ADI-preconditioned Krylov subspace methods for Lyapunov-type equations

Tobias Damm

AG Technomathematik
Fachbereich Mathematik
TU Kaiserslautern

Overview

- Lyapunov-type equations: Applications
  - Stability
  - Gramians and model reduction
- Lyapunov-type equations: Theory
- Lyapunov-type equations: Numerical solution
  - Krylov subspace iterations
  - ADI-preconditioning

Lyapunov-type operators and stability criteria

Type of system		Stability criterion $\exists P > 0:$
$\dot{x} = Ax$	$\Longleftrightarrow$	$A^*P + PA < 0$
$dx = Ax\,dt + A_0 x\,dw$	$\Longleftrightarrow$	$A^*P + PA + A_0^*PA_0 < 0$
$x_{k+1} = Ax_k$	$\Longleftrightarrow$	$A^*PA - P < 0$
$x_{k+1} = Ax_k + A_0 x\, w_k$	$\Longleftrightarrow$	$A^*PA - P + A_0^*PA_0 < 0$
$\dot{x} = Ax(t) + A_0x(t-h)$	$\overset{\forall h}{\Longleftarrow}$	$A^*P + PA + P + A_0^*PA_0 < 0$

Lyapunov-type operators and stability criteria

Type of system		Stability criterion $\exists P > 0$:
$\dot{x} = Ax$	$\iff$	$A^*P + PA < 0$
$dx = Ax\,dt + A_0 x\,dw$	$\iff$	$A^*P + PA + A_0^* P A_0 < 0$
$x_{k+1} = Ax_k$	$\iff$	$A^*PA - P < 0$
$x_{k+1} = Ax_k + A_0 x\,w_k$	$\iff$	$A^*PA - P + A_0^* P A_0 < 0$
$\dot{x} = Ax(t) + A_0 x(t-h)$	$\overset{\forall h}{\Leftarrow}$	$A^*P + PA + P + A_0^* P A_0 < 0$

The generalized Lyapunov operator

$$T : P \mapsto A^*P + PA + \sum_{j=1}^{N} A_j^* P A_j$$

generates a positive semigroup on $\mathcal{H}^n$.

Reachability and observability Gramians of bilinear systems

Linear system: $\dot{x} = Ax + Bu \,, \quad y = Cx \,, \quad \sigma(A) \subset \mathbb{C}_-$

Reachability and observability Gramians P, Q:

$$AP + PA^* = -BB^* \,, \qquad A^*Q + QA = -C^*C$$

Energy functionals

$$\begin{aligned} E_c(x_0) &= \min_{\substack{u \in L^2(-\infty,0) \\ x(-\infty,u)=0,\, x(0,u)=x_0}} \|u\|^2_{L^2} &&= x_0^T P^{-1} x_0 \\ E_o(x_0) &= \|y(\cdot, x_0)\|^2_{L^2} &&= x_0^T Q x_0 \end{aligned}$$

$$\left.\begin{matrix} x_0^T P x_0 \\ x_0^T Q x_0 \end{matrix}\right\} \text{small} \Rightarrow \text{state } x_0 \text{ hard} \left\{\begin{matrix} \text{to reach} \\ \text{to observe} \end{matrix}\right.$$

Reachability and observability Gramians of bilinear systems

Stochastic system: $dx = (Ax + Bu)\,dt + A_0 x\,dw\ , \quad y = Cx\ ,$

Reachability and observability Gramians

$$AP + PA^* + A_0 P A_0^* = -BB^*\ , \qquad A^* Q + QA + A_0^* Q A_0 = -C^* C$$

Energy functionals

$$\begin{aligned} E_c(x_0) &= \min_{\substack{u \in L_w^2(-\infty,0) \\ x(-\infty,u)=0,\, x(0,u)=x_0}} \|u\|_{L_w^2}^2 &&= x_0^T P^{-1} x_0 \\ E_o(x_0) &= \|y(\cdot, x_0)\|_{L_w^2}^2 &&= x_0^T Q x_0 \end{aligned}$$

$$\left.\begin{array}{l} x_0^T P x_0 \\ x_0^T Q x_0 \end{array}\right\} \text{small} \Rightarrow \text{state } x_0 \text{ hard} \left\{\begin{array}{l} \text{to reach} \\ \text{to observe} \end{array}\right.$$

