Greedy low-rank approaches to general linear matrix equations

Daniel Kressner and Petar Sirković

EPF Lausanne, MATHICSE, ANCHP

Disentis, 14-16 August, 2013

Outline

1. General linear matrix equations
2. Low-rank approximations
3. Greedy rank-1 updates
4. Improvements
5. Conclusions

Outline

1. General linear matrix equations
2. Low-rank approximations
3. Greedy rank-1 updates
4. Improvements
5. Conclusions

General linear matrix equations

Solve in $X \in \mathbb{R}^{n\times n}$

$$\sum_{k=1}^{K} A_k X B_k^T = CD^T \qquad \text{(GLME)}$$

- $A_1, \ldots, A_K, B_1, \ldots, B_K \in \mathbb{R}^{n\times n}, C, D \in \mathbb{R}^{n\times \ell}$
- usually $\ell \ll n$
- n^2 unknowns = entries of X
- applications in control theory (Simoncini, 2013), image science (Bouhamidi et al., 2012), Focker-Planck equation (Hartmann et al., 2010)
- recent survey paper by V. Simoncini: "The efficient numerical solution to (GLME) thus represents the next frontier for linear matrix equations . . . "

Important special case - Lyapunov equation

Solve in $X \in \mathbb{R}^{n \times n}$

$$AX + XA^T = -BB^T \qquad \text{(LYAP)}$$

- ubiquitous in control theory
- various efficient methods available (Simoncini, 2013) such as Bartels-Stewart, Krylov subspace methods, low-rank ADI

Derivation of Lyapunov equation I

Given the control system

$$x'(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t) + Du(t)$$

Controllability Gramian P is defined as

$$P = \int_0^\infty e^{At} BB^T e^{A^T t} dt.$$

It can easily be shown that P is the solution of Lyapunov equation

$$AP + PA^T + BB^T = 0.$$

Derivation of Lyapunov equation II

$$
\begin{aligned}
AP + PA^T &= \int_0^\infty (Ae^{At}BB^Te^{A^Tt} + e^{At}BB^Te^{A^Tt}A^T)dt \\
&= \int_0^\infty \frac{\partial}{\partial t}(e^{At}BB^Te^{A^Tt})dt \\
&= (e^{At}BB^Te^{A^Tt}) \mid_{t=0}^{\infty} \\
&= 0 - BB^T = -BB^T
\end{aligned}
$$

Derivation of generalized Lyapunov equation

Given the control system

$$x'(t) = Ax(t) + \sum_{k=1}^{K} N_k x(t)u(t) + Bu(t)$$
$$y(t) = Cx(t) + Du(t)$$

Controllability Gramian P is defined similarly as before, and it can be shown that P is the solution of generalized Lyapunov equation

$$AP + PA^T + \sum_{k=1}^{K} N_k P N_k^T + BB^T = 0.$$

Special case - Generalized Lyapunov equation

Solve in $X \in \mathbb{R}^{n\times n}$

$$AX + XA^T + \sum_{k=1}^{K} N_k X N_k^T = -BB^T \qquad \text{(GLYAP)}$$

- applications in bilinearization of nonlinear problems, Focker-Planck equation, heat equation with Robin boundary conditions

Solving GLME

General linear matrix equation

$$\sum_{k=1}^{K} A_k X B_k^T = CD^T$$

- naive approach = transform GLME into linear system of size $n^2 \times n^2$:

$$\sum_{k=1}^{K} (B_k \otimes A_k) \operatorname{vec}(X) =: \mathcal{A} \operatorname{vec}(X) = \operatorname{vec}(CD^T) \quad \text{(vGLME)}$$

  $\Rightarrow$ severe limitation on n with classical methods

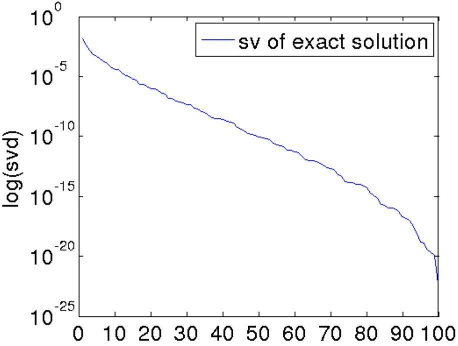
- most techniques for solving Lyapunov equations (e.g., Krylov subspace methods) **do not** extend to (GLME) directly

Outline

1. General linear matrix equations
2. Low-rank approximations
3. Greedy rank-1 updates
4. Improvements
5. Conclusions

Singular value decay and low-rank approximations I

Solution of (GLME) often exhibits very strong singular value decay.



