On the singular values decay of solutions to a class of generalized Sylvester equations and efficient Krylov methods

Davide Palitta*, Elias Jarlebring†, Giampaolo Mele† and Emil Ringh†

*Dipartimento di Matematica - Università di Bologna, Italy (davide.palitta3@unibo.it), †KTH - Royal Institute of Technology, Sweden (eliasj,gmele,eringh@kth.se)

Generalized Sylvester equations

Consider the generalized Sylvester matrix equation

$$AX + XB^T + \sum_{i=1}^{m} N_i X M_i^T = C_1 C_2^T \quad (1)$$

where $A, B, N_i, M_i \in \mathbb{R}^{n\times n}$, for all $i = 1, \dots, m \ll n$, with n (very) large, $C_1, C_2 \in \mathbb{R}^{n\times r}$, $r \ll n$

Applications: MOR for bilinear systems, analysis of linear stochastic differential equations [2], discretization of elliptic PDEs [4]

Further assumptions (not uncommon in applications):

- $\rho(\mathcal{L}^{-1}\Pi) < 1$, $\mathcal{L}(X) := AX + XB^T$, $\Pi(X) := \sum_{i=1}^m N_i X M_i^T$
- If $\mathrm{com}(F, G) := FG - GF$, $F, G \in \mathbb{R}^{n\times n}$, then, for all $i = 1, \dots, m$,

$$\mathrm{com}(A, N_i) = U_i \widetilde{U}_i^T, \quad \mathrm{com}(B, M_i) = Q_i \widetilde{Q}_i^T$$

where $U_i, \widetilde{U}_i \in \mathbb{R}^{n\times s_i}$, $s_i \ll n$, and $Q_i, \widetilde{Q}_i \in \mathbb{R}^{n\times t_i}$, $t_i \ll n$

Closed-form solution and existence of low-rank approximations

The solution X to (1) can be written as

$$X = \sum_{j=0}^{\infty} (-1)^j Y_j$$

$$AY_0 + Y_0 B^T = C_1 C_2^T, \quad AY_j + Y_j B^T = \sum_{i=1}^{m} N_i Y_{j-1} M_i^T, \ j \geq 1$$

Approximation given by

$$X_\ell := \sum_{j=0}^{\ell} (-1)^j Y_j \quad (2)$$

s.t. $\|X - X_\ell\| \leq \|\mathcal{L}^{-1}(C_1 C_2^T)\| \dfrac{\rho(\mathcal{L}^{-1}\Pi)^{\ell+1}}{1 - \rho(\mathcal{L}^{-1}\Pi)}$

Theorem

Let X_ℓ be as in (2). Then there exists a low-rank matrix $\bar{X}_\ell$ s. t.

$$\|X_\ell - \bar{X}_\ell\| \leq \bar{K} e^{-\pi\sqrt{k}}$$

where $\bar{K} > 0$ only depends on $\mathcal{L}$ and ℓ. Moreover,

$$\mathrm{rank}(\bar{X}_\ell) \leq (2k+1)r + \ell(2k+1)^{\ell+1} m^\ell r$$

A new approximation space

For simplicity: $m = 1$ (not restrictive)

- Extended Krylov subspace: effective for (standard) Sylv eqs ($m = 0$) [6]

$\mathbf{EK}_k^{\square}(A, C) := \mathrm{Range}\left\{\left[C, A^{-1}C, \dots, A^{k-1}C, A^{-k}C\right]\right\}$

Lemma

If $\mathrm{com}(A, N) = U\widetilde{U}^T$, $U, \widetilde{U} \in \mathbb{R}^{n\times s}$, $s \ll n$,

$$N \cdot \mathbf{EK}_k^{\square}(A, C_1) \subseteq \mathbf{EK}_k^{\square}(A, [NC_1, U])$$

Compute $\bar{X}_\ell \approx X_\ell$ as $\bar{X}_\ell = S_1 S_2^T$, $S_1, S_2 \in \mathbb{R}^{n\times d}$, $d \ll n$, where

$$S_1 \in \mathbf{EK}_k^{\square}(A, D_L), \quad S_2 \in \mathbf{EK}_k^{\square}(B, D_R)$$

$$D_L := [C_1, NC_1, U, \dots, N^\ell C_1, N^{\ell-1}U], \ D_R := [C_2, MC_2, Q, \dots, M^\ell C_2, M^{\ell-1}Q]$$

In many applications,

- small values of ℓ provide good accuracy
- dim $(\mathrm{Range}\{D_L\}) \ll r + \ell(r+s)$, dim $(\mathrm{Range}\{D_R\}) \ll r + \ell(r+t)$

Projection methods for (1) - Algorithm

Algorithm: Galerkin projection method for generalized Sylv eqs ($m = 1$)

Input: $A, B, N, M \in \mathbb{R}^{n\times n}$, $U \in \mathbb{R}^{n\times s}$, $Q \in \mathbb{R}^{n\times t}$, $C_1, C_2 \in \mathbb{R}^{n\times r}$, $\ell > 0$