Reachability and observability Gramians of bilinear systems

Bilinear **system:** $\dot{x} = Ax + \sum_{j=1}^{m} A_j x u_j + Bu \ , \quad y = Cx \ ,$

Reachability and observability Gramians [Al-Baiyat/Bettayeb]

$$AP + PA^* + \sum A_j P A_j^* = -BB^* \ , \quad A^* Q + QA + \sum A_j^* Q A_j = -C^* C$$

Energy functionals [Scherpen, Gray/Mesko]

$$\begin{aligned} E_c(x_0) &= \min_{\substack{u \in L^2(-\infty,0) \\ x(-\infty,u)=0, x(0,u)=x_0}} \|u\|_{L^2}^2 && \geq x_0^T P^{-1} x_0 \\ E_o(x_0) &= \max_{u \in L^2(0,\infty), \|u\|_{L^2} \leq 1} \|y(\cdot, x_0, u)\|_{L^2}^2 && \leq x_0^T Q x_0 \end{aligned}$$

$$\left.\begin{matrix} x_0^T P x_0 \\ x_0^T Q x_0 \end{matrix}\right\} \text{small} \Rightarrow \text{state } x_0 \text{ hard} \left\{\begin{matrix} \text{to reach} \\ \text{to observe} \end{matrix}\right.$$

Balanced realization

Factorizations $P = LL^T$, $L^T QL = U\Sigma^2 U^T$
give transformation $T = LU\Sigma^{-1/2}$ and equivalent system

$$\tilde{A} = T^{-1}AT\,,\ \tilde{A}_j = T^{-1}A_j T\,, \quad \tilde{B} = T^{-1}B\,, \quad \tilde{C} = CT\,.$$

This is balanced:

$$\tilde{P} = \tilde{Q} = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}, \quad \text{Hankel singular values}$$

If $\sigma_{r+1}, \ldots, \sigma_n$ is small: State neglectable!
Projection on $\mathbb{R}^r$, $(r \overset{!}{\ll} n)$

Model reduction by balanced truncation

Partition: $T = [T_1, T_2],\ T^{-1} = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix}.$

Truncation: $\begin{array}{rclcrcl} \tilde{A}^{(r)} & = & S_1 A T_1 & & \tilde{A}_j^{(r)} & = & S_1 A_j T_1 \\ \tilde{B}^{(r)} & = & S_1 B & & \tilde{C}^{(r)} & = & C T_1 \end{array}$

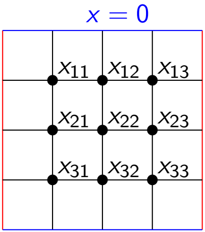
Reduced model:

$$\dot{\tilde{x}}_r = A^{(r)} \tilde{x}_r + \sum_{j=1}^{m} A_j^{(r)} \tilde{x}_r u_j + B^{(r)} u \qquad \tilde{y} = C^{(r)} \tilde{x}_r \ .$$

Literature:

- Extended ("H_2 model reduction") e.g. in Zhang/Lam 2002
- Only very small examples (e.g. $n = 4,\ r = 2$)
- No theory for generalized Lyapunov equation?

Example: A heat conduction problem

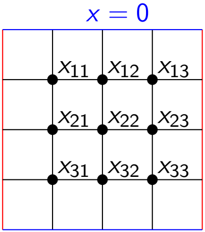


$$
\begin{aligned}
\dot{x} &= \Delta x \\
n \cdot \nabla x &= u(x-1)
\end{aligned}
$$

$$\Delta x_{ij} \approx -\frac{1}{h^2}\left(4x_{ij} - x_{i+1,j} - x_{i,j+1} - x_{i-1,j} - x_{i,j-1}\right)$$

$$x_{10} \approx x_{11} - hu(x_{11} - 1),\; x_{20} \approx x_{21} - hu(x_{21} - 1), \ldots$$

