GMRES Acceleration Analysis for a Convection Diffusion Model Problem

Jurjen Duintjer Tebbens

jointly with Zdeněk Strakoš

Institute of Computer Science
Academy of Sciences of the Czech Republic
email: tebbens@cs.cas.cz

Annual GAMM meeting 2005, Lëtzebuerg, April 1, 2005

Outline

1. The convection diffusion model problem
2. Analysis of acceleration of GMRES convergence speed
3. Numerical examples

1. The convection diffusion model problem

We consider a model problem that has been studied by many authors, e.g. [Fischer, Ramage, Sylvester, Wathen - 1999], [Ernst - 2000], [Elman, Ramage - 2001, 2002], [Liesen, Strakoš - 2004]:

$$\begin{aligned} -\nu\nabla^2 u + w\cdot\nabla u &= 0 \quad \text{in} \quad \Omega = (0,1)\times(0,1), \\ u &= g \quad \text{on} \quad \partial\Omega, \end{aligned}$$

- ν: scalar diffusion coefficient
- $w = [0,1]^T$: convective velocity field
- g : discontinuous inflow boundary conditions

$$\begin{aligned} u(x,0) = u(1,y) &= 1 \quad \text{for} \quad \frac{1}{2} \le x \le 1 \quad \text{and} \quad 0 \le y < 1, \\ u(x,y) &= 0 \quad \text{elsewhere on } \partial\Omega \end{aligned}$$

We use

- SUPG discretization (see for example [Brooks, Hughes - 1990])
- Discrete sine transformation [Liesen, Strakoš - 2004]
- A 35×35 grid (results also hold for other grid sizes)
- The zero initial guess in GMRES processes

This yields a linear system with a $35^2 \times 35^2$ block diagonal matrix $\mathbf{A}$ with tridiagonal Toeplitz blocks of dimension 35×35 :

$$\mathbf{A}\,x = \begin{pmatrix} \mathbf{T}_1 & & & & \\ & \mathbf{T}_2 & & & \\ & & \ddots & & \\ & & & \mathbf{T}_{34} & \\ & & & & \mathbf{T}_{35} \end{pmatrix} x = \begin{pmatrix} b^{(1)} \\ b^{(2)} \\ \vdots \\ b^{(34)} \\ b^{(35)} \end{pmatrix} = b,$$

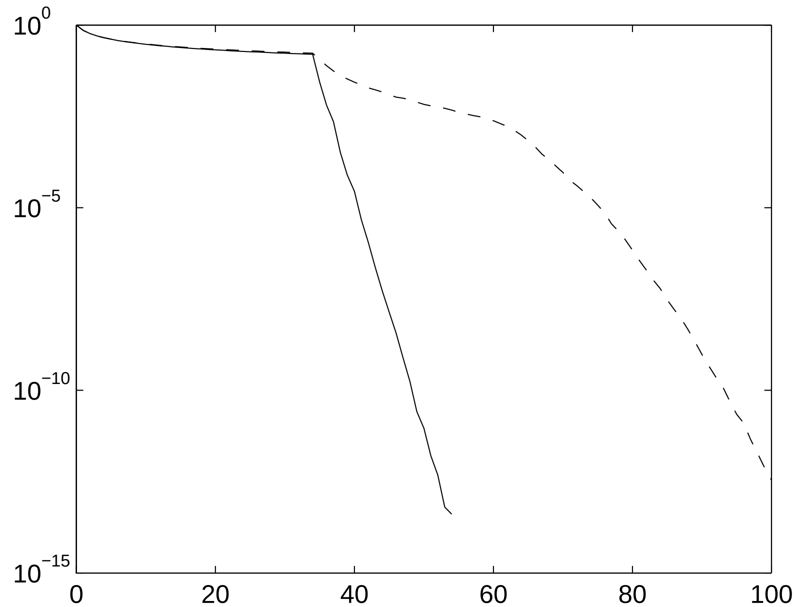
where b is partitioned according to the block-diagonal structure and

$$\mathbf{T}_i = \begin{pmatrix} \gamma_i\tau_i & \gamma_i\zeta_i & & & \\ \gamma_i & \gamma_i\tau_i & \gamma_i\zeta_i & & \\ & \ddots & \ddots & \ddots & \\ & & \gamma_i & \gamma_i\tau_i & \gamma_i\zeta_i \\ & & & \gamma_i & \gamma_i\tau_i \end{pmatrix} = \gamma_i \left(\mathbf{S} + \tau_i \mathbf{I} + \zeta_i \mathbf{S}^T \right),$$

and where $\mathbf{I}$ is the identity matrix and $\mathbf{S}$ is the downshift matrix $\mathbf{S} = (e_2, \ldots, e_{35}, 0)$.

