Università di Bologna

Solving ill-posed linear systems with GMRES

V. Simoncini

Università di Bologna (Italy)
valeria.simoncini@unibo.it

Joint work with Lars Eldèn, Linköping University, Sweden

The problem

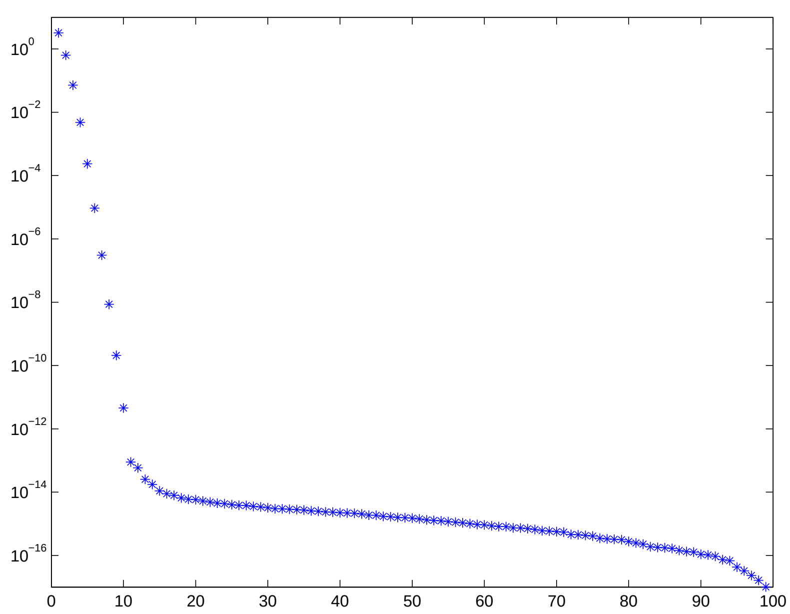
Solve $\qquad Ax = b \qquad$ with GMRES assuming

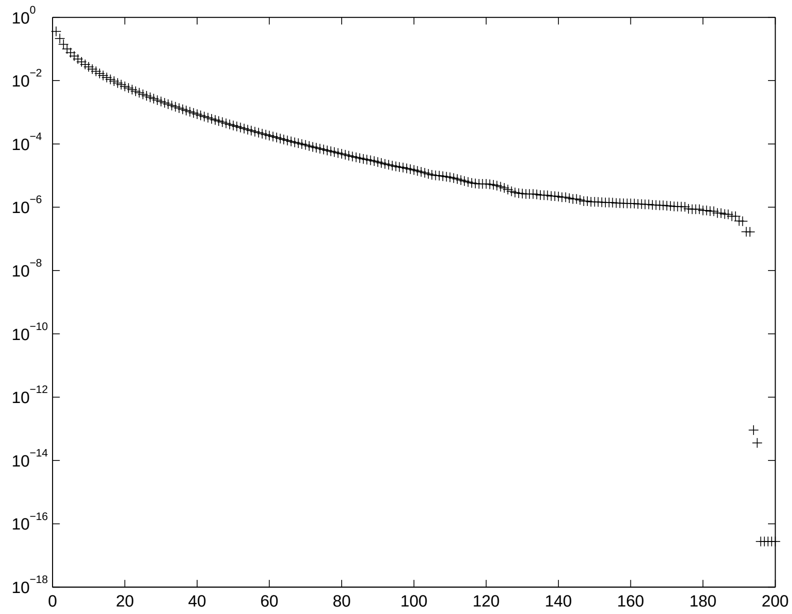
- $A \in \mathbb{C}^{n \times n}$ large, nonsymmetric
- A almost singular (or *numerically singular*)
- b noise-perturbed version of "consistent" rhs

The problem

Solve $\qquad Ax = b \qquad$ with GMRES assuming

- $A \in \mathbb{C}^{n \times n}$ large, nonsymmetric
- A almost singular (or *numerically singular*)
- b noise-perturbed version of "consistent" rhs