Example of GLYAP - Singular value decay

Singular value decay and low-rank approximations II

Natural assumption:

Exact solution can be well approximated by low-rank matrix.

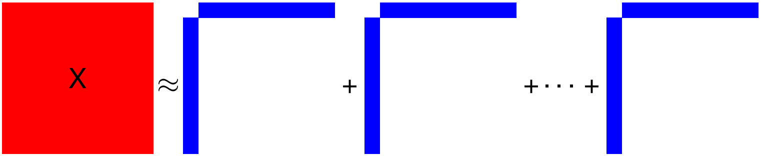
Many existing algorithms exploit this idea for (LYAP). Existing low-rank approaches to (GLYAP):

- fixed-point method with ADI-preconditioning (Damm, 2008)
- preconditioned Krylov subspace methods (Benner et al., 2010),

Both mainly limited to the case where Lyapunov part $AX + XA^T$ dominates (GLYAP).

Low-rank approximations

X has low-rank structure $\Rightarrow$ can be written as a sum of rank-1 matrices.



Outline

1. General linear matrix equations
2. Low-rank approximations
3. Greedy rank-1 updates
4. Improvements
5. Conclusions

Rank-1 updates

Idea how to exploit this: **Greedy updates** (inspired by Chinesta et al., 2010)

- Assume current approximation $X_{\mathrm{old}} =$ sum of i rank-1 matrices
- Get new approximation $X_{\mathrm{new}} \leftarrow X_{\mathrm{old}} + uv^T$ by choosing uv^T optimally
- Optimality depends on the choice of target functional $\mathcal{J}$. Two possibilities:
  - energy norm $\mathcal{J}(X_a, u, v) = || \operatorname{vec}(X_a + uv^T) - \operatorname{vec}(X)||_{\mathcal{A}}$
  - residual $\mathcal{J}(X_a, u, v) = ||\mathcal{A} \operatorname{vec}(X_a + uv^T - X)||_2$
- For either criterion, ALS is used to determine minimum $\Rightarrow$ solution of **one** $n \times n$ **linear system** in every iteration

ALS minimization I

Goal: Minimize $|| \operatorname{vec}(X_a + uv^T) - \operatorname{vec}(X)||_{\mathcal{A}}$

This is equivalent to

$$\min_{u,v} \operatorname{tr}(vu^T(\sum_{k=1}^{K} A_k uv^T B_k^T)) - 2\operatorname{tr}(vu^T Q_i)$$

We alternate between optimization over u and v, other variable stays fixed. For a fixed v, optimal $\hat{u}$ is a local minimum $\Rightarrow$ satisfies

$$\operatorname{tr}(v\hat{u}^T(\sum_{k=1}^{K} A_k \hat{u} v^T B_k^T)) - 2\operatorname{tr}(v\hat{u}^T Q_i) \approx$$

$$\operatorname{tr}(v(\hat{u}+\Delta)^T(\sum_{k=1}^{K} A_k (\hat{u}+\Delta) v^T B_k^T)) - 2\operatorname{tr}(v(\hat{u}+\Delta)^T Q_i),$$

for all small Δ.

ALS minimization II

After disregarding second-order terms we get following equation

$$\frac{1}{2}(\sum_{k=1}^{K}(A_k\hat{u}v^TB_k^Tv + A_k^T\hat{u}v^TB_kv)) - Q_iv = 0$$

Since v^TB_kv is a scalar we get

$$\frac{1}{2}(\sum_{k=1}^{K}(v^TB_k^TvA_k + v^TB_kvA_k^T))\hat{u} = Q_iv.$$

To compute $\hat{u}$, $n \times n$ linear system has to be solved.

For fixed u, we get similar equation for $\hat{v}$.

Algorithm - Greedy rank-1 updates

Require: $A_1, \ldots, A_K, B_1, \ldots, B_K, C, D$
Ensure: low rank approximation X_a

$X_a = 0$
$Q = CD^T$
for $i = 1, \ldots,$ *#maxrank* **do**
  u_i, v_i random $n \times 1$ matrices
  for until convergence **do**
    $u_i \leftarrow \arg\min_{u_i} \mathcal{J}(X_a, u_i, v_i)$
    $v_i \leftarrow \arg\min_{v_i} \mathcal{J}(X_a, u_i, v_i)$
  end for
  $X_a \leftarrow X_a + u_i v_i^T$
  $Q \leftarrow Q - \sum_{k=1}^{K} A_k u_i v_i^T B_k^T$
end for
X_a wanted approximation

