \xi_I$
- $p > 0$
- $B = pI - A$
- $M = I - 2(p - \xi_R)B^{-1} + |p - \xi|^2 B^{-2}$
- $f = B^{-1}(v + (p - \xi_R)B^{-1}v)$
- x vettore reale

allora

$M\,x = f$ è l' *equazione normale* di

$$(A - pI)^{-1}(A - \xi I)x = (A - pI)^{-1}v$$

Scelta del parametro

Prop: *M è simmetrica e definita negativa; inoltre, se $\max\{0, \xi_R\} < p < |\xi|$ allora*

$$cond(M) \ \leq \ \frac{|\xi|^2}{\xi_I^2} W_\xi(p)$$

Minimizzando $W_{\xi_j}(p) = \frac{|p-\xi_j|^2}{p^2}$, $\ j = 1, \ldots, \nu$ ricaviamo

$$p_{opt} \ = \ \min\left\{ \min_{1 \leq j \leq \nu} |\xi_j|, \frac{|\xi_\nu|^2}{\Re(\xi_\nu)} \right\}$$

se ξ_ν è il polo con parte reale maggiore.

Esperimenti Numerici

$A \in \mathbb{R}^{N\times N}$, $v \in \mathbb{R}^N$ vettore random, $||v|| = 1$

Approssimiamo $\exp(0.1A)v$

- *PFE*: Calcolo della Partial Fraction Expansion risolvendo esplicitamente ciascun sistema simmetrico complesso
- *Standard Lanczos*: approccio di Lanczos classico
- *AK*: Variante del metodo di Axelsson-Kucherov
- *SI* : Shift-and-Invert Lanczos
- *tol*: accuratezza finale richiesta

Quando possibile le matrici sono riordinate (`symamd` o `symrcm`)

Usando metodi diretti per la risoluzione dei sistemi lineari

N	tol	Standard Lanczos	PFE	AK	Shift-Invert Lanczos
10000	10^{-5}	615 (406)	1.24	1.39 (9)	0.67 (11)
	10^{-8}	610 (406)	1.87	2.53 (25)	0.66 (11)
	10^{-11}	1221 (484)	2.55	4.71 (47)	0.94 (17)
	10^{-14}	- ($>$ 500)	3.20	7.49 (82)	1.24 (23)
15625	10^{-5}	2.69 (89)	30.35	22.05 (11)	11.49 (10)
	10^{-8}	2.95 (93)	51.61	31.60 (27)	11.88 (11)
	10^{-11}	4.76 (113)	69.03	54.68 (51)	14.22 (17)
	10^{-14}	7.25 (130)	90.20	77.31 (84)	16.96 (24)

Tempi e numero di iterate

Usando metodi iterativi per la risoluzione dei sistemi lineari

N	tol	Standard Lanczos	PFE+ QMR (avg its.)	SI+ PCG (out/avg in)	AK+ Iterativo per B (out/avg in)
10000	10^{-5}	615	2.46 (24)	4.4 (11/21)	3.6 (32/10)
	10^{-8}	610	4.92 (35)	5.5 (11/27)	8.6 (27/ 7)
	10^{-11}	1221	8.17 (43)	9.8 (17/32)	17.5 (92/15)
	10^{-14}	-	11.74 (51)	15.4 (13/37)	29.5 (95/13)
15625	10^{-5}	2.69	5.29 (11)	4.05 (10/10)	3.20 (23/7)
	10^{-8}	2.95	9.36 (17)	5.37 (11/13)	5.88 (29/7)
	10^{-11}	4.76	14.29 (22)	8.87 (17/15)	13.42 (74/12)
	10^{-14}	7.25	19.52 (27)	14.39 (24/18)	22.10 (86/12)

Tempi e numero di iterate

Conclusioni

- Abbiamo analizzato due tecniche di accelerazione: SI e AK

  $\therefore$ Esse possono essere ricondotte allo sviluppo in fratti semplici di un'approssimazione razionale dell'esponenziale

- Abbiamo confrontato numericamente diverse tecniche per il calcolo di $\exp(A)v$

  $\therefore$ Non è possibile individuare “la” tecnica migliore

Alcune referenze

- O. Axelsson and A. Kucherov,*Real valued iterative methods for solving complex symmetric linear systems, Numer. Linear Algebra Appl., 2000*
- M. Hochbruck and C. Lubich, *On Krylov subspace approximations to the matrix exponential operator, SIAM J. Numer. Anal., 1997*
- L. Lopez and V. Simoncini, *Analysis of projection methods for rational function approximation to the matrix exponential , SIAM J. Numer. Anal., 2006*
- I. Moret and P. Novati, *RD-Rational Approximations of the Matrix Exponential, BIT, Numerical Mathematics, 2004*
- M. P. and V. Simoncini *Acceleration Techniques for Approximating the Matrix Exponential Operator, preprint, 2006*
- J. van den Eshof and M. Hochbruck, *Preconditioning Lanczos approximations to the matrix exponential , SIAM J. Sci. Comput., 2006*