Example: A heat conduction problem



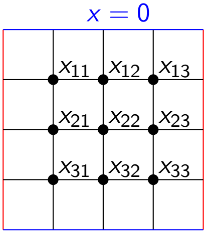
$$\begin{aligned} \dot{x} &= \Delta x \\ n \cdot \nabla x &= u(x-1) \end{aligned}$$

$$\Delta x_{ij} \approx -\frac{1}{h^2}\left(4x_{ij} - x_{i+1,j} - x_{i,j+1} - x_{i-1,j} - x_{i,j-1}\right)$$

$$x_{10} \approx x_{11} - hu(x_{11}-1),\ x_{20} \approx x_{21} - hu(x_{21}-1), \ldots$$

$$\dot{x} = \frac{1}{h^2}\left[\begin{array}{ccc|ccc|c} -3 & 1 & & 1 & & & \\ 1 & -3 & 1 & & 1 & & \\ & 1 & -3 & & & 1 & \\ \hline 1 & & & -4 & 1 & & \\ & 1 & & 1 & -4 & 1 & \star \\ & & 1 & & 1 & -4 & \\ \hline & & & & \star & & \star \end{array}\right] \underbrace{\left[\begin{array}{c} x_{11} \\ x_{21} \\ x_{31} \\ \hline x_{12} \\ x_{22} \\ x_{32} \\ \hline \star \end{array}\right]}_{=:x} + \frac{1}{h}\left(\left[\begin{array}{c} I \\ \hline 0 \\ \hline \star \end{array}\right] x - \left[\begin{array}{c} 1 \\ 1 \\ 1 \\ \hline 0 \\ 0 \\ 0 \\ \hline \star \end{array}\right]\right) u$$

Example: A heat conduction problem



$$\begin{aligned} \dot{x} &= \Delta x \\ n \cdot \nabla x &= u(x-1) \end{aligned}$$

$$\Delta x_{ij} \approx -\frac{1}{h^2}\left(4x_{ij} - x_{i+1,j} - x_{i,j+1} - x_{i-1,j} - x_{i,j-1}\right)$$

$$x_{10} \approx x_{11} - hu(x_{11}-1),\ x_{20} \approx x_{21} - hu(x_{21}-1), \ldots$$

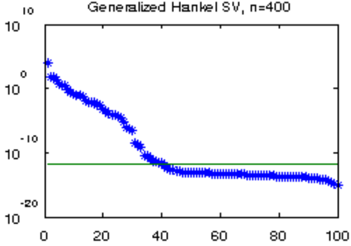
$$\dot{x} = \frac{1}{h^2}\left[\begin{array}{ccc|ccc|c} -3 & 1 & & 1 & & & \\ 1 & -3 & 1 & & 1 & & \\ & 1 & -3 & & & 1 & \\ \hline 1 & & & -4 & 1 & & \\ & 1 & & 1 & -4 & 1 & \star \\ & & 1 & & 1 & -4 & \\ \hline & & & & \star & & \star \end{array}\right] \underbrace{\left[\begin{array}{c} x_{11} \\ x_{21} \\ x_{31} \\ \hline x_{12} \\ x_{22} \\ x_{32} \\ \hline \star \end{array}\right]}_{=:x} + \frac{1}{h}\left(\left[\begin{array}{c} I \\ \hline 0 \\ \hline \star \end{array}\right] x - \left[\begin{array}{c} 1 \\ 1 \\ 1 \\ \hline 0 \\ 0 \\ 0 \\ \hline \star \end{array}\right]\right) u$$

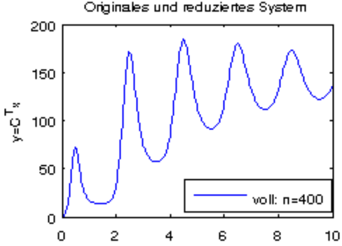
$$\boxed{\dot{x} = Ax + (A_0 x)u + Bu}$$

Model reduction for bilinear systems: A visual check

$$\boxed{\dot{x} = Ax + \sum_{j=1}^{2} A_j x u_j + Bu, \quad y = Cx, \quad u_j(t) = \frac{A \cdot \cos(j\pi t)}{|t|+1}}$$

$u : \mathbb{R}_+ \to \mathbb{R}^2$ is an L^2-test signal, $y : \mathbb{R}_+ \to \mathbb{R}$ the average temperature