Lyapunov equation - convergence

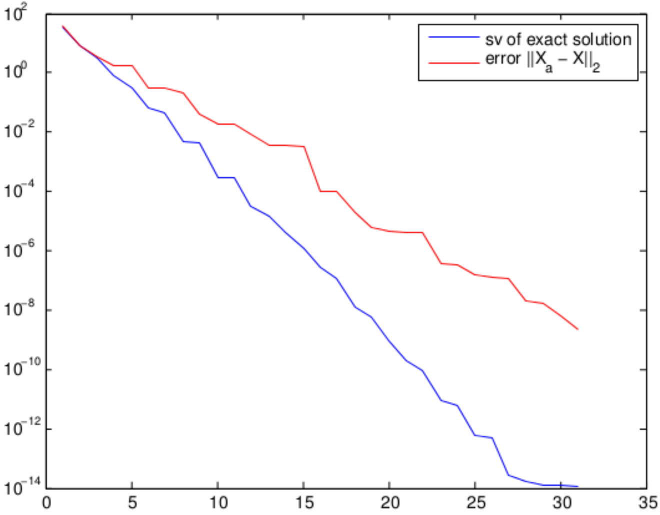


Figure: Convergence of successive rank-1 updates for symmetric positive definite (LYAP) - comparison of singular values and error.

Analysis

Theorem

For symmetric positive-definite (LYAP) *with symmetric semidefinite right-hand side, minimum of ALS is always attained in a point where* $U = V$.

Corollary

For symmetric positive-definite (LYAP) *with symmetric semidefinite right-hand side convergence is monotonic in Löwner ordering on positive semidefinite matrices.*

Generalized Lyapunov equation I

Heat equation with the control in the boundary condition:

$$\begin{aligned} x_t &= \Delta x && \\ n \cdot \nabla x &= 0.5 \cdot u(x-1) && \text{on } \Gamma_1, \Gamma_2 \\ x &= 0 && \text{on } \Gamma_3, \Gamma_4 \end{aligned}$$

Each Robin boundary condition introduces a coupling between $x(t)$ and $u(t)$ $\Rightarrow$ two matrices N_i $\Rightarrow$ resulting equation:

$$AX + XA^T + N_1 X N_1^T + N_2 X N_2^T = -BB^T$$

System matrix

$$\mathcal{A} = (A \otimes I) + (I \otimes A) + (N1 \otimes N1) + (N2 \otimes N2)$$

Generalized Lyapunov equation II

Convergence of algorithm depends on the fact if the system matrix $\mathcal{A}$ in (vGLME) is positive definite.

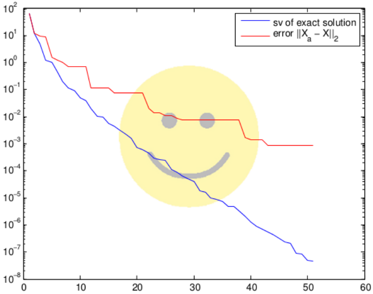


Figure: GLYAP heat equation with Robin b.c. - positive definite $\mathcal{A}$, minimization of energy norm

Figure: GLYAP heat equation with Robin b.c. - indefinite $\mathcal{A}$, minimization of residual

Outline

1. General linear matrix equations
2. Low-rank approximations
3. Greedy rank-1 updates
4. Improvements
5. Conclusions

Galerkin projections

Greedy rank-1 updates

$$X_a = u_1 v_1^T + u_2 v_2^T + \cdots + u_m v_m^T$$

- Idea: Collect direction of updates in subspaces

$$\mathcal{U} = \text{span}(\{u_1, \ldots, u_m\}), \mathcal{V} = \text{span}(\{v_1, \ldots, v_m\})$$

- Obtain (hopefully) improved approximation by Galerkin projection on $\mathcal{V} \otimes \mathcal{U}$

$$\Updownarrow$$

  approximate solution $X_a = UYV^T, Y \in \mathbb{R}^{m \times m}$
- Cost = solving linear system of size $m^2 \times m^2$

Galerking projections example

Approach actually fixes convergence problems for indefinite (GLYAP).