Typical singular value distributions

Two different perspectives

For the analysis of the GMRES behavior:

- Ill-posed problem framework suggests using sing.values

  (e.g., Jensen, Hansen 2006, Hanke, Nagy, Plemmons 1993, Brianzi, Favati, Menchi, Romani 2006)

- The Krylov subspace setting suggests using spectral information

  (e.g., Calvetti, Lewis, Reichel 2002)

Some preliminary considerations

Assume A is **exactly** singular: $\qquad Ax = b$

GMRES: Given $x_0 \in \mathbb{C}^n$, and $r_0 = b - Ax_0$,

Find $x_k \in x_0 + K_k(A, r_0)$ such that $x_k = \arg \min_{x \in x_0 + K_k(A,r_0)} \|b - Ax\|$

Brown, Walker 1997, Hayami, Sugihara 2011:

GMRES determines a least squares solution x_* of a singular system $Ax = b$, for all b and starting approximations x_0, without breakdown, if and only if

$$\mathcal{N}(\mathcal{A}) = \mathcal{N}(\mathcal{A}^*)$$

Furthermore, if the system is consistent and $x_0 \in \mathcal{R}(\mathcal{A})$, then x_* is a minimum norm solution.

Some preliminary considerations

If $Ax = b$ is written as

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix} = \begin{bmatrix} b^{(1)} \\ b^{(2)} \end{bmatrix}, \qquad A_{11} \in \mathbb{C}^{m \times m}$$

then

$$\mathcal{N}(A) = \mathcal{N}(A^*) \qquad \Leftrightarrow \qquad A_{12} = 0$$

and

$$\text{consistency} \qquad \Leftrightarrow \qquad b^{(2)} = 0$$

Clearly, $A_{12} = 0$ corresponds to solving $A_{11}x^{(1)} = b^{(1)}$

In practice: $A_{12} \approx 0$ and $b^{(2)} \approx 0$ (but nonzero)

Our setting

Consider a preconditioned least squares problem

$$\min_y \|(AM_m^\dagger)y - b\|$$

Discrepancy principle:

Determine approx $\hat{x}$ with residual $\|A\hat{x} - b\| \approx \delta$

(δ prespecified, measure of data noise level)

$\Rightarrow$ *right* preconditioning

$\Rightarrow$ *stopping criterion:* $\delta = 1.1 \cdot (\text{data noise}) \cdot / \|b\|$

Our setting

Consider a preconditioned least squares problem

$$\min_y \|(AM_m^\dagger)y - b\|$$

Discrepancy principle:

Determine approx $\hat{x}$ with residual $\|A\hat{x} - b\| \approx \delta$

(δ prespecified, measure of data noise level)

$\Rightarrow$ *right* preconditioning

$\Rightarrow$ *stopping criterion:* $\delta = 1.1 \cdot (\mathrm{data\,noise}) \cdot /\|b\|$

* Schur decomposition $AM_m^\dagger = UBU^*$ with U unitary. Then

$$\min_{\begin{bmatrix} d^{(1)} \\ d^{(2)} \end{bmatrix}} \left\| \underbrace{\begin{bmatrix} L_1 & G \\ 0 & L_2 \end{bmatrix}}_{B} \begin{bmatrix} d^{(1)} \\ d^{(2)} \end{bmatrix} - \begin{bmatrix} c^{(1)} \\ c^{(2)} \end{bmatrix} \right\| \qquad c = U^*b, \; d = U^*y$$

Our setting

$$\min_{\begin{bmatrix} d^{(1)} \\ d^{(2)} \end{bmatrix}} \left\| \begin{bmatrix} L_1 & G \\ 0 & L_2 \end{bmatrix} \begin{bmatrix} d^{(1)} \\ d^{(2)} \end{bmatrix} - \begin{bmatrix} c^{(1)} \\ c^{(2)} \end{bmatrix} \right\|$$

with

$$|\lambda_{\min}(L_1)| \gg |\lambda_{\max}(L_2)|, \qquad \|c^{(1)}\| \gg \|c^{(2)}\| = \delta$$