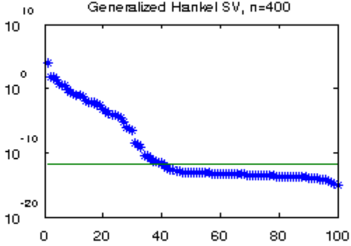


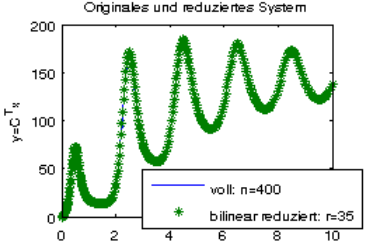


Model reduction for bilinear systems: A visual check

$$\boxed{\dot{x} = Ax + \sum_{j=1}^{2} A_j x u_j + Bu, \quad y = Cx, \quad u_j(t) = \frac{A \cdot \cos(j\pi t)}{|t|+1}}$$

$u : \mathbb{R}_+ \to \mathbb{R}^2$ is an L^2-test signal, $y : \mathbb{R}_+ \to \mathbb{R}$ the average temperature

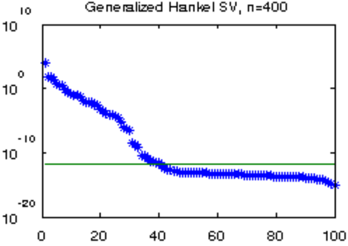


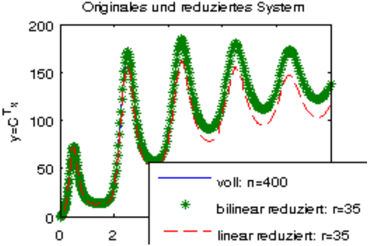


Model reduction for bilinear systems: A visual check

$$\boxed{\dot{x} = Ax + \sum_{j=1}^{2} A_j x u_j + Bu, \quad y = Cx, \quad u_j(t) = \frac{A \cdot \cos(j\pi t)}{|t|+1}}$$

$u : \mathbb{R}_+ \to \mathbb{R}^2$ is an L^2-test signal, $y : \mathbb{R}_+ \to \mathbb{R}$ the average temperature



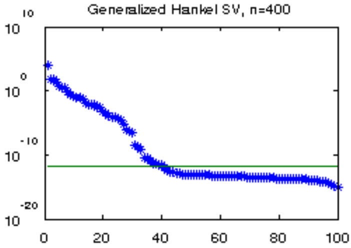


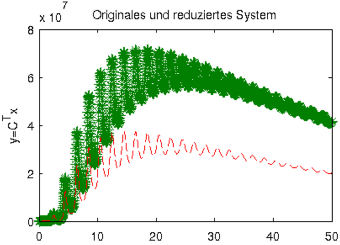
Comparison: P, Q bilinear Gramian
P, Q linear Gramian, $(AP + PA^T = -BB^T, \ldots)$

Model reduction for bilinear systems: A visual check

$$\boxed{\dot{x} = Ax + \sum_{j=1}^{2} A_j x u_j + Bu, \quad y = Cx, \quad u_j(t) = \frac{A \cdot \cos(j\pi t)}{|t|+1}}$$

$u : \mathbb{R}_+ \to \mathbb{R}^2$ is an L^2-test signal, $y : \mathbb{R}_+ \to \mathbb{R}$ the average temperature





Comparison: P, Q bilinear Gramian

P, Q linear Gramian, $(AP + PA^T = -BB^T, \ldots)$

Excursion: Is the idea applicable to delay equations?

$$\boxed{\dot{x}(t) = Ax(t) + A_0 x(t-h) + Bu(t), \quad y(t) = Cx(t)}$$

Assume: $\exists P > 0$, $Q > 0$

$$\begin{aligned} A^* P + PA + P + A_0^* P A_0 &= -BB^T\ , \\ AQ + QA^* + Q + A_0^* Q A_0 &= -C^T C\ . \end{aligned}$$

Some experiments

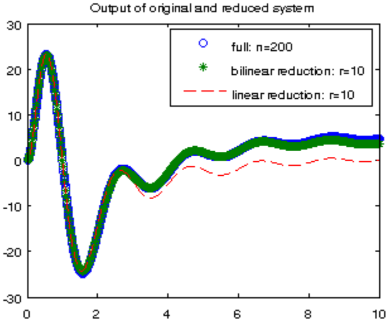
Excursion: Is the idea applicable to delay equations?

$$\boxed{\dot{x}(t) = Ax(t) + A_0 x(t-h) + Bu(t), \quad y(t) = Cx(t)}$$

Assume: $\exists P > 0,\ Q > 0$

$$\begin{aligned} A^*P + PA + P + A_0^* P A_0 &= -BB^T\ , \\ AQ + QA^* + Q + A_0^* Q A_0 &= -C^T C\ . \end{aligned}$$

Some experiments



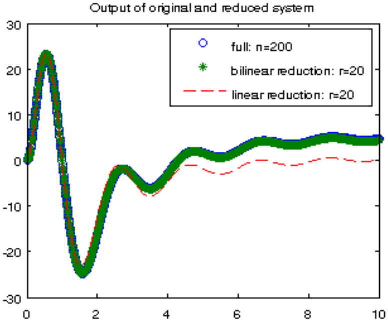
Excursion: Is the idea applicable to delay equations?

$$\boxed{\dot{x}(t) = Ax(t) + A_0 x(t-h) + Bu(t), \quad y(t) = Cx(t)}$$

Assume: $\exists P > 0, Q > 0$

$$\begin{aligned} A^*P + PA + P + A_0^* P A_0 &= -BB^T \,, \\ AQ + QA^* + Q + A_0^* Q A_0 &= -C^T C \,. \end{aligned}$$

Some experiments



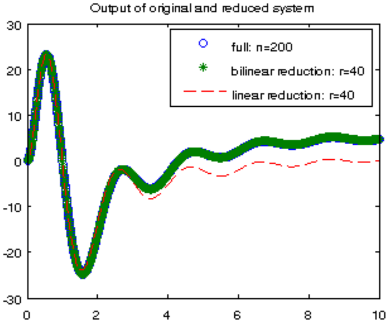
Excursion: Is the idea applicable to delay equations?

$$\boxed{\dot{x}(t) = Ax(t) + A_0x(t-h) + Bu(t), \quad y(t) = Cx(t)}$$

Assume: $\exists P > 0, Q > 0$

$$\begin{aligned} A^*P + PA + P + A_0^*PA_0 &= -BB^T \,, \\ AQ + QA^* + Q + A_0^*QA_0 &= -C^TC \,. \end{aligned}$$

Some experiments



Overview

- Lyapunov-type equations: Applications
  - Stability
  - Gramians and model reduction
- Lyapunov-type equations: Theory
- Lyapunov-type equations: Numerical solution
  - Krylov subspace iterations
  - ADI-preconditioning

Resolvent positive operators on $\mathcal{H}^n$

Definition: $T : \mathcal{H}^n \to \mathcal{H}^n$, linear,

- *positive* $(T \geq 0) \quad \iff \quad T(\mathcal{H}^n_+) \subset \mathcal{H}^n_+$
- *resolvent positive* $\quad \iff \quad \forall \alpha \gg 0\text{: } (\alpha I - T)^{-1} \geq 0$

$$\iff \quad \forall t \geq 0\text{: } e^{Tt} \geq 0$$

Resolvent positive operators on $\mathcal{H}^n$

Definition: $T : \mathcal{H}^n \to \mathcal{H}^n$, linear,

- *positive* ($T \geq 0$) $\iff$ $T(\mathcal{H}^n_+) \subset \mathcal{H}^n_+$
- *resolvent positive* $\iff$ $\forall \alpha \gg 0$: $(\alpha I - T)^{-1} \geq 0$

  $\iff$ $\forall t \geq 0$: $e^{Tt} \geq 0$

Generalized Lyapunov operators are resolvent positive.

Resolvent positive operators on $\mathcal{H}^n$

Definition: $T : \mathcal{H}^n \to \mathcal{H}^n$, linear,

- *positive* $(T \geq 0) \quad \iff \quad T(\mathcal{H}^n_+) \subset \mathcal{H}^n_+$
- *resolvent positive* $\quad \iff \quad \forall \alpha \gg 0$: $(\alpha I - T)^{-1} \geq 0$

  $\iff \quad \forall t \geq 0$: $e^{Tt} \geq 0$

Generalized Lyapunov operators are resolvent positive.

Theorem: [H. Schneider]
For T resolvent positive, the following are equivalent:

- $\exists X > 0 : T(X) < 0$
- $\sigma(T) \subset \mathbb{C}_-$
- $-T^{-1} > 0$
- If $T = \mathcal{L} + \Pi$ with $\mathcal{L}$ res.pos. and $\Pi \geq 0$, then $\sigma(\mathcal{L}) \subset \mathbb{C}_-$ and $\rho(\mathcal{L}^{-1}\Pi) < 1$.

Overview

- Lyapunov-type equations: Applications
  - Stability
  - Gramians and model reduction
- Lyapunov-type equations: Theory
- Lyapunov-type equations: Numerical solution
  - Krylov subspace iterations
  - ADI-preconditioning

Generalized Lyapunov equation

$$\mathcal{L}_A(X) + \Pi(X) := AX + XA^T + \sum A_j X A_j^T = -Y$$

Lyapunov operator: $\mathcal{L}_A(X) = AX + XA^T$

Positive operator: $\Pi(X) = \sum A_j X A_j^T$

Convergent iteration: $X_{k+1} = -\mathcal{L}_A^{-1}\Pi(X_k) - \mathcal{L}_A^{-1}(Y)$

$\Rightarrow \mathcal{L}_A^{-1}$ can be used as preconditioner.

Krylov subspace methods

Why Krylov subspace methods?

$$\boxed{X_{k+1} = -\mathcal{L}_A^{-1}\Pi(X_k) - \mathcal{L}_A^{-1}(Y)}$$

Instead of the iterates $X_0, X_1, \ldots, X_k$ use an „optimal“ linear combination $\tilde{X}_k = \sum_{j=0}^{k} a_j X_j$.

- If X_k converges, then so does $\tilde{X}_k$.
- A good preconditioner is essential.
- Inversion of $\mathcal{L}_A$ can be to expensive.
- In the following: Approximation of $\mathcal{L}_A^{-1}$ via ADI.
- ADI: alternate direction implicit

A slightly more general case

Bilinear system (with $M \geq 0$)

$$M\dot{x} = Sx + \sum_{j=1}^{m} A_j x u_j + Bu\ , \quad y = Cx\ .$$

Leads to $MXS^T + SXM + \sum A_j X A_j^T = -Y$

Similar as before, if M invertible and $\sigma(M^{-1}S) \subset \mathbb{C}_-$.

Important: For arbitrary $p \in \mathbb{R}$:

$$\begin{aligned} & MXS^T + SXM \\ & = \frac{1}{2p}\Big((S+pM)X(S+pM)^T - (S-pM)X(S-pM)^T\Big) \end{aligned}$$

Idea of ADI-iteration

e.g. Wachspress, Smith, Penzl, Li:

$$\begin{aligned} MXS^T + SXM &= -Y \\ \iff \quad (S - pM)X(S - pM)^T &= (S + pM)X(S + pM)^T + 2pY \end{aligned}$$

yields the fixed point equation

$$X = \frac{S + pM}{S - pM} X \frac{S^T + pM}{S^T - pM} + 2p \frac{1}{(S - pM)} Y \frac{1}{(S - pM)^T}$$

Idea of ADI-iteration

e.g. Wachspress, Smith, Penzl, Li:

$$\begin{aligned} MXS^T + SXM &= -Y \\ \iff \quad (S - pM)X(S - pM)^T &= (S + pM)X(S + pM)^T + 2pY \end{aligned}$$

yields the fixed point equation

$$X = \prod_j \frac{S + p_j M}{S - p_j M} X \prod_j \frac{S^T + p_j M}{S^T - p_j M} + \prod_j \frac{\sqrt{2p_j}}{S - p_j M} Y \prod_j \frac{\sqrt{2p_j}}{S^T - p_j M}$$

Idea of ADI-iteration

e.g. Wachspress, Smith, Penzl, Li:

$$\begin{array}{rcl} MXS^T + SXM & = & -Y \\ \iff \quad (S - pM)X(S - pM)^T & = & (S + pM)X(S + pM)^T + 2pY \end{array}$$

yields the fixed point equation

$$X = \prod_j \frac{S + p_j M}{S - p_j M} X \prod_j \frac{S^T + p_j M}{S^T - p_j M} + \prod_j \frac{\sqrt{2p_j}}{S - p_j M} Y \prod_j \frac{\sqrt{2p_j}}{S^T - p_j M}$$

Choice of parameters: For $\varepsilon < 1$, find $p_1, \ldots, p_\ell$:

$$\rho\Big(\prod_j \frac{S + p_j M}{S - p_j M}\Big) \le \max_{\lambda \in D} \prod_j \left| \frac{\lambda - p_j}{\lambda + p_j} \right| \overset{!}{\le} \varepsilon, \qquad \sigma(M^{-1}S) \subset D$$

Classical problem, solution known e.g. for real spectra.

