

ADI shift parameter computation for large scale algebraic Riccati and Lyapunov equations arising in the LQR problem for parabolic PDEs

Jens Saak

joint work with

Peter Benner (MiIT) and Hermann Mena (EPN Quito Ecuador)

Professur Mathematik in Industrie und Technik (MiIT)
Fakultät für Mathematik
Technische Universität Chemnitz

July 25, 2006



Outline

- 1 Problem background in control of parabolic PDEs
 - Abstract Cauchy problems in Hilbert spaces
 - Semidiscretization and finite dimensional systems
- 2 LRCF Newton Method for the ARE
 - Newton's method for solving the ARE
 - Cholesky factor ADI for Lyapunov equations
- 3 ADI shift parameter computation
 - The ADI Min-Max-problem
 - Parameter choice
- 4 Numerical results
 - Parameter comparison for an FDM discretized convection diffusion example
 - Parameter comparison for an FEM discretized convection diffusion example



Problem background in control of parabolic PDEs

Abstract Cauchy problems in Hilbert spaces

We examine optimal control problems for

convection-diffusion-reaction equations

$$\frac{\partial}{\partial t} \mathbf{x} + \nabla \cdot (\mathbf{C}(\mathbf{x}) - \mathbf{K}(\nabla \mathbf{x})) - \mathbf{Q}(\mathbf{x}) = \mathcal{B}\mathbf{u}(t), \quad t \in [0, T_f], \quad (1)$$

on $\Omega \in \mathbb{R}^d$, $d = 1, 2, 3$, with appropriate initial and boundary data. Here \mathbf{C} is the convective part, \mathbf{K} is the diffusive part and \mathbf{Q} is the reactive part.

After variational formulation and linearization these can be written as

abstract Cauchy problems

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t). \quad (2)$$

The state space \mathcal{X} and the control space \mathcal{U} are considered Hilbert spaces (e.g. $\mathcal{X} = H^2(\Omega)$ and $\mathcal{U} = L^2(\Omega)$ for the Dirichlet problem for the heat equation with distributed/point control on the unit square).



Problem background in control of parabolic PDEs

Abstract Cauchy problems in Hilbert spaces

Only certain measurements of \mathbf{x} available as outputs $\mathbf{y} \in \mathcal{Y} \Rightarrow$

output equation

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t). \quad (3)$$

The linear-quadratic regulator (LQR) problem for (1) can then be expressed as

LQR problem for the abstract Cauchy equation

Minimize the **quadratic** cost function

$$J(\mathbf{u}) = \int_0^{\infty} \langle \mathbf{y}, \mathbf{Q}\mathbf{y} \rangle_{\mathcal{Y}} + \langle \mathbf{u}, \mathbf{R}\mathbf{u} \rangle_{\mathcal{U}} dt, \quad (4)$$

with respect to the **linear** constraints (2), (3).



Problem background in control of parabolic PDEs

Abstract Cauchy problems in Hilbert spaces

[GIBSON '78, BALAKRISHNAN '77, LASIECKA/TRIGGIANI '00] discuss that under suitable conditions on $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{Q}$ and \mathbf{R} , \mathbf{u} is given as the

optimal feedback

$$\mathbf{u} = -\mathbf{R}^{-1}\mathbf{B}^*\mathbf{P}_\infty\mathbf{x}(t), \quad (5)$$

where $\mathbf{P}_\infty = \mathbf{P}_\infty^*$ is the stabilizing solution of the

algebraic **operator** Riccati equation

$$0 = \mathbf{A}^*\mathbf{P} + \mathbf{P}\mathbf{A} - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^*\mathbf{P} + \mathbf{C}\mathbf{Q}\mathbf{C} =: \mathfrak{R}(\mathbf{P}). \quad (6)$$

Note: $0 = \mathfrak{R}(\mathbf{P}) \Leftrightarrow \langle \mathbf{v}, \mathfrak{R}(\mathbf{P})\mathbf{w} \rangle = 0 \quad \forall \mathbf{v}, \mathbf{w} \in \text{dom } \mathbf{A}$