$\Rightarrow L_2$ and $c^{(2)}$ correspond to "noise"

Moreover,

- $\|L_1^{-1}\|$ moderate; $\|G\|$ moderate (high non-normality excluded)

Our setting

$$\min_{\begin{bmatrix} d^{(1)} \\ d^{(2)} \end{bmatrix}} \left\| \begin{bmatrix} L_1 & G \\ 0 & L_2 \end{bmatrix} \begin{bmatrix} d^{(1)} \\ d^{(2)} \end{bmatrix} - \begin{bmatrix} c^{(1)} \\ c^{(2)} \end{bmatrix} \right\| \qquad (1)$$

with

$$|\lambda_{\min}(L_1)| \gg |\lambda_{\max}(L_2)|, \qquad \|c^{(1)}\| \gg \|c^{(2)}\| = \delta$$

$\Rightarrow$ L_2 and $c^{(2)}$ correspond to "noise"

Moreover,

- $\|L_1^{-1}\|$ moderate; $\|G\|$ moderate (high non-normality excluded)

The problem above may be viewed as a perturbation of

$$\min_{\begin{bmatrix} d^{(1)} \\ d^{(2)} \end{bmatrix}} \left\| \begin{bmatrix} L_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} d^{(1)} \\ d^{(2)} \end{bmatrix} - \begin{bmatrix} c^{(1)} \\ c^{(2)} \end{bmatrix} \right\| \qquad (2)$$

Is it possible to "solve" (1) as efficiently as we would do with (2) ?

Spectral decomposition

$$B = \begin{bmatrix} L_1 & G \\ 0 & L_2 \end{bmatrix} = XB_0X^{-1} =: [X_1, X_2] \begin{bmatrix} L_1 & 0 \\ 0 & L_2 \end{bmatrix} \begin{bmatrix} Y_1^* \\ Y_2^* \end{bmatrix}$$

where $[Y_1, Y_2]^* = [X_1, X_2]^{-1}$, and

$$[X_1, X_2] = \begin{bmatrix} I & P \\ 0 & I \end{bmatrix}, \qquad [Y_1, Y_2] = \begin{bmatrix} I & 0 \\ -P^* & I \end{bmatrix},$$

and P is the unique soln of the Sylvester eqn $L_1P - PL_2 = -G$

Note that

$$\|X_2\| \leq 1 + \|P\|, \quad \|Y_1\| \leq 1 + \|P\|, \quad \text{where} \quad \|P\| \leq \frac{\|G\|}{\mathrm{sep}(L_1, L_2)}$$

It also follows that: $\qquad \|X\| \leq 1 + \|P\|$

The residual polynomial

$$Bd = c$$

For any polynomial φ_m,

$$\varphi_m(B)c = [X_1, X_2] \begin{bmatrix} \varphi_m(L_1)Y_1^* c \\ \varphi_m(L_2)Y_2^* c \end{bmatrix} = X_1\varphi_m(L_1)Y_1^* c + X_2\varphi_m(L_2) \underbrace{Y_2^* c}_{=c^{(2)}},$$

so that

$$\|\varphi_m(B)c\| \le \|\varphi_m(L_1)Y_1^* c\| + \|X_2\varphi_m(L_2)c^{(2)}\|$$

An explanatory example

`wing` example from Regularization Matlab Toolbox (Hansen, 1994-2007)

(Discretization of a 1st kind Fredholm integral eqn, discontinuous soln)

spec(A):

$$3.74 \cdot 10^{-1},\ -2.55 \cdot 10^{-2},\ 7.65 \cdot 10^{-4},\ -1.48 \cdot 10^{-5},$$
$$2.13 \cdot 10^{-7},\ -2.45 \cdot 10^{-9}, 2.33 \cdot 10^{-11},\ -1.89 \cdot 10^{-13},\ 1.32 \cdot 10^{-15},\ \ldots$$

perturbed rhs: $b = b_e + \varepsilon p$ (p with `randn` entries and $\|p\| = 1$)