ADI-preconditioner for generalized Lyapunov

Also: $SXM + MXS^T + \Pi(X) = Y \iff$

$$X = (S - pM)^{-1}\Big((S + pM)X(S + pM)^T + 2p(\Pi(X) - Y)\Big)(S - pM)^{-T}$$

Preconditioned fixed point equation

Iteration: Choose $p_1, \ldots, p_\ell$ (e.g. according to Wachspress).

$$\begin{aligned}
X_1 &= (S - p_1 M) \setminus [(S + p_1 M)X_0(\ldots)^T + 2p_1(\Pi(X_0) - Y)]/(S - p_1 M)^T \\
X_2 &= (S - p_2 M) \setminus [(S + p_1 M)X_1(\ldots)^T + 2p_2(\Pi(X_0) - Y)]/(S - p_2 M)^T \\
&\vdots \\
X_\ell &= (S - p_\ell M) \setminus [(S + p_\ell M)X_{\ell-1}(\ldots)^T + 2p_\ell(\Pi(X_0) - Y)]/(S - p_\ell M)^T
\end{aligned}$$

$X_0 \mapsto X_\ell$: Another preconditioner.
Cheaper than $\mathcal{L}_A^{-1}$ e.g. if M, S sparse. Usually $\ell \approx 4$ suffices.

ADI-preconditioner for generalized Lyapunov

Also: $SXM + MXS^T + \Pi(X) = Y \iff$

$$X = (S - pM)^{-1}\Big((S + pM)X(S + pM)^T + 2p(\Pi(X) - Y)\Big)(S - pM)^{-T}$$

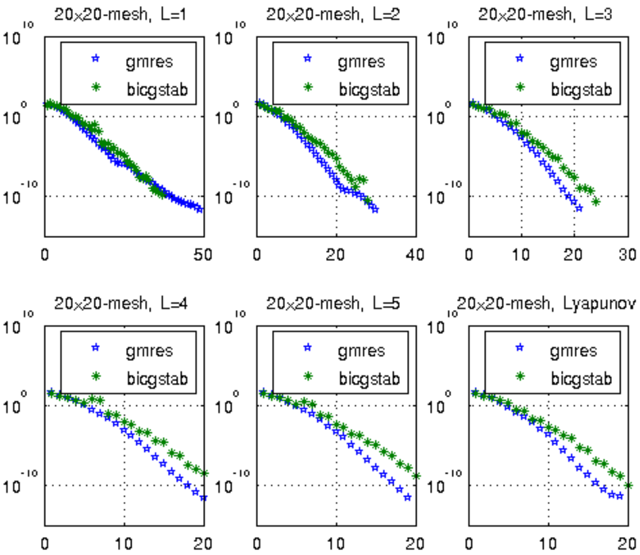
Preconditioned fixed point equation

Iteration: Choose $p_1, \ldots, p_\ell$ (e.g. according to Wachspress).

$$\begin{aligned}
X_1 &= (S - p_1 M) \setminus [(S + p_1 M)X_0(\ldots)^T + 2p_1(\Pi(X_0) - Y)]/(S - p_1 M)^T \\
X_2 &= (S - p_2 M) \setminus [(S + p_1 M)X_1(\ldots)^T + 2p_2(\Pi(X_1) - Y)]/(S - p_2 M)^T \\
&\vdots \\
X_\ell &= (S - p_\ell M) \setminus [(S + p_\ell M)X_{\ell-1}(\ldots)^T + 2p_\ell(\Pi(X_{\ell-1}) - Y)]/(S - p_\ell M)^T
\end{aligned}$$

$X_0 \mapsto X_\ell$: Another preconditioner.
Cheaper than $\mathcal{L}_A^{-1}$ e.g. if M, S sparse. Usually $\ell \approx 4$ suffices.

Convergence histories for Poisson example



Summary

- Relation between Gramians of bilinear systems and Lyapunov equations
- Model reduction can be based on these Gramians.
- Suggestion for their computation:
  ADI-preconditioned Krylov subspace method.
- Unclear: Low-rank approximation of the Gramians