Problem background in control of parabolic PDEs

Semidiscretization and finite dimensional systems

Spatial semidiscretization of (1): $\mathcal{X} \rightsquigarrow \mathcal{X}_h$ in (2) and $\mathcal{Y} \rightsquigarrow \mathcal{Y}_h$ in (3) \Rightarrow large scale sparse **ODE** system

$$\dot{x}_h = A_h x_h + B_h u, \quad y_h = C_h x_h$$

with cost function

$$J_h(u) = \int_0^\infty \langle y_h, Q_h y_h \rangle + \langle u, R u \rangle dt,$$

and u is given in feedback form as

$$u = -R^{-1} B_h^T P_h x_h$$

where P_h is the minimal selfadjoint solution of the algebraic **matrix** Riccati equation (ARE)

$$0 = A_h^T P + P A_h - P B_h R^{-1} B_h^T P + C_h Q_h C_h =: \mathfrak{R}_h(P). \quad (7)$$



Problem background in control of parabolic PDEs

Semidiscretization and finite dimensional systems

Spatial semidiscretization of (1): $\mathcal{X} \rightsquigarrow \mathcal{X}_h$ in (2) and $\mathcal{Y} \rightsquigarrow \mathcal{Y}_h$ in (3) \Rightarrow large scale sparse **ODE** system

$$M\dot{x}_h = A_h x_h + B_h u, \quad y_h = C_h x_h$$

with cost function

$$J_h(u) = \int_0^\infty \langle y_h, Q_h y_h \rangle + \langle u, R u \rangle dt,$$

and u is given in feedback form as

$$u = -R^{-1} B_h^T P_h M x_h$$

where P_h is the minimal selfadjoint solution of the algebraic **matrix** Riccati equation (ARE)

$$0 = A_h^T P M + M P A_h - M P B_h R^{-1} B_h^T P M + C_h Q_h C_h =: \mathfrak{R}_h(P). \quad (7)$$



LRCF Newton Method for the ARE

Newton's method for solving the ARE

Newton's iteration for the ARE

$$\mathfrak{X}'_h|_P(N_I) = -\mathfrak{X}_h(P_I), \quad P_{I+1} = P_I + N_I,$$

where the **Frechét derivative** of \mathfrak{X}_h at P is the **Lyapunov operator**

$$\mathfrak{X}'_h|_P : Q \mapsto (A_h - B_h R^{-1} B_h^T P)^T Q + Q (A_h - B_h R^{-1} B_h^T P),$$

can be rewritten as

one iteration step

$$(A_h - B_h R^{-1} B_h^T P_I)^T P_{I+1} + P_{I+1} (A_h - B_h R^{-1} B_h^T P_I) = -C_h^T Q_h C_h - P_I B_h R^{-1} B_h^T P_I$$

i.e. in every Newton step we have to solve a

Lyapunov equation

$$F^T P + P F = -G G^T. \quad (8)$$



LRCF Newton Method for the ARE

Cholesky factor ADI for Lyapunov equations

Peaceman Rachford ADI:

Consider $Au = s$ where $A \in \mathbb{R}^{n \times n}$ spd, $s \in \mathbb{R}^n$. ADI Iteration Idea:
 Decompose $A = H + V$ with $H, V \in \mathbb{R}^{n \times n}$ such that

$$\begin{aligned}(H + pI)v &= r \\ (V + pI)w &= t\end{aligned}$$

can be solved easily.

ADI Iteration

If H, V spd $\Rightarrow \exists p_j, j = 1, 2, \dots, J$ such that

$$\begin{aligned}u_0 &= 0 \\ (H + p_j I)u_{j-\frac{1}{2}} &= (p_j I - V)u_{j-1} + s \\ (V + p_j I)u_j &= (p_j I - H)u_{j-\frac{1}{2}} + s\end{aligned} \tag{9}$$

converges to $u \in \mathbb{R}^n$ solving $Au = s$.



LRCF Newton Method for the ARE

Cholesky factor ADI for Lyapunov equations

Peaceman Rachford ADI:

Consider $Au = s$ where $A \in \mathbb{R}^{n \times n}$ spd, $s \in \mathbb{R}^n$. ADI Iteration Idea:
 Decompose $A = H + V$ with $H, V \in \mathbb{R}^{n \times n}$ such that

$$\begin{aligned}(H + pI)v &= r \\ (V + pI)w &= t\end{aligned}$$

can be solved easily.

ADI Iteration

If H, V spd $\Rightarrow \exists p_j, j = 1, 2, \dots, J$ such that

$$\begin{aligned}u_0 &= 0 \\ (H + p_j I)u_{j-\frac{1}{2}} &= (p_j I - V)u_{j-1} + s \\ (V + p_j I)u_j &= (p_j I - H)u_{j-\frac{1}{2}} + s\end{aligned} \tag{9}$$

converges to $u \in \mathbb{R}^n$ solving $Au = s$.



LRCF Newton Method for the ARE

Cholesky factor ADI for Lyapunov equations

The Lyapunov operator

$$\mathcal{L} : P \mapsto F^T P + P F$$

can be decomposed into the linear operators

$$\mathcal{L}_H : P \mapsto F^T P \quad \mathcal{L}_V : P \mapsto P F.$$

Such that in analogy to (4) we find the

ADI iteration for the Lyapunov equation (8)

$$\begin{aligned} P_0 &= 0 \\ (F^T + p_j I) P_{j-\frac{1}{2}} &= -G G^T - P_{j-1} (F - p_j I) \\ (F^T + p_j I) P_j^T &= -G G^T - P_{j-\frac{1}{2}}^T (F - p_j I) \end{aligned}$$

Can be rewritten to iterate on the low rank Cholesky factors Z_j of P_j to exploit $\text{rk}(P_j) \ll n$. [LI/WHITE 2002; PENZL 1999; BENNER/LI/PENZL 2000]



ADI shift parameter computation

The ADI Min-Max-problem

Optimal parameters solve the

min-max-problem

$$\min_{\{p_j \in \mathbb{R} | j=1, \dots, J\} \subset \mathbb{R}} \max_{\lambda \in \sigma(H), \gamma \in \sigma(V)} \left| \prod_{j=1}^J \frac{(p_j - \lambda)(p_j - \gamma)}{(p_j + \lambda)(p_j + \gamma)} \right|.$$

Remark

- Also known as rational Zolotarev problem since he solved it first on real intervals enclosing the spectra in 1877.
- Another solution for the real case was presented by Wachspress/Jordan in 1963.
- Wachspress and Starke presented different strategies to compute suboptimal shifts for the complex case around 1990.



ADI shift parameter computation

The ADI Min-Max-problem

Optimal parameters solve the

min-max-problem

$$\min_{\{p_j \in \mathbb{R} | j=1, \dots, J\} \subset \mathbb{R}} \max_{\lambda \in \sigma(F)} \left| \prod_{j=1}^J \frac{(p_j - \lambda)}{(p_j + \lambda)} \right|.$$

Remark

- Also known as rational Zolotarev problem since he solved it first on real intervals enclosing the spectra in 1877.
- Another solution for the real case was presented by Wachspress/Jordan in 1963.
- Wachspress and Starke presented different strategies to compute suboptimal shifts for the complex case around 1990.



ADI shift parameter computation

Parameter choice

We will discuss three possible parameter choices here:

- 1 (sub-) optimal parameters [WACHSPRESS: ADI model problem '95] :
 - optimal for real spectra
 - Solve the min-max-problem on an elliptic functions region.
 - Computation needs knowledge of the complete spectrum of F .
 - parameters real if imaginary parts are "small"
- 2 heuristic parameters [PENZL: Lyapack '99]
 - use Ritz values as shifts
 - suboptimal \Rightarrow convergence might be slow
 - in general complex for complex spectra
- 3 semi-optimal parameters: Idea: combine the advantages of these methods
 - use Arnoldi's method to approximate the outer spectrum
 - compute the (sub-)optimal real parameters for this approximation



ADI shift parameter computation

Parameter choice

We will discuss three possible parameter choices here:

- 1 (sub-) optimal parameters [WACHSPRESS: ADI model problem '95] :
 - optimal for real spectra
 - Solve the min-max-problem on an elliptic functions region.
 - Computation needs knowledge of the complete spectrum of F .
 - parameters real if imaginary parts are "small"
- 2 heuristic parameters [PENZL: Lyapack '99]
 - use Ritz values as shifts
 - suboptimal \Rightarrow convergence might be slow
 - in general complex for complex spectra
- 3 semi-optimal parameters: Idea: combine the advantages of these methods
 - use Arnoldi's method to approximate the outer spectrum
 - compute the (sub-)optimal real parameters for this approximation



ADI shift parameter computation

Parameter choice

We will discuss three possible parameter choices here:

- 1 (sub-) optimal parameters [WACHSPRESS: ADI model problem '95] :
 - optimal for real spectra
 - Solve the min-max-problem on an elliptic functions region.
 - Computation needs knowledge of the complete spectrum of F .
 - parameters real if imaginary parts are "small"
- 2 heuristic parameters [PENZL: Lyapack '99]
 - use Ritz values as shifts
 - suboptimal \Rightarrow convergence might be slow
 - in general complex for complex spectra
- 3 semi-optimal parameters: Idea: combine the advantages of these methods
 - use Arnoldi's method to approximate the outer spectrum
 - compute the (sub-)optimal real parameters for this approximation



Numerical results

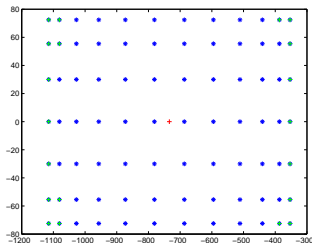
Parameter comparison for an FDM discretized convection diffusion example

$$\dot{x}_t = \Delta x + \begin{pmatrix} 20 \\ 0 \end{pmatrix} \cdot \nabla x - 180x + Bu$$

on $\Omega = \{(0, 1) \times (0, 1)\} \times (0, \infty)$. Finite difference semidiscretization leads to a system of the form

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned}$$

where A is nonsymmetric and its spectrum is complex.



Source: Slicot Working Note 2002-2: A collection of Benchmark examples for model reduction of linear time invariant dynamical systems [CHAHLAOUI/VAN DOOREN 2002]



Numerical results

Parameter comparison for an FDM discretized convection diffusion example

Sizes of the low rank
Cholesky factors:

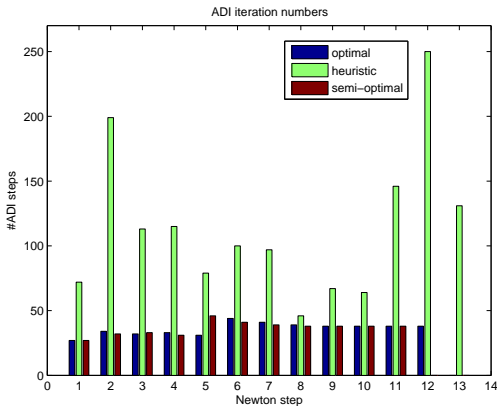
heuristic: $m=262$

optimal/

semi-optimal:

$m=76$

all of them are real in
this case.





Numerical results

Parameter comparison for an FEM discretized convection diffusion example

Chemical reactor inflow heating:

$$\frac{\partial T}{\partial t} - \alpha \Delta T + \nabla T \cdot \begin{pmatrix} v(\xi_2) \\ 0 \end{pmatrix} = 0 \quad \text{in } \Omega$$

$$T = T_0 \quad \text{on } \Gamma_{in}$$

$$\frac{\partial T}{\partial n} = \sigma(T_{ext} - T) \quad \text{on } \Gamma_{heat1} \cup \Gamma_{heat2}$$

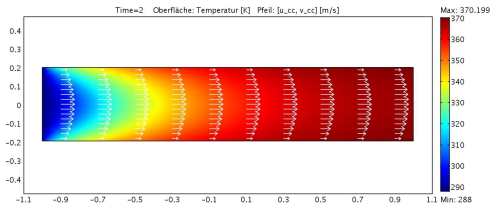
$$\frac{\partial T}{\partial n} = 0 \quad \text{on } \Gamma_{out}$$

FEM discretization leads to:

$$M\dot{x} = Ax + Bu$$

$$y = Cx$$

where: $u \equiv T_{ext}$

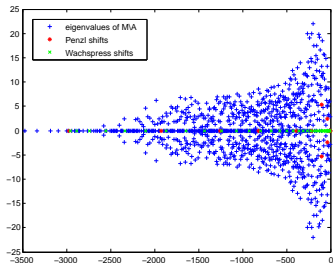




Numerical results

Parameter comparison for an FEM discretized convection diffusion example

Problem size: $n_{state} = 1090$, $n_{in} = 1$, $n_{out} = 3$



semi-optimal convergence in 40-50 ADI steps per Newton step

heuristic complex computations needed \Rightarrow Lyapack fails!

only real Ritz values slow convergence in ADI method (stopped by `max_iter` constant) \Rightarrow no convergence in outer Newton iteration.



Conclusions

- efficient for tested problems
- optimal parameter choice for arbitrary complex spectra is an open problem
- special ordering of the parameters for optimal error reduction should be investigated



Conclusions

- efficient for tested problems
- optimal parameter choice for arbitrary complex spectra is an open problem
- special ordering of the parameters for optimal error reduction should be investigated

Thank you for your attention!