L_1: corresponds to the abs largest six eigenvalues

$\|G\| = 2.29 \cdot 10^{-5}$, $\|P\| = 10.02$

For $\varepsilon = 10^{-7}$: $\|Y_1^* c\| = 1$ and $\|Y_2^* c\| = 6.7 \cdot 10^{-7}$

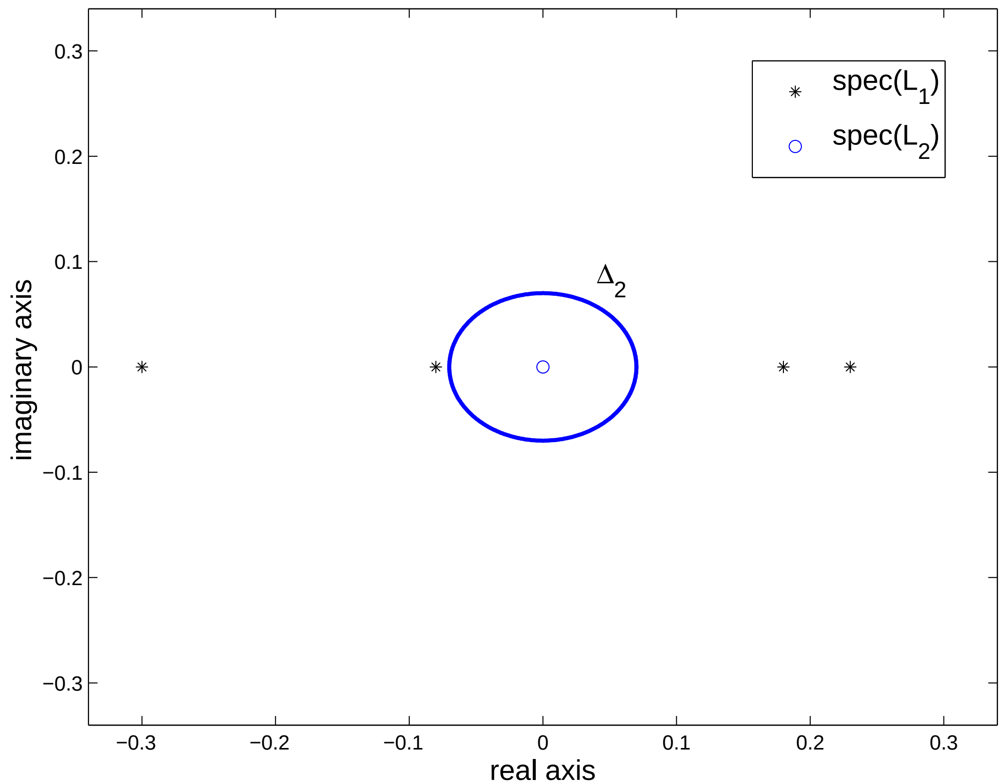
For $\varepsilon = 10^{-5}$: $\|Y_2^* c\| = 6.49 \cdot 10^{-5}$

GMRES residual bounds

m_*: grade of L_1 with respect to $Y_1^* c$ $\quad$ ($\exists \phi_{m_*}$ s.t. $\phi_{m_*}(L_1) Y_1^* c = 0$)

r_k: GMRES residual after k iterations

Δ_2: circle centered at the origin and having radius ρ s.t. $\mathrm{spec}(L_2) \subset \Delta_2$



GMRES residual bounds

m_*: grade of L_1 with respect to $Y_1^* c$ $\qquad$ ($\exists \phi_{m_*}$ s.t. $\phi_{m_*}(L_1) Y_1^* c = 0$)

r_k: GMRES residual after k iterations

Δ_2: circle centered at the origin and having radius ρ s.t. $\mathrm{spec}(L_2) \subset \Delta_2$

i) If $k < m_*$, let $s_k^{(1)} = \phi_k(L_1) Y_1^* c$ be the GMRES residual associated with $L_1 z = Y_1^* c$. Then

$$\|r_k\| \leq \|s_k^{(1)}\| + \|X_2\| \gamma_k \tau, \qquad \tau = \rho \max_{z \in \Delta_2} \|(zI - L_2)^{-1} c^{(2)}\|,$$

where $\gamma_k = \max_{z \in \Delta_2} \prod_{i=1}^{k} |\theta_i - z| / |\theta_i|$ and θ_i are the roots of ϕ_k

GMRES residual bounds

m_*: grade of L_1 with respect to $Y_1^* c$ $\qquad$ $(\exists \phi_{m_*}$ s.t. $\phi_{m_*}(L_1) Y_1^* c = 0)$

r_k: GMRES residual after k iterations

Δ_2: circle centered at the origin and having radius ρ s.t. $\mathrm{spec}(L_2) \subset \Delta_2$

i) If $k < m_*$, let $s_k^{(1)} = \phi_k(L_1) Y_1^* c$ be the GMRES residual associated with $L_1 z = Y_1^* c$. Then

$$\|r_k\| \le \|s_k^{(1)}\| + \|X_2\| \gamma_k \tau, \qquad \tau = \rho \max_{z \in \Delta_2} \|(zI - L_2)^{-1} c^{(2)}\|,$$

where $\gamma_k = \max_{z \in \Delta_2} \prod_{i=1}^{k} |\theta_i - z| / |\theta_i|$ and θ_i are the roots of ϕ_k

Wing data. $m_* = 6$, $\|G\| = 2.29 \cdot 10^{-5}$ and $\|P\| = 10.02$. radius $\rho = 2 \cdot 10^{-9}$

ε	k	$\|s_k^{(1)}\|$	$\|X_2\|\gamma_k\tau$	Sum	$\|r_k\|$
10^{-7}	2	1.640e-03	6.770e-06	1.647e-03	1.640e-03
	3	3.594e-05	6.770e-06	4.271e-05	3.573e-05
10^{-5}	2	1.621e-03	6.770e-04	2.298e-03	1.640e-03
	3	6.568e-05	6.770e-04	7.427e-04	7.568e-05

GMRES residual bounds

m_*: grade of L_1 with respect to $Y_1^* c$ $\qquad (\exists \phi_{m_*}$ s.t. $\phi_{m_*}(L_1) Y_1^* c = 0)$

r_k: GMRES residual after k iterations

Δ_2: circle centered at the origin and having radius ρ s.t. $\mathrm{spec}(L_2) \subset \Delta_2$

ii) If $k = m_* + j$, $j \geq 0$, let $s_j^{(2)} = \varphi_j(L_2) c^{(2)}$ be the GMRES residual associated with $L_2 z = c^{(2)}$ after j iterations (note that $\|s_j^{(2)}\| \leq \|c^{(2)}\|$). Then

$$\|r_k\| \leq \rho \gamma_{m_*} \|s_j^{(2)}\| \|X_2\| \max_{z \in \Delta_2} \|(zI - L_2)^{-1}\|$$

where $\gamma_{m_*} = \max_{z \in \Delta_2} \prod_{i=1}^{m_*} |\theta_i - z| / |\theta_i|$ and θ_i are the roots of the grade polyn of L_1

GMRES residual bounds

m_*: grade of L_1 with respect to $Y_1^* c$ $\quad$ $(\exists \phi_{m_*}$ s.t. $\phi_{m_*}(L_1) Y_1^* c = 0)$

r_k: GMRES residual after k iterations

Δ_2: circle centered at the origin and having radius ρ s.t. $\mathrm{spec}(L_2) \subset \Delta_2$

ii) If $k = m_* + j$, $j \geq 0$, let $s_j^{(2)} = \varphi_j(L_2) c^{(2)}$ be the GMRES residual associated with $L_2 z = c^{(2)}$ after j iterations (note that $\|s_j^{(2)}\| \leq \|c^{(2)}\|$). Then

$$\|r_k\| \leq \rho \gamma_{m_*} \|s_j^{(2)}\| \|X_2\| \max_{z \in \Delta_2} \|(zI - L_2)^{-1}\|$$

where $\gamma_{m_*} = \max_{z \in \Delta_2} \prod_{i=1}^{m_*} |\theta_i - z| / |\theta_i|$ and θ_i are the roots of the grade polyn of L_1

Wing data. $m_* = 6$, $\|G\| = 2.29 \cdot 10^{-5}$ and $\|P\| = 10.02$. radius $\rho = 2 \cdot 10^{-9}$

ε	k	$\|s_k^{(1)}\|$	$\|X_2\| \gamma_k \tau$	Bound	$\|r_k\|$
10^{-7}	10			6.712e-06	6.311e-07
10^{-5}	10			6.442e-04	6.308e-05

Application to singular preconditioning

Let $M_m^\dagger \in \mathbb{C}^{n\times n}$ be a rank-m approximation of A^{-1}

Then $\mathrm{rank}(AM_m^\dagger) = m$,

$$B = U^*(AM_m^\dagger)U = \begin{bmatrix} L_1 & G \\ 0 & 0 \end{bmatrix}$$

and the least squares problem reads

$$\min_d \{\|[L_1, G]d - c^{(1)}\|^2 + \|c^{(2)}\|^2\}, \qquad c = U^*b = \begin{bmatrix} c^{(1)} \\ c^{(2)} \end{bmatrix}$$

A preconditioned 2D ill-posed elliptic problem

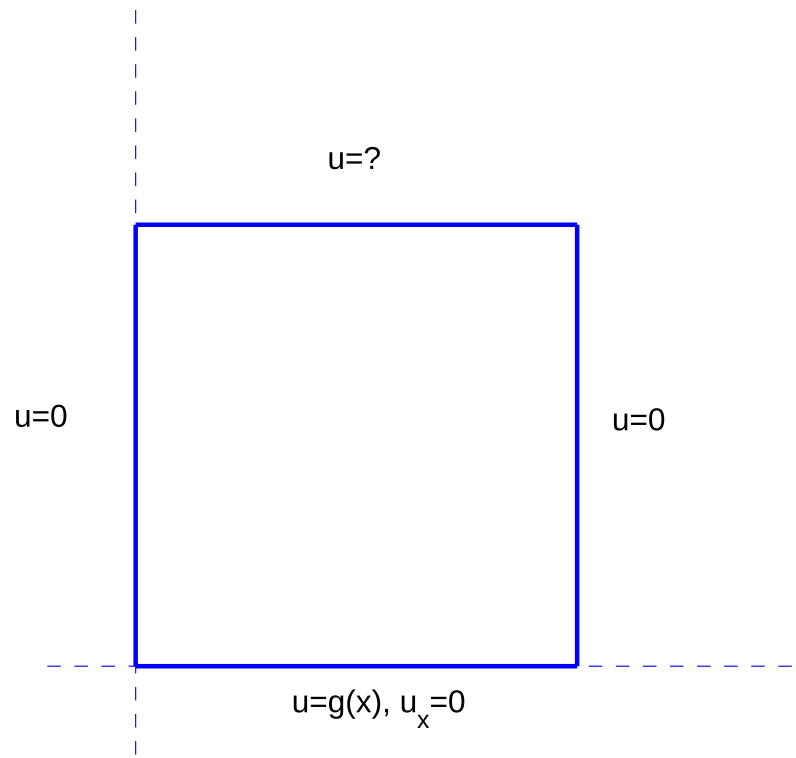
$$(\beta(y)u_y)_y + (\alpha u_x)_x + \gamma u_x = 0$$

with $\alpha = 1$, $\gamma = 2$, and

$$\beta(y) = \begin{cases} 50, & 0 \le y \le 0.5, \\ 8 & 0.5 < y \le 1. \end{cases}$$

`randn` perturbation s.t.

$\|g - g_{pert}\| / \|g\| \approx 1.8 \cdot 10^{-3}$



A preconditioned 2D ill-posed elliptic problem

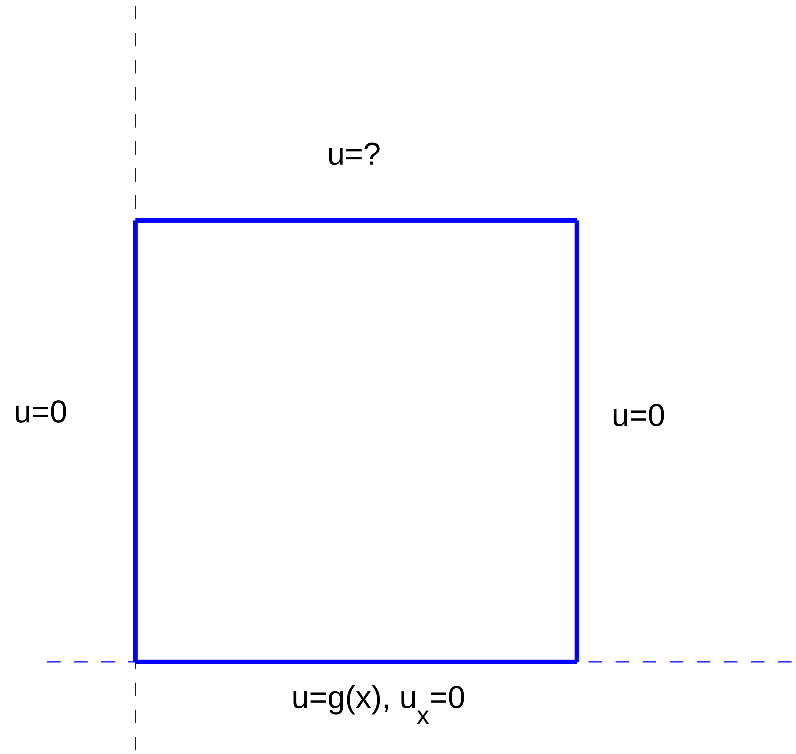
$$(\beta(y)u_y)_y + (\alpha u_x)_x + \gamma u_x = 0$$

with $\alpha = 1$, $\gamma = 2$, and

$$\beta(y) = \begin{cases} 50, & 0 \le y \le 0.5, \\ 8 & 0.5 < y \le 1. \end{cases}$$

`randn` perturbation s.t.

$\|g - g_{pert}\|/\|g\| \approx 1.8 \cdot 10^{-3}$



Rank-m preconditioner $M_m^\dagger$ ($m = 9$):

approx to the exact soln operator: $f_0 = \cosh\left((\frac{1}{\beta_0}L_m)^{\frac{1}{2}}\right)$, $\beta_0 = \overline{\beta(y)}$

(Eldèn, Simoncini, 2009)

The preconditioned problem solved by GMRES

$$\min_y \|(AM_m^\dagger)y - g\|, \qquad M_m^\dagger = \cosh\left((\frac{1}{\beta_0} L_m)^{\frac{1}{2}} \right)$$

* Operation Av: solve the well-posed problem with b.c. $u(x,1) = v(x)$ replacing $u(x,0) = g(x)$ (dim. 10000)