Output: $S_1, S_2 \in \mathbb{R}^{n\times d}$, $d \ll n$

1. Set $\beta_1 = \|C_1\|_F$, $\beta_2 = \|C_2\|_F$
2. Set $D_L = [C_1, NC_1, U, \dots, N^\ell C_1, N^{\ell-1}U]$
3. Set $D_R = [C_2, MC_2, Q, \dots, M^\ell C_2, M^{\ell-1}Q]$
4. Perform rank-revealing QR factorization, $D_L = V_1\gamma_L$, $D_R = W_1\gamma_R$
5. Set $\mathcal{V}_1 \equiv V_1$ and $\mathcal{W}_1 \equiv W_1$
6. Set $\theta_L = \gamma_L(1:r, 1:r)$ and $\theta_R = \gamma_R(1:r, 1:r)$
7. **For** $k = 1, 2, \dots,$ **till convergence**
8. Compute next basis blocks $\mathcal{V}_k$, $\mathcal{W}_k$
9. Set $V_k = [V_{k-1}, \mathcal{V}_k]$, $W_k = [W_{k-1}, \mathcal{W}_k]$
10. Update $T_k = V_k^T A V_k$, $H_k = W_k^T B W_k$, $G_k = V_k^T N V_k$, $F_k = W_k^T M W_k$
11. Solve $T_k Z_k + Z_k H_k^T + G_k Z_k F_k^T = E_1 \theta_L \theta_R^T E_1^T$
12. Compute $\|R_k\|_F^2 = \|Z_k E_k h_{k+1}^T\|_F^2 + \|\tau_{k+1} E_k Z_k\|_F^2$
13. If $\|R_k\|_F / \beta_1\beta_2$ is small enough **Stop**
14. **EndDo**
15. Compute the SVD of $Z_k = \Theta\Sigma\Upsilon^T$
16. Set $S_1 = V_k\left(\Theta\sqrt{\Sigma}\right)$, $S_2 = W_k\left(\Upsilon\sqrt{\Sigma}\right)$

Numerical experiments: low-rank Π

Consider

$$AX + XA^T + uv^T X vu^T = CC^T$$

where $A = n^2\,\mathrm{tridiag}(1, -\underline{2}, 1) \in \mathbb{R}^{n\times n}$, $n = 10000$, and $u, v, C \in \mathbb{R}^n$ random vectors. Left and right spaces coincide: $\mathbf{EK}_k^{\square}(A, [C, u])$

`tol`$= 10^{-8}$

	Its.	Memory	rank(X)	Lin. solves	CPU time (s)
BilADI (4 Wach. shifts) [1]	88	72	72	4869	9.50
BilADI (8 $\mathcal{H}_2$-opt. shifts) [1]	60	**71**	71	2752	7.71
GLEK [5]	**7**	332	**57**	1053	14.67
$\mathbf{EK}_k^{\square}(A, [C, u])$	57	228	63	**114**	**3.12**

- For problems with low-rank Π, the proposed method works well even removing the assumption $\rho(\mathcal{L}^{-1}\Pi) < 1$, while other methods do not!

Numerical experiments: MIMO★

Consider

$$AX + XA^T + \gamma^2 N_1 X N_1^T + \gamma^2 N_2 X N_2^T = CC^T$$

where $\gamma = 1/4$, $A = \mathrm{tridiag}(2, -\underline{5}, 2)$, $N_1 = \mathrm{tridiag}(3, \underline{0}, -3) \in \mathbb{R}^{n\times n}$, $n = 10000$, $N_2 = -N_1 + I$, $C = \mathtt{randn}(n, r)$, $r = 2$. `tol`$= 10^{-8}$

- $\mathrm{com}(A, N_1) = -\mathrm{com}(A, N_2) = 12[e_1, e_n][e_1, -e_n]^T$
- $\mathrm{Range}\{[N_1C, N_2C]\} = \mathrm{Range}\{N_1C\}$
- Left and right spaces coincide: $\mathbf{EK}_k^{\square}(A, [C, N_1C, e_1, e_n])$

	Its.	Memory	rank(X)	Lin. solves	CPU time (s)
BilADI (4 Wach. shifts) [1]	34	**95**	95	2452	13.12
BilADI (8 $\mathcal{H}_2$-opt. shifts) [1]	32	**95**	95	2268	13.50
GLEK [5]	29	498	138	4965	29.53
$\mathbf{EK}_k^{\square}(A, [C, N_1C, e_1, e_n])$	**10**	120	**89**	**60**	**3.08**

★ slight modification of Example 2 in [3]

References:

[1] P. Benner and T. Breiten, *Low rank methods for a class of generalized Lyapunov equations and related issues*, Numerische Mathematik, 124(3): 441-470, 2013

[2] P. Benner and T. Damm, *Lyapunov equations, energy functionals, and model order reduction of bilinear and stochastic systems*, SIAM J. Control Optim., 49(2): 686-711, 2011

[3] Y., Lin, L., Bao and Y., Wei, *Order reduction of bilinear MIMO dynamical systems using new block Krylov subspaces*, Comput. Math. Appl., 58(6): 1093-1102, 2009

[4] D. Palitta and V. Simoncini, *Matrix-equation-based strategies for convection-diffusion equations*, BIT Numerical Mathematics, 56(2): 751-776, 2016

[5] S.D., Shank, V. Simoncini and D.B. Szyld, *Efficient low-rank solution of generalized Lyapunov equations*, Numerische Mathematik, 134(2): 327-342, 2016

[6] V. Simoncini, *A new iterative method for solving large-scale Lyapunov matrix equations*, SIAM J. Scient. Computing, 29(3): 1268-1288, 2007

ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA