Krylov methods for large–scale generalized Sylvester equations with low-rank commuting coefficients

Giampaolo Mele

KTH Royal Institute of technology

13 February 2017

Joint work with D. Palitta, E. Ringh, E. Jarlebring

METT 2017, Pisa

Framework

Framework

- Sylvester equations: $AX + XB^T = C_1 C_2^T$

Framework

- Sylvester equations: $AX + XB^T = C_1 C_2^T$
  Low rank solution: $\lambda(A) \cup \lambda(B) \subset \Omega \subset \mathbb{C}_-$
  Methods: Krylov, ADI, etc

Framework

- Sylvester equations: $AX + XB^T = C_1 C_2^T$
  Low rank solution: $\lambda(A) \cup \lambda(B) \subset \Omega \subset \mathbb{C}_-$
  Methods: Krylov, ADI, etc
- Generalized Sylvester equation:

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

Framework

- Sylvester equations: $AX + XB^T = C_1 C_2^T$
  Low rank solution: $\lambda(A) \cup \lambda(B) \subset \Omega \subset \mathbb{C}_-$
  Methods: Krylov, ADI, etc
- Generalized Sylvester equation:

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

  Low rank solution: N_i, M_i low rank and [Benner,Breiten '13]
  Methods: Krylov, ADI, etc [Shank et al '15], [Benner,Damm '11]
  if $\|N_i\| \ll \|A\|$, $\|M_i\| \ll \|B\|$

Framework

- Sylvester equations: $AX + XB^T = C_1 C_2^T$
  Low rank solution: $\lambda(A) \cup \lambda(B) \subset \Omega \subset \mathbb{C}_-$
  Methods: Krylov, ADI, etc
- Generalized Sylvester equation:

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

  Low rank solution: N_i, M_i low rank and [Benner,Breiten '13]
  Methods: Krylov, ADI, etc [Shank et al '15], [Benner,Damm '11]
  if $\|N_i\| \ll \|A\|$, $\|M_i\| \ll \|B\|$

- We consider $(\star)$. Our assumptions:

$$AN_i - N_i A = U_i \tilde{U}_i^T, \qquad BM_i - M_i B = Q_i \tilde{Q}_i^T$$

Outline

- Neumann series expansion
- Krylov method: exploiting the low rank commutation
- Low rank numerical solutions
- Numerical experiments

Neumann series expansion

Neumann series expansion

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

Solution as a Neumann series

Let $\mathcal{L}(X) := AX + XB^T$ and $\Pi(X) := \sum_{i=1}^{m} N_i X M_i^T$. Assume $\|\mathcal{L}^{-1}\Pi\| < 1$, then the unique solution satisfies

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

where $\mathcal{L}(Y_0) = C_1 C_2^T$ and $\mathcal{L}(Y_{j+1}) = \Pi(Y_j)$

Neumann series expansion

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

Solution as a Neumann series

Let $\mathcal{L}(X) := AX + XB^T$ and $\Pi(X) := \sum_{i=1}^{m} N_i X M_i^T$. Assume $\|\mathcal{L}^{-1}\Pi\| < 1$, then the unique solution satisfies

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

where $\mathcal{L}(Y_0) = C_1 C_2^T$ and $\mathcal{L}(Y_{j+1}) = \Pi(Y_j)$

Proof:

$$X = (I + \mathcal{L}^{-1}\Pi)^{-1} \mathcal{L}^{-1}(C_1 C_2^T) = \sum_{j=0}^{\infty} (-1)^j (\mathcal{L}^{-1}\Pi)^j \mathcal{L}^{-1}(C_1 C_2^T)$$

Neumann series expansion

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

Solution as a Neumann series

Let $\mathcal{L}(X) := AX + XB^T$ and $\Pi(X) := \sum_{i=1}^{m} N_i X M_i^T$. Assume $\|\mathcal{L}^{-1}\Pi\| < 1$, then the unique solution satisfies

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

where $\mathcal{L}(Y_0) = C_1 C_2^T$ and $\mathcal{L}(Y_{j+1}) = \Pi(Y_j)$

- Approximation:

$$X_N = \sum_{j=0}^{N} (-1)^j Y_j$$

Neumann series expansion

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

Solution as a Neumann series

Let $\mathcal{L}(X) := AX + XB^T$ and $\Pi(X) := \sum_{i=1}^{m} N_i X M_i^T$. Assume $\|\mathcal{L}^{-1}\Pi\| < 1$, then the unique solution satisfies

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

where $\mathcal{L}(Y_0) = C_1 C_2^T$ and $\mathcal{L}(Y_{j+1}) = \Pi(Y_j)$

- Approximation:
$$X_N = \sum_{j=0}^{N} (-1)^j Y_j$$
- Error:
$$\|X - X_N\| \leq \|\mathcal{L}^{-1}(C)\| \frac{\| \mathcal{L}^{-1} \Pi\|^{N+1}}{1 - \| \mathcal{L}^{-1} \Pi\|}$$

Neumann series expansion

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

Solution as a Neumann series

Let $\mathcal{L}(X) := AX + XB^T$ and $\Pi(X) := \sum_{i=1}^{m} N_i X M_i^T$. Assume $\|\mathcal{L}^{-1}\Pi\| < 1$, then the unique solution is approximated by

$$X_N = \sum_{j=0}^{N} (-1)^j Y_j$$

where

$$AY_0 + Y_0 B^T = C_1 C_2^T$$

$$AY_{j+1} + Y_{j+1} B^T = \sum_{i=1}^{m} N_i Y_j M_i^T$$

Krylov method: exploiting the low rank commutation

Projection method for Sylvester equations

$$AX + XB^T = C_1 C_2^T$$

Given $\mathcal{K}_{k-1} \subset \mathcal{K}_k \subset \mathbb{R}^n, \mathcal{H}_{k-1} \subset \mathcal{H}_k \subset \mathbb{R}^n$ nested subspaces, the approximation is computed as the product of low-rank matrices,

$$X_k = V_k Z_k W_k^T$$

V_k and W_k are orthogonal and s.t. $\text{span}(V_k) = \mathcal{K}_k$, $\text{span}(W_k) = \mathcal{H}_k$.
Z_k satisfy (Galerkin orth. condition)

$$\tilde{A}_k Z_k + Z_k \tilde{B}_k^T = \tilde{C}_1 \tilde{C}_2^T$$

Projection method for Sylvester equations

$$AX + XB^T = C_1 C_2^T$$

Given $\mathcal{K}_{k-1} \subset \mathcal{K}_k \subset \mathbb{R}^n, \mathcal{H}_{k-1} \subset \mathcal{H}_k \subset \mathbb{R}^n$ nested subspaces, the approximation is computed as the product of low-rank matrices,

$$X_k = V_k Z_k W_k^T$$

V_k and W_k are orthogonal and s.t. $\text{span}(V_k) = \mathcal{K}_k$, $\text{span}(W_k) = \mathcal{H}_k$.
Z_k satisfy (Galerkin orth. condition)

$$\tilde{A}_k Z_k + Z_k \tilde{B}_k^T = \tilde{C}_1 \tilde{C}_2^T$$

Our choice: $\mathcal{K}_k = \mathbf{EK}_k^{\square}(A, C_1), \mathcal{H}_k = \mathbf{EK}_k^{\square}(B, C_2)$.

Projection method for Sylvester equations

$$AX + XB^T = C_1 C_2^T$$

Given $\mathcal{K}_{k-1} \subset \mathcal{K}_k \subset \mathbb{R}^n, \mathcal{H}_{k-1} \subset \mathcal{H}_k \subset \mathbb{R}^n$ nested subspaces, the approximation is computed as the product of low-rank matrices,

$$X_k = V_k Z_k W_k^T$$

V_k and W_k are orthogonal and s.t. $\text{span}(V_k) = \mathcal{K}_k$, $\text{span}(W_k) = \mathcal{H}_k$.
Z_k satisfy (Galerkin orth. condition)

$$\tilde{A}_k Z_k + Z_k \tilde{B}_k^T = \tilde{C}_1 \tilde{C}_2^T$$

Our choice: $\mathcal{K}_k = \mathbf{EK}_k^{\square}(A, C_1), \mathcal{H}_k = \mathbf{EK}_k^{\square}(B, C_2)$.

Observation

There are $S_1, S_2 \in \mathbb{C}^{n \times kr}$ s.t.
$\text{span}(S_1) \subseteq \mathbf{EK}_k^{\square}(A, C_1), \text{span}(S_2) \subseteq \mathbf{EK}_k^{\square}(B, C_2)$

$$X_k = S_1 S_2^T$$

Theorem: low rank commuting and Krylov spaces

Consider the generalized Sylvester equation

$$AX + XB^T + NXM^T = C_1 C_2^T$$

such that $\mathrm{com}(A, N) = U\tilde{U}^T$ and $\mathrm{com}(B, M) = Q\tilde{Q}^T$.

Theorem: low rank commuting and Krylov spaces

Consider the generalized Sylvester equation

$$AX + XB^T + NXM^T = C_1 C_2^T$$

such that $\mathrm{com}(A, N) = U\tilde{U}^T$ and $\mathrm{com}(B, M) = Q\tilde{Q}^T$. Let $\tilde{Y}_i$ be the low–rank numerical solution of

$$AY_0 + Y_0 B^T = C_1 C_2^T$$
$$AY_{j+1} + Y_{j+1} B^T = N\tilde{Y}_j M^T,$$

obtained with the Extended Krylov method with k iterations.

Theorem: low rank commuting and Krylov spaces

Consider the generalized Sylvester equation

$$AX + XB^T + NXM^T = C_1 C_2^T$$

such that $\text{com}(A, N) = U\tilde{U}^T$ and $\text{com}(B, M) = Q\tilde{Q}^T$. Let $\tilde{Y}_i$ be the low–rank numerical solution of

$$AY_0 + Y_0 B^T = C_1 C_2^T$$
$$AY_{j+1} + Y_{j+1} B^T = N\tilde{Y}_j M^T,$$

obtained with the Extended Krylov method with k iterations. Let $\tilde{X}_N = \sum_{j=0}^{N} (-1)^j \tilde{Y}_j$.

Theorem: low rank commuting and Krylov spaces

Consider the generalized Sylvester equation

$$AX + XB^T + NXM^T = C_1 C_2^T$$

such that $\text{com}(A, N) = U\tilde{U}^T$ and $\text{com}(B, M) = Q\tilde{Q}^T$. Let $\tilde{Y}_i$ be the low–rank numerical solution of

$$AY_0 + Y_0 B^T = C_1 C_2^T$$
$$AY_{j+1} + Y_{j+1} B^T = N\tilde{Y}_j M^T,$$

obtained with the Extended Krylov method with k iterations. Let $\tilde{X}_N = \sum_{j=0}^{N} (-1)^j \tilde{Y}_j$. Then there exist $S_1 \in \mathbf{EK}^{\square}_{(N+1)k}(A, \hat{C}_1^{(N)})$ and $S_2 \in \mathbf{EK}^{\square}_{(N+1)k}(B, \hat{C}_2^{(N)})$ such that $\tilde{X}_N = S_1 S_2^T$

Theorem: low rank commuting and Krylov spaces

Consider the generalized Sylvester equation

$$AX + XB^T + NXM^T = C_1 C_2^T$$

such that $\text{com}(A, N) = U\tilde{U}^T$ and $\text{com}(B, M) = Q\tilde{Q}^T$. Let $\tilde{Y}_i$ be the low–rank numerical solution of

$$AY_0 + Y_0 B^T = C_1 C_2^T$$
$$AY_{j+1} + Y_{j+1} B^T = N\tilde{Y}_j M^T,$$

obtained with the Extended Krylov method with k iterations. Let $\tilde{X}_N = \sum_{j=0}^{N} (-1)^j \tilde{Y}_j$. Then there exist $S_1 \in \mathbf{EK}^{\square}_{(N+1)k}(A, \hat{C}_1^{(N)})$ and $S_2 \in \mathbf{EK}^{\square}_{(N+1)k}(B, \hat{C}_2^{(N)})$ such that $\tilde{X}_N = S_1 S_2^T$ where

$$\hat{C}_1^{(N)} = [C_1, NC_1, \dots, N^N C_1, U, NU, \dots, N^{N-1} U]$$
$$\hat{C}_2^{(N)} = [C_2, MC_2, \dots, M^N C_2, Q, MQ, \dots, M^{N-1} Q]$$

Sketch/Illustration of the proof

Lemma

Assume that $A \in \mathbb{R}^{n \times n}$ is nonsingular and let $N \in \mathbb{R}^{n \times n}$ such that $\mathrm{com}(A, N) = U\tilde{U}^T$ with $U, \tilde{U} \in \mathbb{R}^{n \times s}$. Let $C \in \mathbb{R}^{n \times r}$, then

$$N \cdot \mathbf{EK}_d^{\square}(A, C) \subseteq \mathbf{EK}_d^{\square}(A, [NC, U])$$

Sketch/Illustration of the proof

Lemma

Assume that $A \in \mathbb{R}^{n \times n}$ is nonsingular and let $N \in \mathbb{R}^{n \times n}$ such that $\mathrm{com}(A, N) = U\tilde{U}^T$ with $U, \tilde{U} \in \mathbb{R}^{n \times s}$. Let $C \in \mathbb{R}^{n \times r}$, then

$$N \cdot \mathbf{EK}_d^{\square}(A, C) \subseteq \mathbf{EK}_d^{\square}(A, [NC, U])$$

$$AY_0 + Y_0 B^T = C_1 C_2^T$$

Sketch/Illustration of the proof

Lemma

Assume that $A \in \mathbb{R}^{n \times n}$ is nonsingular and let $N \in \mathbb{R}^{n \times n}$ such that $\text{com}(A, N) = U\tilde{U}^T$ with $U, \tilde{U} \in \mathbb{R}^{n \times s}$. Let $C \in \mathbb{R}^{n \times r}$, then

$$N \cdot \mathbf{EK}_d^{\square}(A, C) \subseteq \mathbf{EK}_d^{\square}(A, [NC, U])$$

$$AY_0 + Y_0 B^T = C_1 C_2^T \xrightarrow[\text{Krylov}]{\text{Extended}}$$

Sketch/Illustration of the proof

Lemma

Assume that $A \in \mathbb{R}^{n \times n}$ is nonsingular and let $N \in \mathbb{R}^{n \times n}$ such that $\text{com}(A, N) = U\tilde{U}^T$ with $U, \tilde{U} \in \mathbb{R}^{n \times s}$. Let $C \in \mathbb{R}^{n \times r}$, then

$$N \cdot \mathbf{EK}_d^{\square}(A, C) \subseteq \mathbf{EK}_d^{\square}(A, [NC, U])$$

$$AY_0 + Y_0 B^T = C_1 C_2^T \xrightarrow[\text{Krylov}]{\text{Extended}} Y_0 = F_0 R_0^T, \quad \begin{array}{l} F_0 \in \mathbf{EK}_k^{\square}(A, C_1) \\ R_0 \in \mathbf{EK}_k^{\square}(B, C_2) \end{array}$$

Sketch/Illustration of the proof

Lemma

Assume that $A \in \mathbb{R}^{n \times n}$ is nonsingular and let $N \in \mathbb{R}^{n \times n}$ such that $\text{com}(A, N) = U \tilde{U}^T$ with $U, \tilde{U} \in \mathbb{R}^{n \times s}$. Let $C \in \mathbb{R}^{n \times r}$, then

$$N \cdot \mathbf{EK}_d^{\square}(A, C) \subseteq \mathbf{EK}_d^{\square}(A, [NC, U])$$

$$AY_0 + Y_0 B^T = C_1 C_2^T \xrightarrow[\text{Krylov}]{\text{Extended}} Y_0 = F_0 R_0^T, \quad \begin{matrix} F_0 \in \mathbf{EK}_k^{\square}(A, C_1) \\ R_0 \in \mathbf{EK}_k^{\square}(B, C_2) \end{matrix}$$

$$AY_1 + Y_1 B^T = N Y_0 M^T$$

Sketch/Illustration of the proof

Lemma

Assume that $A \in \mathbb{R}^{n \times n}$ is nonsingular and let $N \in \mathbb{R}^{n \times n}$ such that $\text{com}(A, N) = U\tilde{U}^T$ with $U, \tilde{U} \in \mathbb{R}^{n \times s}$. Let $C \in \mathbb{R}^{n \times r}$, then

$$N \cdot \mathbf{EK}_d^{\square}(A, C) \subseteq \mathbf{EK}_d^{\square}(A, [NC, U])$$

$$AY_0 + Y_0 B^T = C_1 C_2^T \xrightarrow[\text{Krylov}]{\text{Extended}} Y_0 = F_0 R_0^T, \quad \begin{array}{l} F_0 \in \mathbf{EK}_k^{\square}(A, C_1) \\ R_0 \in \mathbf{EK}_k^{\square}(B, C_2) \end{array}$$

$$AY_1 + Y_1 B^T = NY_0 M^T = (NF_0)(MR_0)^T$$

Sketch/Illustration of the proof

Lemma

Assume that $A \in \mathbb{R}^{n \times n}$ is nonsingular and let $N \in \mathbb{R}^{n \times n}$ such that $\text{com}(A, N) = U\tilde{U}^T$ with $U, \tilde{U} \in \mathbb{R}^{n \times s}$. Let $C \in \mathbb{R}^{n \times r}$, then

$$N \cdot \mathbf{EK}_d^{\square}(A, C) \subseteq \mathbf{EK}_d^{\square}(A, [NC, U])$$

$$AY_0 + Y_0 B^T = C_1 C_2^T \xrightarrow[\text{Krylov}]{\text{Extended}} Y_0 = F_0 R_0^T, \quad \begin{array}{l} F_0 \in \mathbf{EK}_k^{\square}(A, C_1) \\ R_0 \in \mathbf{EK}_k^{\square}(B, C_2) \end{array}$$

$$AY_1 + Y_1 B^T = NY_0 M^T = (NF_0)(MR_0)^T \xrightarrow[\text{Krylov}]{\text{Extended}}$$

Sketch/Illustration of the proof

Lemma

Assume that $A \in \mathbb{R}^{n \times n}$ is nonsingular and let $N \in \mathbb{R}^{n \times n}$ such that $\text{com}(A, N) = U\tilde{U}^T$ with $U, \tilde{U} \in \mathbb{R}^{n \times s}$. Let $C \in \mathbb{R}^{n \times r}$, then

$$N \cdot \mathbf{EK}_d^{\square}(A, C) \subseteq \mathbf{EK}_d^{\square}(A, [NC, U])$$

$$AY_0 + Y_0 B^T = C_1 C_2^T \xrightarrow[\text{Krylov}]{\text{Extended}} Y_0 = F_0 R_0^T, \quad \begin{matrix} F_0 \in \mathbf{EK}_k^{\square}(A, C_1) \\ R_0 \in \mathbf{EK}_k^{\square}(B, C_2) \end{matrix}$$

$$AY_1 + Y_1 B^T = N Y_0 M^T = (NF_0)(MR_0)^T \xrightarrow[\text{Krylov}]{\text{Extended}}$$

$$Y_1 = F_1 R_1^T, \quad \begin{matrix} F_1 \in \mathbf{EK}_k^{\square}(A, NF_0) \\ R_1 \in \mathbf{EK}_k^{\square}(B, MR_0) \end{matrix}$$

Sketch/Illustration of the proof

Lemma

Assume that $A \in \mathbb{R}^{n \times n}$ is nonsingular and let $N \in \mathbb{R}^{n \times n}$ such that $\text{com}(A, N) = U\tilde{U}^T$ with $U, \tilde{U} \in \mathbb{R}^{n \times s}$. Let $C \in \mathbb{R}^{n \times r}$, then

$$N \cdot \mathbf{EK}_d^{\square}(A, C) \subseteq \mathbf{EK}_d^{\square}(A, [NC, U])$$

$$AY_0 + Y_0 B^T = C_1 C_2^T \xrightarrow[\text{Krylov}]{\text{Extended}} Y_0 = F_0 R_0^T, \quad \begin{array}{l} F_0 \in \mathbf{EK}_k^{\square}(A, C_1) \\ R_0 \in \mathbf{EK}_k^{\square}(B, C_2) \end{array}$$

$$AY_1 + Y_1 B^T = NY_0 M^T = (NF_0)(MR_0)^T \xrightarrow[\text{Krylov}]{\text{Extended}}$$

$$Y_1 = F_1 R_1^T, \quad \begin{array}{l} F_1 \in \mathbf{EK}_k^{\square}(A, NF_0) \\ R_1 \in \mathbf{EK}_k^{\square}(B, MR_0) \end{array}$$

$$NF_0 \in N \cdot \mathbf{EK}_k^{\square}(A, C_1)$$
$$MR_0 \in M \cdot \mathbf{EK}_k^{\square}(B, C_2)$$

Sketch/Illustration of the proof

Lemma

Assume that $A \in \mathbb{R}^{n \times n}$ is nonsingular and let $N \in \mathbb{R}^{n \times n}$ such that $\mathrm{com}(A, N) = U\tilde{U}^T$ with $U, \tilde{U} \in \mathbb{R}^{n \times s}$. Let $C \in \mathbb{R}^{n \times r}$, then

$$N \cdot \mathbf{EK}_d^{\square}(A, C) \subseteq \mathbf{EK}_d^{\square}(A, [NC, U])$$

$$AY_0 + Y_0 B^T = C_1 C_2^T \xrightarrow[\text{Krylov}]{\text{Extended}} Y_0 = F_0 R_0^T, \quad \begin{matrix} F_0 \in \mathbf{EK}_k^{\square}(A, C_1) \\ R_0 \in \mathbf{EK}_k^{\square}(B, C_2) \end{matrix}$$

$$AY_1 + Y_1 B^T = NY_0 M^T = (NF_0)(MR_0)^T \xrightarrow[\text{Krylov}]{\text{Extended}}$$

$$Y_1 = F_1 R_1^T, \quad \begin{matrix} F_1 \in \mathbf{EK}_k^{\square}(A, NF_0) \\ R_1 \in \mathbf{EK}_k^{\square}(B, MR_0) \end{matrix}$$

$$\begin{matrix} NF_0 \in N \cdot \mathbf{EK}_k^{\square}(A, C_1) & \subseteq \mathbf{EK}_k^{\square}(A, [NC_1, U]) \\ MR_0 \in M \cdot \mathbf{EK}_k^{\square}(B, C_2) & \subseteq \mathbf{EK}_k^{\square}(B, [MC_2, Q]) \end{matrix}$$

Projection method for generalized Sylvester equations

$$AX + XB^T + NXM^T = C_1 C_2^T$$

Given $\mathcal{K}_{k-1} \subset \mathcal{K}_k \subset \mathbb{R}^n, \mathcal{H}_{k-1} \subset \mathcal{H}_k \subset \mathbb{R}^n$ nested subspaces, the approximation is computed as the product of low-rank matrices,

$$X_k = V_k Z_k W_k^T$$

V_k and W_k are orthogonal and s.t. $\text{span}(V_k) = \mathcal{K}_k$, $\text{span}(W_k) = \mathcal{H}_k$.
Z_k satisfy (Galerkin orth. condition)

$$\tilde{A}_k Z_k + Z_k \tilde{B}_k^T + \tilde{N}_k Z_k \tilde{M}_k^T = \tilde{C}_1 \tilde{C}_2^T$$

$$\mathcal{K}_k = \mathbf{EK}_d^{\square}(A, [C_1, NC_1, \dots, N^N C_1, U, NU, \dots, N^{N-1} U])$$
$$\mathcal{H}_k = \mathbf{EK}_d^{\square}(B, [C_2, MC_2, \dots, M^N C_2, Q, MQ, \dots, M^{N-1} Q])$$

Low rank numerical solutions

Low rank approximations

Let $\mathcal{L}(X) := AX + XB^T$, $\Pi(X) := \sum_{i=1}^{m} N_i X M_i^T$ and $C_1, C_2 \in \mathbb{C}^{n \times r}$

Low rank approximations

Let $\mathcal{L}(X) := AX + XB^T$, $\Pi(X) := \sum_{i=1}^m N_i X M_i^T$ and $C_1, C_2 \in \mathbb{C}^{n \times r}$

Theorem [Grasedyck '04]: low rank Sylvester eq.

Let $\mathcal{L}(X) = C_1 C_2^T$. Then there exists an $\bar{X}$ such that

$$\text{rank}(\bar{X}) \leq (2k+1)r$$
$$\|X - \bar{X}\| \leq K(\mathcal{L}) e^{-\pi\sqrt{k}}$$

Low rank approximations

Let $\mathcal{L}(X) := AX + XB^T$, $\Pi(X) := \sum_{i=1}^{m} N_i X M_i^T$ and $C_1, C_2 \in \mathbb{C}^{n \times r}$

Theorem [Grasedyck '04]: low rank Sylvester eq.

Let $\mathcal{L}(X) = C_1 C_2^T$. Then there exists an $\bar{X}$ such that

$$\operatorname{rank}(\bar{X}) \le (2k+1)r$$
$$\left\| X - \bar{X} \right\| \le K(\mathcal{L}) e^{-\pi\sqrt{k}}$$

Theorem: low rank generalized Sylvester eq.

Let X_N be the matrix obtained by truncating the Neumann series. Then there exists an $\bar{X}_N$ such that

$$\operatorname{rank}(\bar{X}_N) \le (2k+1)r + N(2k+1)^{N+1} m^N r$$
$$\left\| X_N - \bar{X}_N \right\| \le K(\mathcal{L}, N) e^{-\pi\sqrt{k}}$$

Low rank approximations

Let $\mathcal{L}(X) := AX + XB^T$, $\Pi(X) := \sum_{i=1}^{m} N_i X M_i^T$ and $C_1, C_2 \in \mathbb{C}^{n \times r}$

Theorem [Grasedyck '04]: low rank Sylvester eq.

Let $\mathcal{L}(X) = C_1 C_2^T$. Then there exists an $\bar{X}$ such that

$$\operatorname{rank}(\bar{X}) \leq (2k+1)r$$
$$\|X - \bar{X}\| \leq K(\mathcal{L}) e^{-\pi\sqrt{k}}$$

Theorem: low rank generalized Sylvester eq.

Let X_N be the matrix obtained by truncating the Neumann series. Then there exists an $\bar{X}_N$ such that

$$\operatorname{rank}(\bar{X}_N) \leq (2k+1)r + N(2k+1)^{N+1} m^N r$$
$$\|X_N - \bar{X}_N\| \leq K(\mathcal{L}, N) e^{-\pi\sqrt{k}}$$

Similar result for $\Pi(X)$ low rank [Benner,Breiten '13]

Numerical experiments

MIMO: multiple input multiple output

Application: bilinear systems (stability)

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

MIMO: multiple input multiple output

Application: bilinear systems (stability)

$$AX + XA^T + \gamma^2(N_1XN_1^T + N_2XN_2^T) = CC^T$$

- $\gamma > 0$ small

MIMO: multiple input multiple output

Application: bilinear systems (stability)

$$AX + XA^T + \gamma^2(N_1XN_1^T + N_2XN_2^T) = CC^T$$

- $\gamma > 0$ small
- $$A = \begin{pmatrix} -5 & 2 & & \\ 2 & \ddots & \ddots & \\ & \ddots & & 2 \\ & & 2 & -5 \end{pmatrix} \qquad N_1 = \begin{pmatrix} 0 & -3 & & \\ 3 & \ddots & \ddots & \\ & \ddots & & -1 \\ & & 3 & 0 \end{pmatrix}$$

MIMO: multiple input multiple output

Application: bilinear systems (stability)

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

- $\gamma > 0$ small
- $$A = \begin{pmatrix} -5 & 2 & & \\ 2 & \ddots & \ddots & \\ & \ddots & & 2 \\ & & 2 & -5 \end{pmatrix} \qquad N_1 = \begin{pmatrix} 0 & -3 & & \\ 3 & \ddots & \ddots & \\ & \ddots & & -1 \\ & & 3 & 0 \end{pmatrix}$$
- $N_2 = -N_1 + I$

MIMO: multiple input multiple output

Application: bilinear systems (stability)

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

- $\gamma > 0$ small
- $$A = \begin{pmatrix} -5 & 2 & & \\ 2 & \ddots & \ddots & \\ & \ddots & & 2 \\ & & 2 & -5 \end{pmatrix} \qquad N_1 = \begin{pmatrix} 0 & -3 & & \\ 3 & \ddots & \ddots & \\ & \ddots & & -1 \\ & & 3 & 0 \end{pmatrix}$$
- $N_2 = -N_1 + I$
- $\text{com}(A, N_1) = -\text{com}(A, N_2) = 12[e_1, e_n][e_1, -e_n]^T$

MIMO: multiple input multiple output

Application: bilinear systems (stability)

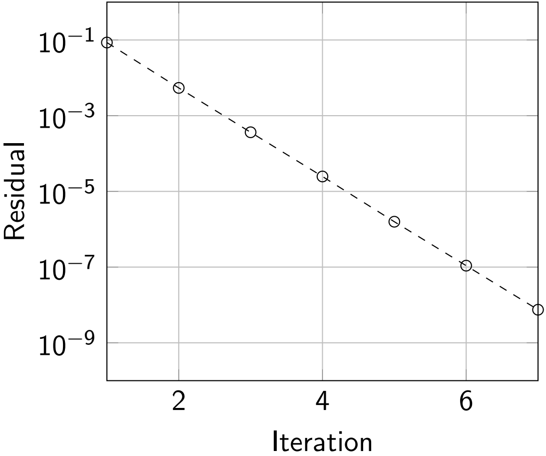
$$AX + XA^T + \gamma^2(N_1 X N_1^T + N_2 X N_2^T) = CC^T$$

- $\gamma > 0$ small
- $$A = \begin{pmatrix} -5 & 2 & & \\ 2 & \ddots & \ddots & \\ & \ddots & & 2 \\ & & 2 & -5 \end{pmatrix} \qquad N_1 = \begin{pmatrix} 0 & -3 & & \\ 3 & \ddots & \ddots & \\ & \ddots & & -1 \\ & & 3 & 0 \end{pmatrix}$$
- $N_2 = -N_1 + I$
- $\text{com}(A, N_1) = -\text{com}(A, N_2) = 12[e_1, e_n][e_1, -e_n]^T$
- $\mathbf{EK}_d^{\square}(A, [C, N_1 C, [e_1, e_n]])$

MIMO: multiple input multiple output

Application: bilinear systems (stability)

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



MIMO: comparison with other methods

	γ	Its.	Memory	rank(X)	Lin. solves
Ext. Krylov (low rank-comm)	1/6	8	7.32MB	64	48
BilADI [1] (4 Wach. shifts)	1/6	15	5.18MB	68	591
BilADI (8 $\mathcal{H}_2$-opt. shifts)	1/6	14	5.18MB	68	522
GLEK [2]	1/6	13	16.78MB	52	1549
Ext. Krylov (low rank-comm)	1/5	8	7.32MB	72	48
BilADI (4 Wach. shifts)	1/5	20	5.95MB	78	990
BilADI (8 $\mathcal{H}_2$-opt. shifts)	1/5	20	5.95MB	78	987
GLEK	1/5	17	20.30MB	59	2309
Ext. Krylov (low rank-comm)	1/4	10	9.16MB	89	60
BilADI (4 Wach. shifts)	1/4	30	7.25MB	95	1978
BilADI (8 $\mathcal{H}_2$-opt. shifts)	1/4	33	7.25MB	95	2269
GLEK	1/4	30	33.42MB	118	5330

[1] [Benner,Breiten '13]

[2] [Shank,Simoncini,Szyld '16]

Poisson problem: generalized Sylvester equation

Poisson–Chi problem

$$\Delta u + \chi u = f \qquad (x, y) \in [0, 1] \times \mathbb{R}$$
$$u(x, 0) = u(x, 1) = 0 \qquad \text{homogeneous Dirichlet b.c.}$$
$$u(x, y + 1) = u(x, y) \qquad \text{periodic b.c.}$$

- f: source term (separable function)
- $$\chi(x, y) = \begin{cases} 1 & x, y > 1/2 \\ 0 & \text{otherwise} \end{cases}$$

Poisson problem: generalized Sylvester equation

Poisson–Chi problem

$\Delta u + \chi u = f$	$(x, y) \in [0, 1] \times \mathbb{R}$
$u(x, 0) = u(x, 1) = 0$	homogeneous Dirichlet b.c.
$u(x, y + 1) = u(x, y)$	periodic b.c.

Discretization

$$AX + XB^T + DXD^T = C_1 C_2^T$$

- A: Circulant tridiagonal with elements $n^2(1, -2, 1)$
- B: Toeplitz tridiagonal with elements $n^2(1, -2, 1)$
- C_1, C_2 low rank, $D = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}$

Poisson problem: generalized Sylvester equation

Poisson–Chi: Sylvester equation

$$AX + XB^T + DXD^T = C_1 C_2^T$$

Properties

- $AD = DA + v_1 w_1^T - w_1 v_1^T - v_2 w_2^T + w_2 v_2^T$
- $BD = DB + v_1 w_1^T - w_1 v_1^T$
- $D^2 = D$
- A: singular

Let $U = [v_1, v_2, w_1, w_2]$ and $Q = [v_1, w_1]$ then

$$\mathcal{K}_d = \mathbf{EK}_d^{\square}(A, [C_1, DC_1, \dots, D^N C_1, U, N, \dots, D^{N-1} U])$$

$$\mathcal{H}_d = \mathbf{EK}_d^{\square}(B, [C_2, DC_2, \dots, D^N C_2, Q, DQ, \dots, D^{N-1} Q])$$

Poisson problem: generalized Sylvester equation

Poisson–Chi: Sylvester equation

$$AX + XB^T + DXD^T = C_1 C_2^T$$

Properties

- $AD = DA + v_1 w_1^T - w_1 v_1^T - v_2 w_2^T + w_2 v_2^T$
- $BD = DB + v_1 w_1^T - w_1 v_1^T$
- $D^2 = D$
- A: singular

Let $U = [v_1, v_2, w_1, w_2]$ and $Q = [v_1, w_1]$ then

$$\mathcal{K}_d = \mathbf{EK}_d^{\square}(A, [C_1, DC_1, U, DU])$$

$$\mathcal{H}_d = \mathbf{EK}_d^{\square}(B, [C_2, DC_2, Q, DQ])$$

Poisson problem: generalized Sylvester equation

Poisson–Chi: Sylvester equation

$$AX + XB^T + DXD^T = C_1 C_2^T$$

Properties

- $AD = DA + v_1 w_1^T - w_1 v_1^T - v_2 w_2^T + w_2 v_2^T$
- $BD = DB + v_1 w_1^T - w_1 v_1^T$
- $D^2 = D$
- A: singular

Let $U = [v_1, v_2, w_1, w_2]$ and $Q = [v_1, w_1]$ then

$$\mathcal{K}_d = \mathbf{EK}_d^{\square}(A, [C_1, DC_1, U, DU])$$

$$\mathcal{H}_d = \mathbf{EK}_d^{\square}(B, [C_2, DC_2, Q, DQ])$$

but A is singular...

Poisson problem: generalized Sylvester equation

Poisson–Chi: Sylvester equation

$$(A+I)X + XB^T + DXD^T - X = C_1 C_2^T$$

Properties

- $AD = DA + v_1 w_1^T - w_1 v_1^T - v_2 w_2^T + w_2 v_2^T$
- $BD = DB + v_1 w_1^T - w_1 v_1^T$
- $D^2 = D$
- A: singular

Let $U = [v_1, v_2, w_1, w_2]$ and $Q = [v_1, w_1]$ then

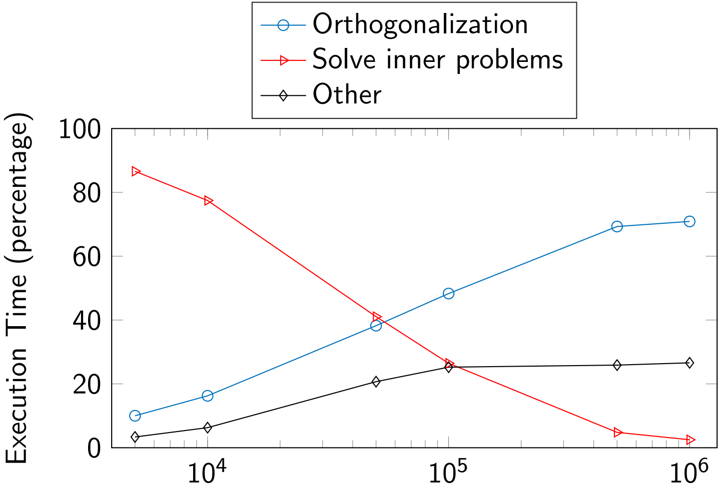
$$\mathcal{K}_d = \mathbf{EK}_d^{\square}(A+I, [C_1, DC_1, U, DU])$$

$$\mathcal{H}_d = \mathbf{EK}_d^{\square}(B, [C_2, DC_2, Q, DQ])$$

Poisson problem: generalized Sylvester equation

Poisson–Chi: Sylvester equation (shifted)

$$(A+I)X + XB^T + DXD^T - X = C_1 C_2^T$$



Conclusion

Scientific contributions:

- ✐ New low rank method for generalized Sylvester equations
- ✐ Structured exploitation for Extended Krylov method
- ✐ Characterization of the low rank numerical solutions

Future of this project:

- Preprint available soon