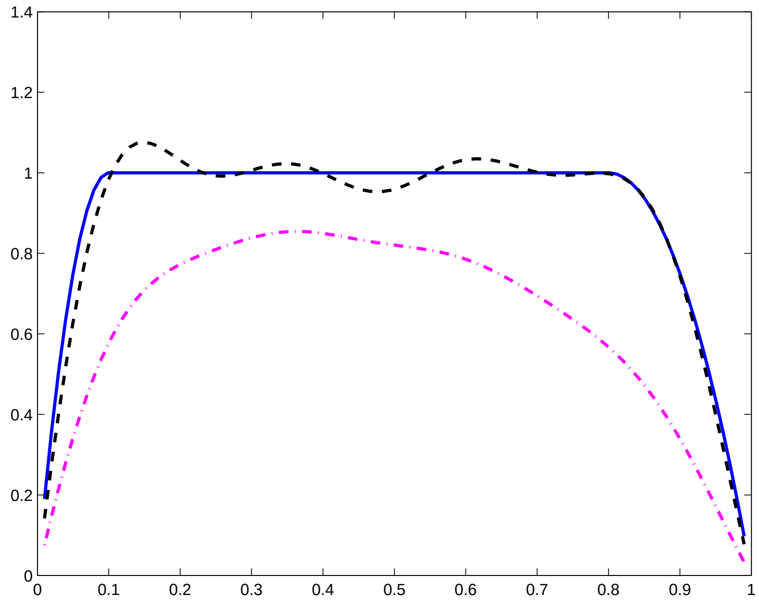
-

-.

The preconditioned problem solved by GMRES

$$\min_y \|(AM_m^\dagger)y - g\|, \qquad M_m^\dagger = \cosh\left((\frac{1}{\beta_0} L_m)^{\frac{1}{2}} \right)$$

* Operation Av: solve the well-posed problem with b.c. $u(x,1) = v(x)$ replacing $u(x,0) = g(x)$ (dim. 10000)

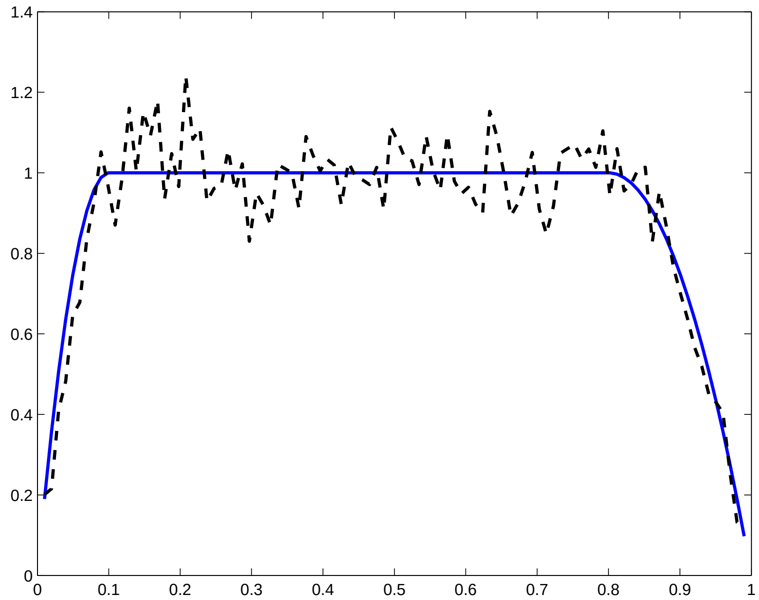


Exact soln ('-'). 3 steps Prec'd GMRES ('- -'). Operation $M_m^\dagger g$ ('-.')

The preconditioned problem solved by GMRES

$$\min_y \|(AM_m^\dagger)y - g\|, \qquad M_m^\dagger = \cosh\left((\frac{1}{\beta_0} L_m)^{\frac{1}{2}} \right)$$

* Operation Av: solve the well-posed problem with b.c. $u(x,1) = v(x)$ replacing $u(x,0) = g(x)$ (dim. 10000)



Exact solution (solid line). Unprec'd GMRES with smallest error (5 steps, dashed)

Conclusions

- Good understanding of GMRES behavior for almost singular systems stemming from ill-posed problems
- Applicability to singular preconditioning

Error bounds (not shown) seem to imply:

- The singular preconditioner acts as regularization operator
- For the residual $r_k = b - Ax_k$, the quantity $\|A^* r_k\|$ could be monitored together with $\|r_k\|$

Reference:

L. Eldèn and V. Simoncini, *Solving Ill-posed Linear Systems with GMRES and a Singular Preconditioner*, SIMAX, v.33 (4), 2012.