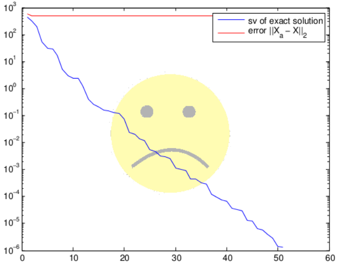


Figure: Greedy rank-1 updates

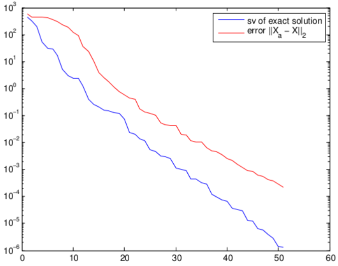


Figure: Greedy rank-1 updates + Galerkin

Bilinearization of RC circuit I

$$v_t = f(v) + bu(t)$$

$$f(v) = [f_k(v)] = \begin{pmatrix} -g(v_1) - g(v_1 - v_2) \\ g(v_1 - v_2) - g(v_2 - v3) \\ \vdots \\ g(v_{N_0 - 1} - v_{N_0}) \end{pmatrix}$$

$$g(v) = exp(40v) + v - 1$$

Carleman bilinearization $\Rightarrow$ matrix equation of size $(N_0 + N_0^2) \times (N_0 + N_0^2)$.

$$AX + XA^T + NXN^T = -BB^T$$

Bilinearization of RC circuit II

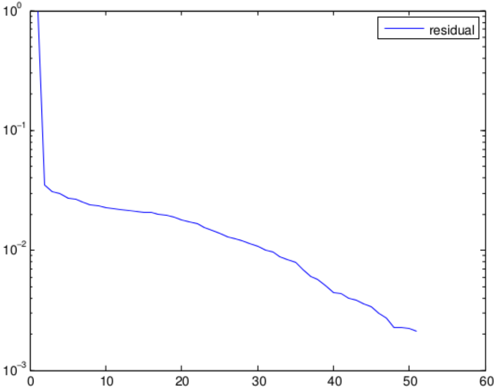


Figure: Bilinearization of RC circuit - convergence of the residual with Galerkin approach

- convergence is slow
- possibly needs preconditioning

Preconditioned residual I

Idea: Inject few dominant vectors of preconditioned residual to the subspaces U and V.

- preconditioner $\mathcal{P}^{-1}$ = one iteration of iterative Lyapunov solver (using matrix sign function)
- $P_U \Sigma P_V = \mathcal{P}^{-1}(Q_i)$ and truncate
- $$\mathcal{U} \leftarrow \mathrm{span}(\mathcal{U} \cup P_U)$$
  $$\mathcal{V} \leftarrow \mathrm{span}(\mathcal{V} \cup P_V)$$
- Galerking projection on $\mathcal{U}$ and $\mathcal{V}$
- truncation of subspaces if needed

Preconditioned residual II

Using preconditioned residual improves the convergence!

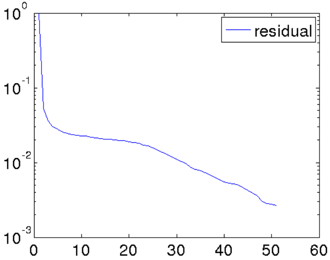


Figure: Bilinear RC circuit - no preconditioning

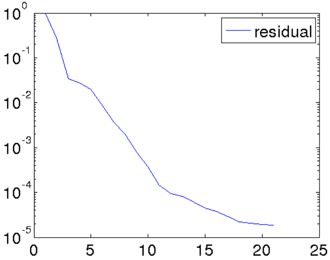


Figure: Bilinear RC circuit - preconditioned residual

Outline

1. General linear matrix equations
2. Low-rank approximations
3. Greedy rank-1 updates
4. Improvements
5. Conclusions

Conclusions

- general linear matrix equations
- low-rank approximation - Greedy rank-1 updates
- Galerkin projection - fixes indefinite case
- using preconditioned residual - accelerates the convergence

Thank you for your attention!

Selected references

- Bai, Z. and Skoogh, D., A projection method for model reduction of bilinear dynamical systems, 2006
- Benner, P. and Breiten, T., Low rank methods for a class of generalized Lyapunov equations and related issues, 2012
- Bouhamidi, A. and Jbilou, K. and Reichel, L. and Sadok, H., A generalized global Arnoldi method for ill-posed matrix equations, 2012
- Chinesta, F. and Ammar, A. and Cueto, E., Recent advances and new challenges in the use of the proper generalized decomposition for solving multidimensional models, 2010
- Damm, T., Direct methods and ADI-preconditioned Krylov subspace methods for generalized Lyapunov equations, 2008
- Simoncini, V., Computational methods for linear matrix equations, 2013