- For very small ν the operator is highly non-normal [Reddy, Trefethen - 1994]
- We consider two cases: $\nu = 0.005$ and $\nu = 0.00005$
- Values $\gamma_i\tau_i$ are close to each other and the values $|\gamma_i\zeta_i| << 1$

GMRES norm reduction for the discretized convection diffusion model problem



Dashed line: $\nu = 0.00005$ - Solid line: $\nu = 0.005$.

Slow convergence during the first 35 steps is quantified by a lower bound that considers the small problems of dimension 35

$$\mathbf{T}_i\, x^{(i)} = b^{(i)}, \qquad i = 1, 2, \ldots, 35,$$

and by analysis of GMRES behavior for the small problems, see [Liesen, Strakoš - 2004].

We cannot use this strategy to explain **accelerated convergence speed after step 35** because

- The strategy is useless after the 35th iteration, the bound should couple the influence of the individual blocks
- Acceleration of convergence needs to be described be an upper bound

We can, however, exploit the analysis for the small problems:

2. Analysis of acceleration of GMRES convergence speed

We generalize a result for Toeplitz matrices proved in [Liesen, Strakoš - 2004] to explain the initial phase of convergence:

Theorem: *Let us consider the residual vectors generated by GMRES applied to* $\mathbf{C}y = d$ *and assume that* $[d, \mathbf{C}d, \ldots, \mathbf{C}^k d]$ *has full column rank. If* $\mathbf{C} = \mathbf{B} + \mathbf{U}$ *with* $\mathbf{UB} = \mathbf{BU}$ *and* $\mathbf{U}d = \tau d$*, then*

$$\begin{aligned} r_k^T &= \|r_k\|^2 [1, -\tau, \ldots, (-\tau)^k] \cdot \left[d, \mathbf{B}d, \ldots, \mathbf{B}^k d\right]^+, \\ \|r_k\| &= \frac{1}{\|[1, -\tau, \ldots, (-\tau)^k] \cdot [d, \mathbf{B}d, \ldots, \mathbf{B}^k d]^+\|}. \end{aligned}$$

We put $\mathbf{C} \equiv \mathbf{A}$, $d \equiv b$ and $\mathbf{U} \equiv \tau \mathbf{I}$, hence $\mathbf{A} = \mathbf{B} + \tau \mathbf{I}$.
With B we can recognize a sudden change at $k = 35$**:**

Recall that we consider a block diagonal system matrix with blocks

$$\mathbf{T}_i = \begin{pmatrix} \gamma_i\tau_i & \gamma_i\zeta_i & & & \\ \gamma_i & \gamma_i\tau_i & \gamma_i\zeta_i & & \\ & \ddots & \ddots & \ddots & \\ & & \gamma_i & \gamma_i\tau_i & \gamma_i\zeta_i \\ & & & \gamma_i & \gamma_i\tau_i \end{pmatrix}, \qquad i = 1, \ldots 35.$$

All $|\gamma_i\zeta_i| << 1$ and all $\gamma_i\tau_i$ are close. Hence for a diagonal translation τ close to all $\gamma_i\tau_i$, the matrix $\mathbf{B} = \mathbf{A} - \tau\mathbf{I}$ has blocks

$$\tilde{\mathbf{T}}_i = \mathbf{T}_i - \tau\mathbf{I} \approx \begin{pmatrix} 0 & 0 & & & \\ \gamma_i & 0 & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & \gamma_i & 0 & 0 \\ & & & \gamma_i & 0 \end{pmatrix}, \qquad i = 1, \ldots 35.$$

Because of

$$\mathbf{B}b = \begin{pmatrix} \tilde{\mathbf{T}}_1 & & & & \\ & \tilde{\mathbf{T}}_2 & & & \\ & & \ddots & & \\ & & & \tilde{\mathbf{T}}_{34} & \\ & & & & \tilde{\mathbf{T}}_{35} \end{pmatrix} \begin{pmatrix} b^{(1)} \\ b^{(2)} \\ \vdots \\ b^{(34)} \\ b^{(35)} \end{pmatrix}, \quad \tilde{\mathbf{T}}_i \approx \begin{pmatrix} 0 & 0 & & & \\ \gamma_i & 0 & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & \gamma_i & 0 & 0 \\ & & & \gamma_i & 0 \end{pmatrix},$$

application of $\mathbf{B}$ to b shifts down one position all entries of all $b^{(i)}$. Hence $\mathbf{B}^{35}b \approx \mathbf{0}$. Now our identity

$$\|r_k\| = \frac{1}{\| [1, -\tau, \ldots, (-\tau)^k] \, [b, \mathbf{B}b, \ldots, \mathbf{B}^k b]^+ \|}$$

yields the upper bound

$$\|r_k\| \quad \leq \quad \frac{\|\mathbf{B}^k b\|}{|\tau|^k},$$

which quantifies the acceleration of convergence.

3. Numerical examples

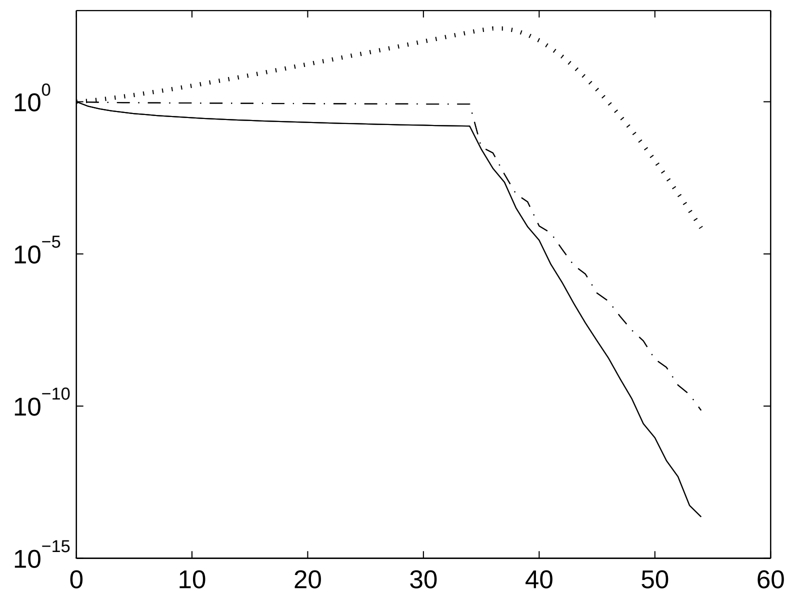
For $\nu = 0.005$:

- The diagonal values $\gamma_i \tau_i$ are particularly close
  $\Rightarrow \quad \tau = (\sum_{i=1}^{35} \gamma_i \tau_i)/35$

- The norms $\|b^{(1)}\|, \|b^{(2)}\| > \|b^{(i)}\|$ for $i \geq 3$
  $\Rightarrow \tau = (\gamma_1 \tau_1 + \gamma_2 \tau_2)/2$

The choices and $\tau = (\sum_{i=1}^{35} \gamma_i \tau_i)/35$ and $\tau = (\gamma_1 \tau_1 + \gamma_2 \tau_2)/2$ yield

Convergence curve and upper bounds for $\nu = 0.005$



Solid line: GMRES residual norm reduction

Dash-dotted line: Upper bound with $\tau = (\gamma_1 \tau_1 + \gamma_2 \tau_2)/2$

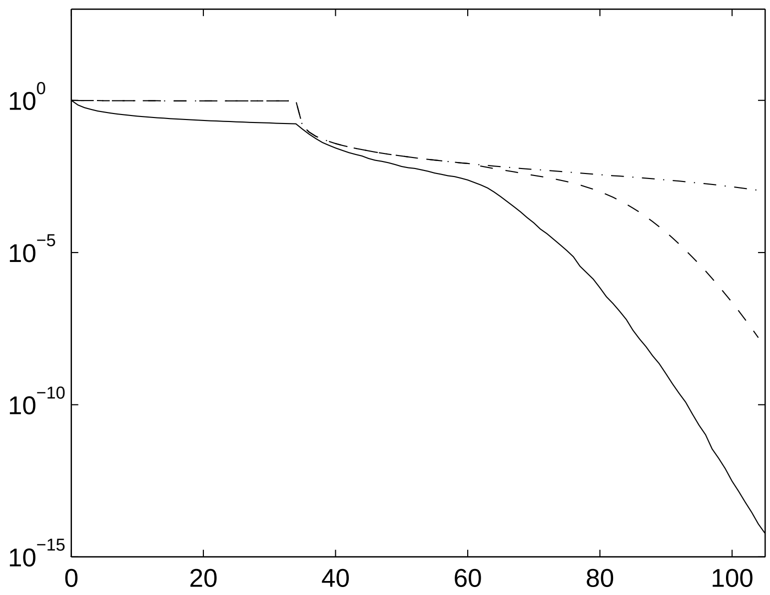
Dotted line: Upper bound with $\tau = (\sum_{i=1}^{35} \gamma_i \tau_i)/35$

For $\nu = 0.00005$:

- The upper-diagonal values $|\gamma_i \zeta_i|$ are particularly small: a highly non-normal linear system
- The diagonal values $\gamma_i \tau_i$ are less close than for $\nu = 0.005$.
- Again, the norms $\|b^{(1)}\|, \|b^{(2)}\| > \|b^{(i)}\|$ for $i \geq 3$

The choice $\tau = (\gamma_1 \tau_1 + \gamma_2 \tau_2)/2$ yields the dash-dotted line in the following graph:

Convergence curve and upper bounds for $\nu = 0.00005$



Solid line: GMRES residual norm reduction

Dash-dotted line: Upper bound with $\tau = (\gamma_1\tau_1 + \gamma_2\tau_2)/2$

Dashed line: Upper bound with adaptive translation

Villmols merci for your attention.

This work is supported by the Program Information Society under project 1ET400300415.